Contents

1. Context

1.1 Overview

1.1.1 This note provides guidance for departments building on chapter 4 (pages 28-30) in the Sourcing Playbook on the use of different testing approaches, including pilots, to improve the success of outsourcing and insourcing projects by providing insight and evidence into what works. Testing and piloting should be considered part of a wider strategic delivery model assessment (previously known as 'Make or Buy' decision), overall commercial strategy and project business case.

1.1.2 This guidance is a collection of best practice to be used by practitioners, including commercial and project delivery professionals, in developing effective testing and pilot approaches, including how to avoid common pitfalls and legal issues to be considered. It should not be considered as legal advice or a replacement for departments to ensure suitably qualified and experienced practitioners run robust procurements with appropriate legal consultation.

1.2 Contact

1.2.1 For departments considering how testing approaches (including pilots) can be included in complex projects you should consult the Cabinet Office. The Sourcing Programme (sourcing.programme@cabinetoffice.gov.uk) provides support to complex outsourcing projects in collaboration with the Complex Transactions Team (complextransactions@cabinetoffice.gov.uk), and other Cabinet Office teams.

2. What is testing and piloting?

2.1 Defining ‘piloting’

2.1.1 Across government, the concept of ‘piloting’ new services has come to mean many different things. This guidance note refers to a wide range of mechanisms for testing services before they are fully implemented, of which ‘pilots’ are only one form. These include:

- Policy Trials
- Proofs of Concept
- Scoping Phases
- Test and Learns
- Pilots

2.1.2 This guidance also considers the use of Agile approaches and Innovation Partnerships. Whilst these are not, in themselves, testing mechanisms, testing of the services and/or products being developed using these approaches needs to be fully embedded within the planning, procurement and implementation design for any “Agile” project or Innovation Partnership.

2.1.3 In many instances it will be appropriate for departments to use one or more testing approaches at earlier stages of project development, with the pilot being the final testing stage prior to a full scale rollout of services. Figure 1 provides a high-level overview of the differences between approaches, with further detail in Table 1 with definitions and key differences between each of the testing approaches. Appendix 2 contains some fictionalised examples of the different approaches to show how these might be used in practice.

2.1.4 Early testing enables departments to understand the viability of a project or outcome at its various stages of development. This allows the department the opportunity to change the course of action at an early stage, limiting cost and time where it becomes apparent that the project will not deliver the required outcome. It is essential to test against user needs to ensure that the final delivery model is fit for purpose. Additional information has been provided on pilots throughout this guidance note as experience has shown us the significant benefits of testing a preferred approach for complex outsourcing projects on a smaller scale to identify learnings for the subsequent implementation.

2.1.5 For any type of testing involving engagement with suppliers, the test and any subsequent procurement must comply with legal requirements (procurement and state aid law) and the contractual relationship with the supplier clearly set out. Legal advice may be required if an external supplier is engaged to provide testing services in order to ensure that they do not gain an unfair advantage for any future procurement of services. Further guidance on legal requirements is provided in section 9.

Figure 1: Overview of differences between testing approaches

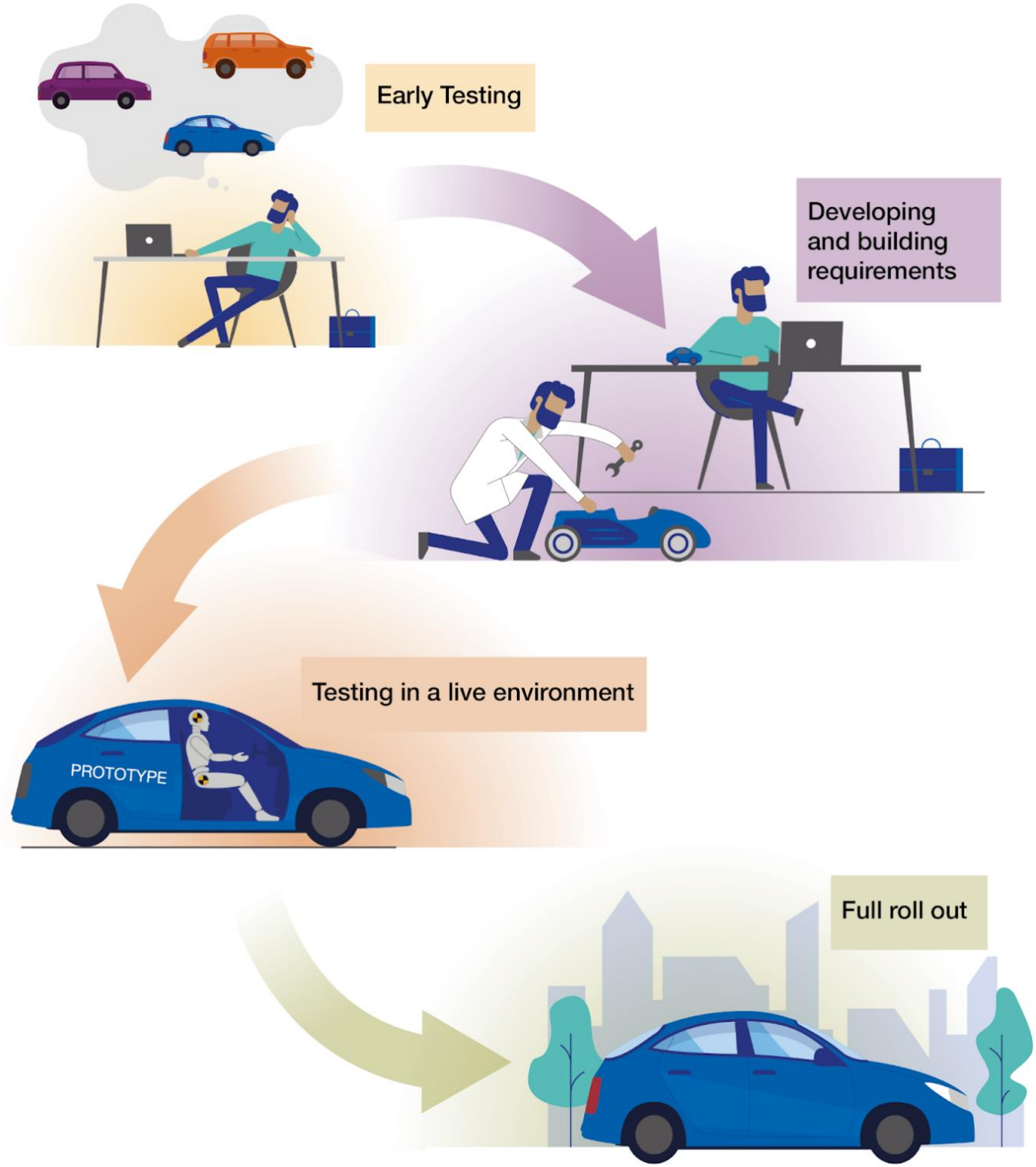


Table 1: Approaches to testing services by project phase

Early testing		Developing and building requirements			Testing in a live environment	
POLICY TRIAL	PROOF OF CONCEPT	SCOPING PHASE	"AGILE" APPROACH[1]	INNOVATION PARTNERSHIPS	TEST AND LEARN	PILOT
Definition - what is it?						
Analysis of the potential impact or outcomes from a range of policy options	A practical demonstration of a product or service to show whether it can potentially meet the broad requirements	An initial phase of a project to develop the detailed scope and/or specifications where these cannot be clearly defined at the project outset	A project management methodology using iterative development that values human communication and feedback, adapting to change and producing working results	A procurement process under PCR2015 designed to support the design, development and acquisition of new and innovative products and services that are not currently available in the market	The implementation of one or more potential or prototype products or solutions (usually in a live environment) for the purposes of validation and/or dev of final requirements	The implementation of a preferred product or solution in a selected area or population prior to full rollout
What is the programme primarily testing/achieving?						
Is this a good idea or not? What are the potential costs/ benefits?	Does the product or service work in a test environment? What would we need to change to meet our specific requirements?	What are the options and their pros and cons? What are the scope and requirements that we want to contract for?	At the end of each "Sprint": Does the product/ service do what it is designed to at this stage? How does it need to evolve to meet emerging needs?	Not applicable - Innovation Partnerships are a procurement mechanism. However, testing mechanisms should be built into projects	How does the product or service options work in a live environment? Which, if any, should we implement and what needs to change?	What are the issues we might encounter in a full-scale roll-out? What adaptations might be needed to detailed service specifications?
Examples of when to use?						
When several different policy options are available, particularly if these would require different approaches to the supply market	When considering whether a product or service that is currently available in the market might be appropriate or could be adapted to meet the requirements	When the service requirements or outcomes are reasonably clear but, for example, where the boundaries between in-house and outsourced services need to be defined or the scope may vary	Where the desired service outcomes are clear but the method of delivery and/or the products required to deliver the service require bespoke development or adaptation	To develop particularly novel, unique or innovative solutions that do not exist in the current marketplace	Where the proposed solution is at a reasonably advanced stage of development but has not been tested in a live environment to demonstrate its potential efficacy	As the final stage of testing prior to a full-scale implementation to ensure that the solution can be effectively rolled out on a larger scale

[1] **This is the GDS standard approach for Digital Delivery as set out in the Technology Code of Practice**

Early testing		Developing and building requirements			Testing in a live environment	
POLICY TRIAL	PROOF OF CONCEPT	SCOPING PHASE	"AGILE" APPROACH[1]	INNOVATION PARTNERSHIPS	TEST AND LEARN	PILOT
Description of the approach - what does it involve?						
Approach(es) are tested against a control environment to determine whether they deliver the desired outcome	The approach/ product is mocked up and tested in a safe environment to check that it will work as intended	The approach is designed with the selected supplier(s) as information on the scale and scope of the deliverables is defined	The end product is developed on an ongoing, iterative basis in partnership with the selected supplier(s)	The Innovation Partnership procurement process is set out under PCR2015	A proposed approach is tested for a defined period, with learnings used to refine the final requirements	The preferred approach is implementation on a smaller scale to identify learnings for the subsequent implementation
Potential outcomes and next steps						
Positive outcomes may lead to further testing, negative outcomes to a redefined approach	Product/service testing in a live environment or full-scale implementation	Product/service testing in a live environment or full-scale implementation	Implementation is phased and iterative with the final service being gradually defined	Not applicable	Final requirement definition (that may then be retendered)	Full-scale implementation with the application of learning points

3. Why use testing and pilots

3.1 Benefits of testing and pilots

3.1.1 Having appropriate testing of a new service delivery model at various stages of development is critical to understanding any challenges in the environment. Testing services will reduce the risk of failure at scale, enable learning and validation of results, improve the delivery and stakeholder relations and provide a feedback loop:

- Testing will help inform if you are ready to scale delivery. Testing can help validate assumptions and determine if any adjustments to delivery are necessary and reveal unforeseen challenges that might arise. In the case of Pilots, this will also give the supplier's team and the implementation team a chance to work together before full implementation

- Testing is an opportunity to gauge your target population's reaction to delivery. For pilots, by selecting a sample group that is representative of the project's target population testing can help confirm whether or not the delivery model is appropriate and provide confidence to end users.

- Testing can help you make better decisions about how to allocate time and resources. For example, you might learn that changes to information technology systems are necessary or that you need to allocate more time for completing the evaluation activities than you had originally anticipated.

- Testing can help ensure that you are well prepared to measure the success of your delivery. Testing can highlight any adjustments to your benefits realisation methodology to ensure that you are measuring the desired outcomes in the best way possible. For example, a Pilot will be an opportunity to test key performance indicators.

"Having appropriate testing of a new service delivery model at various stages of development is critical to understanding any challenges in the environment."

4. Early testing approaches

4.1 Policy Trials

4.1.1 A trial programme is usually used to test whether a proposed policy will deliver the intended outcomes or behaviours. It requires assessment of the proposed policy (or several alternatives) against a control scenario where “no change” is assumed. This may take a more theoretical or “desktop” approach (for example, using market research surveys to test the public’s reaction to a proposed policy change or applying statistical research to compare policy options) or it may involve groups of a targeted population testing new approaches (for example, comparing different approaches to the payment of benefits to ensure payments are made correctly and on time).

4.1.2 Trial programmes can be delivered with internal capability or it may be necessary to engage external providers to manage the trial. If a third party is being used that may be expected to bid for delivery of the end solution, you will need to ensure that involvement in the trial will not give any perceived or real advantage in bidding at later stages in the programme.

4.1.3 Trials will not usually lead immediately into a full-scale implementation of a new service without further product or service development and subsequent testing of the preferred delivery mechanism. The Cabinet Office Policy Lab team is available to support departments with the development and implementation of Policy Trials.

4.2 Proof of Concept

4.2.1 A Proof of Concept (POC) is usually a theoretical demonstration of a product, process or service in order to determine whether it can be turned into reality. This involves testing whether it is likely to be viable and has the potential to be built in a real-life environment.

4.2.2 In software development, a POC is often used to demonstrate whether a product has the capability to provide the desired functionality. For example, a software “sandbox” environment which could be used to demonstrate if a digital outcome for government is viable. In an outsourcing context, as well as building a prototype service design model, a POC project might seek to identify potential technical, logistical or other operational issues that could interfere with the future success of a full-scale implementation. Although a POC will look to test the extent to which the proposed solution will address the department’s operational goals and requirements, it may not identify the final delivery mechanisms or processes that will be required.

4.2.3 The outcome of a POC should be to identify a clear service definition. This is an early version of the final service to be delivered and can then be used to test viability across a range of end users and inform the delivery model assessment for how a service should be delivered. A POC does not replace a Pilot but may be one of the project milestones that precedes it.

4.2.4 In some cases, it might also be useful to consider whether a “Proof of Value” exercise will be more appropriate than a straightforward POC. While a traditional POC project will look to test whether a product or service will work to deliver the functional requirements, a Proof of Value exercise should also determine the economic case for implementation.

5. Developing and building requirements

5.1 Scoping Phase

5.1.1 In many instances at the start of an outsourcing project, although there is a clear issue that needs to be addressed, the final deliverables and the mechanism to deliver the desired outcomes are unclear. In these circumstances, a department may decide that the first phase of work is to define the scope and service requirements (e.g. the specific geographical areas to be covered, which services will be retained, or the prioritisation and timescales for rollout).

5.1.2 The department may have the capability to manage this stage internally but, in many cases, it will be appropriate to use external advisers and/or potential future suppliers to help understand best practice, review options and alternatives, build the business case for change and review the supply market's capability and experience in delivering similar projects. Involving suppliers in developing the department's requirements (especially those that would bid for delivery of the end solution) does have legal implications and there are potential implications for the suppliers involved. See section 9 for practical and legal considerations.

5.1.3 The outcome of the scoping phase (sometimes referred to as a 'development phase') will include a specification of requirements, setting out the department's requirements that can be used in any subsequent procurement exercise and/or contracting for the end deliverables.

5.1.4 Subsequent phases of the project will usually include a form of pilot to test the operational delivery prior to full implementation.

5.2 Agile approach

5.2.1 "Agile" methodology is a project management process, whereby demands and solutions evolve through the collaborative efforts of self-organising and cross-functional teams (often including third party providers) and their end customers. Agile processes are frequently used in software development projects and increasingly the same methodology is being applied to the development and evolution of service solutions.

5.2.2 The use of Agile processes may be appropriate, particularly when using an Innovation Partnership procurement route, where the department and its selected service provider(s) are looking to develop a particularly novel or innovative solution.

5.2.3 In an Agile project, the detailed requirements are not clear at the outset and evolve through an ongoing series of short "sprints", with the conclusion of each sprint being to test the outcomes, validate whether these are in line with expectations and consequently determine the requirements for the next sprint.

5.2.4 Projects using Agile processes may not require a formal pilot stage - as the requirements develop over time and are tested continually, they may naturally and

incrementally move into a live environment and a full pilot might therefore be an unnecessary phase of the project.

5.2.5 Working in an Agile approach does also bring some additional risks to take into consideration, such as the need to maintain effective change and cost control around the original required outcomes. Guidance is provided by the Government Digital Service on using Agile.

5.3 Innovation Partnerships

5.3.1 An Innovation Partnership is one of the procurement procedures set out in the Public Contract Regulations (PCR 2015) and is designed to cover the development and subsequent purchase (in a single procurement procedure) of particularly new or innovative products, services or works.

5.3.2 Innovation Partnerships involve the selection of a preferred partner (or partners) to develop a new product, service or works. However, this procurement method can only be used if the product, service or works is “an innovative solution” not currently available on the market.

5.3.3 Under an Innovation Partnership procedure, following a competitive selection process, the department and its selected partner(s) work together to develop a suitable solution. An Innovation Partnership procedure is structured in successive phases, following the sequence of steps in a research and innovation process. As such, it may follow Agile development principles, which will usually be the best way to ensure the solution evolves iteratively to meet the needs and deliver the desired outcomes.

5.3.4 Whilst an Innovation Partnership procedure is not, therefore, in itself a testing of piloting process, he processes for researching and developing the solution will need to include appropriate testing mechanisms, as with Agile developments.

6. Testing in a live environment

6.1 Test and Learn

6.1.1 A Test and Learn process is frequently used to test, in a live environment, one or more options for the delivery of services when the final specification of services requires further development (and therefore it is not possible to carry out a fully compliant PCR procurement for the end solution) and/or it is impossible to test the desired final deliverable in its full form. For example, due to timescales or the scale of investment needed.

6.1.2 The lessons learned from a Test and Learn process are used to finalise requirements, which can then form the basis of a further competitive procurement exercise to select the provider(s) who will provide the full services.

6.1.3 One of the key considerations in planning for a Test and Learn process will be how to ensure that the full implementation can be procured competitively. A two-step procurement process will usually be appropriate as the requirements may change significantly between the tested model and the version that is recommended for final implementation. In the procurement for a Test and Learn it is important to be clear about how information will be used and the process for the subsequent procurement.

6.2 Pilots

6.2.1 A pilot will usually be the final stage of testing a delivery model prior to the full rollout of new services. It usually involves the implementation of the proposed services on a localised basis, with the purpose of ironing out operational and logistical issues in advance of the large-scale roll-out of services. This may involve services being tested on a regional or local scale, with certain target populations, for only a subsection of the services required in the full roll-out or with a single supplier in instances where the full rollout will use two or more service providers.

6.2.2 A pilot should not, therefore, be seeking to test the appropriateness of policies, look at alternative options or to justify a business case. Other testing approaches will be better suited to these purposes. With respect to award of the contract, this will be carried out through the usual PCR compliant procurement process.

6.2.3 Provided the piloted delivery model is considered successful, lessons learned from the pilot will be used to adapt and evolve the service requirements. The intention will usually be to move directly from the pilot phase into full implementation.

7. When to consider different types of testing and pilots

7.1 Proportional approach

7.1.1 Testing approaches should be proportional to the size, complexity and level of uncertainty in delivering a service. It is good practice to embed robust, proportionate testing approaches in all project plans including second or subsequent generations of outsourcing, where services are being brought back into the department's control ('insourcing') or where any significant change to the operational delivery, policy implementation or commercial models occurs or is being considered.

7.1.2 A comprehensive testing programme should be embedded into the development of all complex outsourcing projects to reduce the risk of delivery and service failure.

7.2 When to pilot a service

7.2.1 As stated in the Sourcing Playbook, **where a service is being outsourced for the first time, a pilot should be run as part of a programme of testing.**

7.2.2 For other types of complex outsourcing projects, it is also almost always worth the effort to pilot following contract award and prior to roll out of the service at scale and, if so, ensuring that this is included within the service specifications in the procurement exercise. Where a pilot is not being recommended as part of the testing programme for a complex outsourcing project, this should be stated clearly in the business case with the appropriate justification.

7.2.3 The implications for project timescales and current operations in managing testing and any consequent procurement exercise must be taken into consideration during planning and business case processes.

7.2.4 Experience tells us that it is good practice to consider a pilot when:

- Significant transformation of service delivery (including 'insourcing');
- Scope of delivery is large;
- Citizens will need to interact differently with the service;
- New product/service could have far-reaching, unintended consequences;
- Implementing the design and/or solutions will be a costly process; and/or
- The delivery would be difficult to reverse.

7.2.5 There may be less value in completing a pilot where:

- The service is a continuation of a current service and there is no change in the core requirements or delivery approach;

- Other methods of testing have already been conducted that have demonstrated the ability to de-risk the full-scale implementation;
- A programme is being delivered using Agile methodology where testing of outcomes is already embedded as each milestone/sprint is implemented.[2]

Planning which testing approaches to include and whether to include a pilot should begin at the earliest strategic stages of a project.

It should start before any procurement process, and it should be incorporated into the delivery model assessment, sourcing strategy, and, where relevant, procurement documents. The overall approach to testing should be outlined in a clear testing programme including a clear purpose and objectives and the associated success metrics for each test. This includes aligning both testing of the policy outcome and the technical deliverability.

7.2.6 Once the decision has been made to outsource a service, the approach to testing including any potential requirements for a pilot should be tested with the market. Suppliers may provide valuable feedback based on their own experience as to the most appropriate ways in which they can be engaged, suggested contractual and financial models to apply and the timescales required to conduct each phase of service testing.

7.2.7 During the early phases of testing, the project team has the opportunity to take the learning from this testing and change direction. Early tests should also be used to inform later tests and there should be sufficient flexibility in the testing programme and project timeline to include learning from tests.

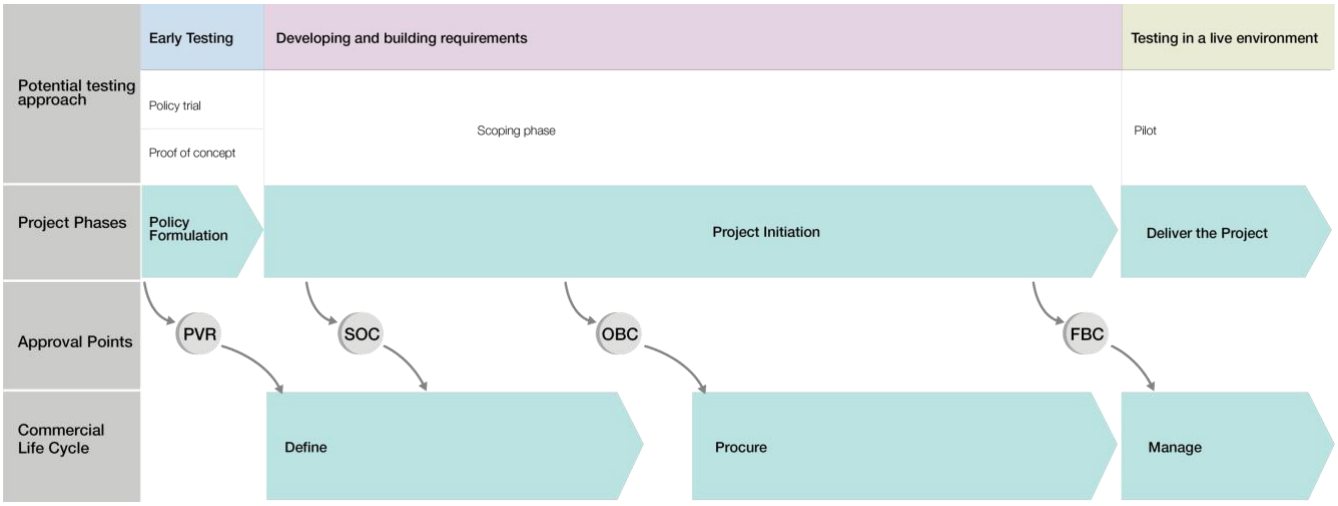
7.2.8 Appendix II provides a list of fictionalised examples of how different types of testing can be used to support projects.

7.2.9 The testing programme should align to key project milestones, be included in the business case and support assurance reviews by providing critical evidence throughout the lifecycle of the project up to full implementation. Where relevant, the testing programme will need to fit within the wider programme Integrated Assurance and Approval Plan agreed with the Infrastructure and Project Authority. An example is outlined in Figure 2 and may include:

- **Trials** of different policy options;
- Developing a **Proof of Concept** to demonstrate that services exist that can deliver the desired functionality;
- Engaging external suppliers to conduct a **Scoping Phase**;
- Completing a **Pilot** to inform full implementation.

[2] **However, the use of Agile methodology should not necessarily preclude you from conducting a Pilot as one of the final milestones prior to implementation.**

Figure 2. Potential single stage procurement approach for testing a service



> "The approach to testing including any potential requirements for a pilot should be tested with the market. Suppliers may provide valuable feedback based on their own experience as to the most appropriate ways in which they can be engaged."

8. Designing, running and evaluating tests and pilots

8.1 Designing tests and pilots

8.1.1 Tests, including pilots, should be developed to ensure success and to mitigate potential risks prior to scaled implementation. Factors to take into consideration when designing effective tests and pilots are outlined in Figure 3.

Figure 3: Considerations when designing an effective test

Objectives and Evaluation

- Objectives for testing should be closely tied to the objectives of the project, however it is important to understand the tests themselves are not delivering the full benefits that will be set out in the overall programme. Some of these benefits may only be achievable due to the scale/volume of services in the full roll-out.
- Objectives should be SMART (specific, measurable, achievable, realistic, timebound) and include defining what needs to be tested or implemented (this could be parts of the solution where risk exists or limited geographic areas but should be representative of different circumstances) and clear success criteria.
- For Pilots, objectives might include testing:
  - The feasibility of **operational procedures**
  - The validity of **financial or operational assumptions**, e.g. the scope and scale of future requirements
  - The **delivery capability** of the solution
  - The practical effects of **policy delivery in the real world**
  - The capability of **systems and IT solutions**
  - Any **interdependencies** between teams, processes or policy decisions
- Develop an evaluation methodology to ensure the tests can be objectively scored against the objectives and identify any remedial actions.
- Consider the potential outcomes of the evaluation and next steps or alternative approaches to take in the event that the test outcomes are not fully successful, including the impact on overall project timescales

Scope and Scale

- Identify the priorities of what will be tested, the size and boundaries measured, how to make the test as realistic and representative as possible, and when and where the tests will be run. Where appropriate, consider the impact of regional differences or other differences in the test population on the results and the extent to which results can be replicated across other locations or populations.

- What you want to do with the product, who owns it, how is it to be taken forward after the end of that stage.

Resources and Governance

- Identify the internal resources required to support the management and evaluation of the test including the materials, assets and IT. Review any training needs for internal resources.
- Agree a governance model and monitoring processes with clear escalation routes in the event of any issues or unintended consequences.
- Ensure sufficient budget is available including any external resources.

Timescales

- Establish the appropriate length of the test (especially when considering Test and Learn and Pilot processes) to effectively test the solution and outputs. Include sufficient time to implement lessons from the test. Build in flexibility to allow for increased time should more testing be required or if requirements scope changes as a result of the testing.
- Ensure that the team and relevant stakeholders have clarity of steps, actions and owners at the start of the testing and meetings are scheduled to update the governance body and relevant stakeholders on progress.
- A simple test process may last a number of weeks. If complex, a Test and Learn or Pilot could take several months or even a year or more to effectively test the end to end service and its impact on complex stakeholder groups. Sufficient time should also be planned in to complete the feedback loop and make any changes prior to scaling up the service.

Commercial

- Before finalising the design, test it with potential suppliers. Suppliers can provide valuable insight into how best to engage the market, what has worked previously and innovative new ways of working.
- Contractual considerations for tests involving suppliers include:
  - The scale, scope and timescale for the delivery
  - The test evaluation process and requirements for management information, reporting etc.
  - Resourcing considerations
  - Management of intellectual property
  - Change controls
  - IPR – how to control and manage this in live running
- Additional contractual considerations for Pilots include:
  - Transition to full implementation - e.g. will there be a cessation of service between the Pilot and implementation to allow for any necessary changes?
  - If a service is multi-sourced, whether the services are being piloted with all suppliers or just a single source/location and, if so, how learnings will be applied and requirements amended or other service providers prior to the rollout (If substantial

changes are necessary, the contract(s) may need to be terminated and the roll-out services retendered)

- Put in place appropriate contract management mechanisms for the test, where using a Pilot, ensure continuity in contract management teams.
- See Section 9 for further detail on commercial considerations.

Legal

- Legal advice should be sought where testing involves suppliers. Tests must be compliant with procurement and state aid law.
- See Section 9 for further detail on legal considerations.

8.2 Running Tests

8.2.1 Progress against the objectives should be regularly reviewed with any risks or issues dealt with promptly. Throughout the test, ensure that data is being captured on internal and external factors which may be influencing delivery.

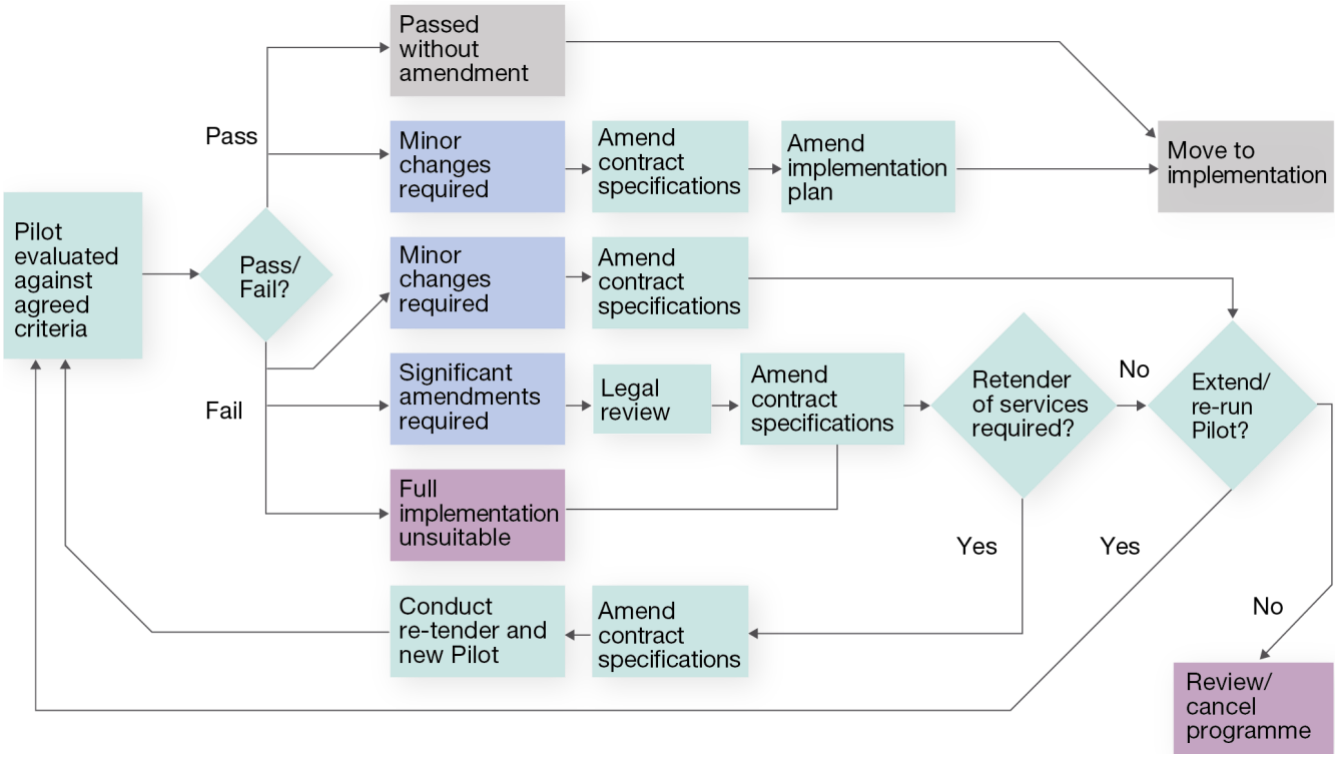
8.2.2 Encourage open methods of communication with all parties during the testing and also at the end to ensure the capture of lessons learned. Consider ideal points of time for feedback data from resources, suppliers and users of the service where the information is of most value e.g. at the end of a phase rather than right at the end of the test or pilot.

8.3 Evaluating the Results

8.3.1 At the end of the test, collect and evaluate the results including what worked, what didn't and what had to be changed or addressed. Further insights should be captured through interviews with key stakeholders including end users.

8.3.2 Develop a final report summarising the findings against the success criteria, identifying any gaps in delivery performance and highlighting if service changes are needed. Consider how to use the findings to inform the next stage of the procurement. Some tests or pilots may serve a particular population and results may vary based on demographics (socio-economic, average age, etc.). Figure 4 outlines the next steps following a Pilot, i.e. where a department is looking to move into implementation phases.

Figure 4: Potential pilot outcomes and next steps



8.3.3 If the piloted delivery model hasn't achieved what you had hoped, it's time to start thinking about how to communicate the outcome and any alternative next steps. Have the outcomes not been delivered completely or do the results simply indicate that more refinement is needed? The reason for completing a Pilot is to learn in a controlled environment where final issues can be resolved before roll out.

8.3.4 If the piloted delivery model goes well, this is the time to review and refine the transition plan that would have been agreed in the original procurement. This is the plan to take the delivery model from the Pilot to roll out. That transition plan could be anywhere from a few weeks to a few months, depending on how much time is needed to scale up the approach. There will still be risks inherent in a "big bang" implementation following a pilot and further staggering of implementation may help mitigate this and reduce the pressure on the implementation team.

"At the end of the test, collect and evaluate the results including what worked, what didn't and what had to be changed or addressed."

9. Developing the commercial strategy and approaching the market

9.1 Incorporating testing and pilots into commercial strategy

9.1.1 The overall commercial strategy for an outsourcing project should take into account the testing programme and align it to the project timeline. Key considerations include:

- Which are the key project milestones where testing of the service will be required? What is the most appropriate approach to this testing?
- Has sufficient time been allocated to testing, evaluation and redesign within the project timelines?
- Is there a solution to providing the service that is being outsourced readily available in the market? If not, how will you develop the capability and/or capacity with potential suppliers and incorporate this into the testing process?
- Consider the appropriate procurement process to be followed: The Outsourcing Playbook sets out that Competitive Dialogue or Competitive Procedure with Negotiation should be used for complex outsourcing projects but an Innovation Partnership may also be appropriate if the solution will require a particularly novel approach.
- At what stage in the process will potential suppliers be engaged? What is the most appropriate way of engaging with the external market? How will you ensure that the testing (and, if appropriate, the subsequent roll out of the services) is compliant with the procurement and state aid rules?
- How will testing be funded and have internal costs and external expenditure been included within the programme business case?
- Is there a need to engage external advisors to support the requirements definition, evaluation or business case development?
- How many stages in the programme will require procurement exercises? (e.g. Phase 1 consultancy support for the development of requirements, Phase 2 to conduct a limited Test and Learn, Phase 3 for the Pilot and rollout)?
- Will the final outcome be a single-source or multi-source solution? If the latter, how will providers be expected to work together and will this therefore need to be incorporated into the testing process?

9.2 Single-stage or multi-stage procurement approaches

9.2.1 Departments will need to consider whether it is appropriate to commit to the final supplier selection in a single-stage procurement exercise or whether two or more procurements will be more appropriate to cover the different phases of developing

the requirements and the testing that is associated with each of these phases. Figure 5 provides three example scenarios. Departments should also consider whether to undertake some or all of these testing in-house if internal expertise and resources are available.

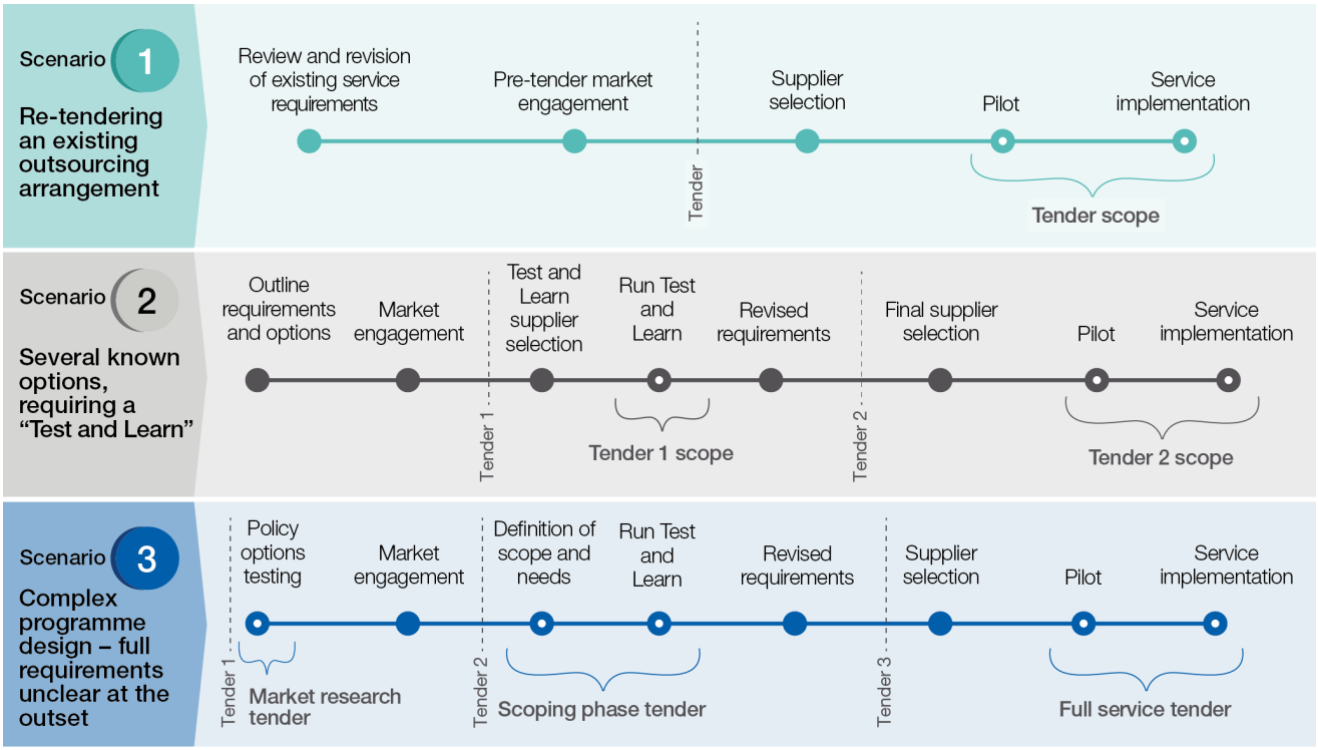
9.2.2 For example, as shown in Figure 5 below, it might be appropriate for a department to conduct several distinct procurement exercises to cover:

- Market research and trials of a policy to test which of several policy options would deliver the desired outcomes.
- Use of external advisory support to scope and build the requirements, which might include a "Test and Learn" phase to validate the proposed requirements.
- The contract(s) for wider implementation of the proposed services, including a regional Pilot to ensure a smooth rollout.

9.2.3 If the services are already outsourced but will be transitioning to a new supplier or if the requirements are easily defined and well-established in the market, a single-stage tender, including a pilot, may be more cost-effective and quicker to implement.

9.2.4 The Model Services Contract (MSC) is a useful starting point for a single stage procurement where a pilot phase forms part of a larger contract for complex services, although it will require some amendment to take into account the mechanics of running the pilot, rolling out the services, termination in the event of failure of the pilot and other specifics relevant to the project.

Figure 5: Examples of possible procurement approaches



9.2.5 When adopting a single-stage tender including a pilot, or procuring a specific test, there are a number of key commercial considerations. Where appropriate, these should be addressed in the tender documentation. These include:

- **Delivery and performance** - who is responsible for delivery of the test or pilot? Are there elements the department takes responsibility for rather than the supplier?

- **Transfers of assets** - if assets are to be transferred to the service provider, when and how will this occur? How will the associated liabilities be transferred?

- **Transfer of Undertakings (Protection of Employment) Regulations ("TUPE")** – where the pilot requires the outsourcing of certain activities, and/or or the transfer of assets, it will be necessary to determine whether or not TUPE applies to departmental employees. It will be necessary to engage with HR and the Government Legal Department Employment Group to assess TUPE implications at an early stage of the planning process. If TUPE applies, employees assigned to the transferring activities or assets will transfer to the service provider by operation of law. Further, they are given enhanced pension protection under New Fair Deal policy, which must also be factored in.

- **Intellectual property ownership and rights to use** - e.g. how can learnings from the test / pilot potentially be shared with other service providers and/or be used in subsequent competitive procurements. The starting position for any test/pilot is that intellectual property ownership will (to the extent legally possible) rest with the department. Where this is not possible, the department (and its suppliers) will require a right to use. You should also consider how intellectual property, for example from service improvements, will be managed throughout the service contract.

- **Cost variation** - how will learnings on costs impact the pricing for full implementation? Will changes in costs proposed by the supplier in its original tender amount to a substantial change to the contract and therefore require a second stage procurement for the roll out?

- **Service specifications** - how will learnings be translated into amendments for the full implementation? Note that any material changes to the service specification may lead to a requirement to re-tender services, unless specifically provided for in the original tender documents

- **Termination rights** - does the department require the right to terminate the contract at the end of the pilot process? Will the right be required both with (e.g. if the pilot fails to meet expectations) and without cause (e.g. because of a change in policy)? What are the consequences of termination? Consider both financial and operational consequences. For example, what costs will the department and the supplier bear if the project is terminated after the pilot? Will operations need to transfer back into the department and, if so, how will this be managed and what support should the supplier provide? Will the services need to be re-tendered?

- **Liability, risk, dependencies and limitations** - What are the financial limitations of liability (and potentially the insurance requirements for service providers)? How does the risk profile change? Are there any service credit reliefs?

9.3 Financial Models

9.3.1 The nature, scope and scale of a test or pilot will vary widely and it is therefore impossible to be prescriptive about the most appropriate financial model to manage costs and payments to the supplier. Some options to consider are outlined in Figure 6.

Figure 6: Potential financial models to manage cost and payments to suppliers

Option	Pros	Cons
Costing the test or pilot as a distinct phase in the implementation, with specific charges. In the case of pilots, having costs detailed separately to full implementation pricing	• Clear and transparent • Suppliers will perceive as "fair" • Limits cost exposure in the event that the test / pilot fails to meet expectations • Potential for clear distinction between pilot costs and costs after full implementation (e.g. cost base may change significantly due to economies of scale) • Gives the potential for a "clean break" if a pilot is unsuccessful or the project is stopped • Justifies IPR transfer to the department • In the case of pilots, where full implementation is priced from the outset, this carries less legal risk	• Most cost risks are transferred to the department • Potentially higher total project cost • Cost risks associated with "scope creep" or if test /pilot timescales overrun • Increased procurement risk
Pilot costs factored into total contract pricing by the provider(s) as part of "cost of sales"	• Transfers cost risks for the pilot to the service providers • Drives efficiency in the Pilot process -	• Pilot costs are not transparent • Costs for the pilot may be factored in to providers' pricing for the fully-implemented

Option	Pros	Cons
	disincentivises time overruns • Motivates provider to ensure Pilot is successful and will move to full implementation • In the case of pilots, where full implementation is priced from the outset, this carries less legal risk	service, increasing total cost of ownership • Providers are likely to require the department to guarantee payment of pilot costs in the event full implementation is cancelled • Incentivises providers to cut costs, scope, duration or scale of the pilot to maximise profitability, reducing the value of the pilot to test full implementation
Testing / pilot costs charged on a cost plus basis	• Clear and transparent • Allows for joint learning around the full costs of implementation - may be suitable where full implementation costs are unclear at pilot stage • Perceived as fair by service providers • Potential to invite bids at cost price only	• All cost risks transferred to the department - requires close cost management • Incentivises cost and timing overruns • In the case of pilots, where full implementation is not priced from the outset, this carries greater legal risk

9.4 Procurement Rules

9.4.1 As with all procurements, departments need to consider compliance with the PCR 2015 and the principles of transparency, equal treatment and non-discrimination ('procurement principles') when considering testing or piloting that involves suppliers at any stage in the process. Pilots can be legally complex and departments should seek legal advice at an early stage to help with structuring the procurement.

9.4.2 Broadly speaking, if the value of the contract is above a certain threshold, the full PCR 2015 regime applies. In this case the department will have to choose which of the procurement procedures set out in the PCRs to follow. Cabinet Office policy on the choice of procurement procedure can be found here. Whilst certain requirements (set out in the PCR 2015) must be met in order to be able to use the competitive dialogue procedure and the competitive procedure with negotiation, given the nature of a pilot, it is likely that one of these procedures will be most appropriate. See the additional guidance note on these procedures here.

9.4.3 Departments should seek legal advice at an early stage to help with structuring the procurement, but things to consider are:

- Whether you have **sufficient certainty over your requirements** to enable the award of the whole of the contract to be identified in a single procurement process or whether the project is best suited to two (or more) separate procurements
- Making sure the **advertisement and procurement documents are clear** about your requirements, how the test/pilot will work, whether you intend to appoint the supplier involved to carry out the roll out or whether the roll out will be subject to a new procurement process, how the evaluation of tenders will be carried out;
- **How you will choose between suppliers** where there are a number of pilots involving different suppliers.

The need to ensure transparency, equal treatment and non-discrimination in procurement can impact suppliers who have been involved in the pilot (or testing).

Where the procurement has been split into separate tendering stages, there is a risk that the supplier involved in testing or a pilot may not be able to be involved in the subsequent roll out. Ways to mitigate this include;

- ensuring that all suppliers bidding for the roll out stage are given the same information that the supplier involved in the Pilot stage has been given; and
- ensuring adequate time limits for the receipt of tenders to ensure suppliers not involved in the previous phases have had the opportunity to familiarise themselves with the project to the same extent as the testing or Pilot supplier.

9.5 Free trials

9.5.1 Caution should be exercised where departments are approached by suppliers offering to take part in any form of testing, including where this is offered by a supplier on a free of charge basis. The PCR 2015 and procurement principles will still need to be considered and the department may find its commercial position compromised if such an offer were to be accepted.

9.5.2 Good practice is to develop a robust testing programme with clear objectives and outcomes defined by the department, rather than being led by the offering of a specific supplier.

10. Further resources

10.1.1 There are a number of resources which may be helpful when considering approaches to testing and piloting services. For any questions on this guidance please contact sourcing.programme@cabinetoffice.gov.uk and complextransactions@cabinetoffice.gov.uk.

- **The Sourcing Playbook.** Guidance on insourcing, outsourcing and the mixed economy delivery of public services. Sets out the policy for piloting and testing services.
- **Green Book**. HM Treasury guidance on how to appraise and evaluate policies, projects and programmes. https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-governent
- **Service Manual**. Government digital service guidance setting out how to deliver IT and digital services. https://www.gov.uk/service-manual
- **Government Technology Innovation Strategy**. Framework for departments in developing innovation strategies. https://www.gov.uk/government/publications/the-government-technology-innovation-strategy/the-government-technology-innovation-strategy
- **Project and Programme management**. Essential resources for project delivery professionals leading, managing or involved in projects or programmes in government from the Infrastructure & Projects Authority. https://www.gov.uk/guidance/project-and-programme-management
- **Innovate UK**. Guidance and examples of innovative solutions. https://www.gov.uk/government/organisations/innovate-uk

11. Appendix I: Common pitfalls and lessons learned

Experience has shown us that there are some specific challenges when testing and running pilots. These should be considered and mitigations put in place to avoid them.

Pitfall	Definition
Testing was not considered in the original project plan and/or commercial strategy	Risks both delays to programme implementation when it becomes clear that timings need to be extended and/or significant "teething problems", which may only arise in a fully live environment. A successful test will require enough time to evaluate constraints, performance and policy outcomes and therefore needs to be factored in to project plans at the outset. There may also be procurement risks if the contract has to be amended to take account of implementation issues - e.g. if implementation needs to be substantially delayed and/or milestones (and any deductions for delay) re-configured.
The test is not allowed to run its course so it gives an inaccurate or incomplete result	It is tempting to end a test as soon as any core implementation issues become apparent and are being built into the plans for future rollout however experience shows that tests, and especially pilots, should run their course to identify any further issues.
Unclear, ambiguous or over ambitious testing objectives	Testing and pilots in particular should attempt to test as many of the service elements, within the operational and logistical constraints. In particular, there may be benefits of scale that can only be delivered when the full programme is up and running. It is also almost inevitable that rolling a programme out at greater scale, in different locations or with varying services will create further issues to be addressed. Conversely, over-engineering a pilot can also delay the benefits of the full rollout.
Unclear or ambiguous test evaluation criteria	Poorly-defined test evaluation criteria will likely lead to poor outcomes. Take time to create a robust evaluation model to identify what needs to be changed prior to implementation.
"Analysis paralysis"	Testing or a pilot cannot test every possible scenario of the full-scale implementation. Whilst a reasonable amount of scenario testing is advisable and sufficient time needs to be allowed to ensure the findings are robust, repeat work and over-testing will simply delay the long-term benefits.
Test scope creep	Extending the scope of a test or pilot may run the risk of being deemed a "substantial amendment" under the PCR 2015 and in any case is likely to increase the cost and delay full implementation. Any changes in scope therefore need to be carefully evaluated and legal advice should be sought.

Pitfall	Definition
Changes during a single stage procurement	Procurements run as a single stage (i.e. where the roll out is included in the contract from the outset) are no different to any other type of procurement. The procurement procedure must be clearly set out, the department's requirements clearly specified and tenders submitted and evaluated on the basis of delivery of the whole of the project. Once the contract has been entered into, there should be no substantial changes - this includes pricing changes. If this is unlikely to be possible, then departments should adopt a two (or more) stage procurement process, with the roll out services being procured separately once the testing has been completed. If substantial changes are necessary in a single stage procurement, the contract may need to be terminated and the roll out services re-tendered - provision should be included in the contract to allow for this.
Failing to assess the challenges of scaling up	Once a supplier has been selected and is running a successful pilot, it is tempting to think that implementation will be simple. It remains critical for the overseeing of the roll-out process to be resourced appropriately, with clear project management roles and responsibilities owned by the department.
Failing to consider how intellectual property rights (IPR) will be managed throughout a contract	Major outsourcing programmes will usually include requirements for suppliers to manage continuous improvements of their services, both by implementing changes to processes and through technology or other product changes. It is important to determine who will own or have the licence to use any IPR in these changes or developments to ensure that services can be transitioned effectively to alternative suppliers and/or brought back in-house at the end of the contract.
Pre-judging the outcomes of a test or pilot	A pilot comes towards the end of a long programme of defining requirements, selecting suppliers and testing policies and options in controlled environments. The project team may be hoping that the pilot is successful so that they can move into implementation as soon as possible. It is critical to minimise the implementation risks by focusing on the practical, commercial and operational issues that the pilot should be designed to test and to ensure that all lessons learned are applied in the final requirements.
Incorrect assumptions/ forecast	Volumes for an outsourced service are likely to be subject to change, especially where these are based upon demand for public services. Some variations may be more predictable (e.g. known seasonal variations) but others may be less so, perhaps due to political intervention or changes in public behaviour. Potential variations should be identified in advance and mitigating actions identified as part of the pilot process.

12. Appendix II: Testing and piloting scenarios

The following are a list of fictionalised examples of how different types of testing can be used to support projects.

POLICY TRIAL

Challenge

A Department is looking to distribute a large number of small-scale grants to SMEs, including sole traders, across the UK. Several options are available for how to pay these grants including direct bank transfers (BACS), the use of pre-paid credit cards or cash transfers onto customers' own cards. The Department wants to understand which of these would be preferable to customers and would be most effective in driving grants to be used appropriately.

Approach

The Department runs a procurement exercise to select a market research provider to survey 1,000 small businesses and sole traders to gauge their reactions to different approaches and how they would use any funds made available

Next Steps

The outcomes of the survey indicate a preference for the use of pre-paid credit cards, so the Department conducts a new procurement to run a Test and Learn programme in the West Midlands to see how the process might work in practice

PROOF OF CONCEPT

Challenge

An Arms-length body (ALB) has developed a high degree of technical training expertise in a very specialised area. This training could be very valuable to other countries with less advanced approaches. The ALB would like to identify a future partner to develop appropriate systems and manage the sales process. Several such providers exist in the private sector but none has ever worked with public sector clients or in such a confidential environment.

Approach

Using pre-tender engagement processes, the ALB invites providers to demonstrate their existing capabilities and configure their technology platforms to validate whether the ALB's concept would be a viable proposition and gain a rough idea of the magnitude of work required to adapt systems and processes to their needs.

Next Steps

Four potential suppliers provide initial demonstrations of their product and convince the ALB of the market's potential to meet their needs. A business case is developed in preparation for a tender using the Competitive Dialogue process.

SCOPING PHASE

Challenge

A Department operates 5 call centres in different locations around the UK handling a mixture of both complex and more simplistic customer contact. The Department's technology, processes and facilities are outdated, requiring significant investment, and the private sector market is very capable of handling the operations on an outsourced basis. However, the confidential nature of some interactions and the potential for poor publicity mean that the future service requirements need to be carefully designed.

Approach

Using an Open tender process, the Department seeks bids to conduct a Scoping Phase with either a consultancy or a service provider to determine the feasibility of outsourcing and, if appropriate, the scale and range of service provision that will be included in the tender.

Next Steps

The Department uses a specialist consultant to develop its future requirements, resulting in a recommendation to consolidate services in two locations, one of which would be managed externally. The Department commences staff consultation processes and prepares tenders for a new internal facility and a managed service provision to meet more straightforward needs.

INNOVATION PARTNERSHIP

Challenge

A Department is looking to renew its fleet of specialist emergency vehicles and the associated maintenance and fleet management. In keeping with governmental environment commitments, the Department would like to use hydrogen fuel cell technology to power its future fleet. However, no such vehicles exist on the open market as yet and there is no ready capability to manage and maintain such a fleet.

Approach

The Department seeks bidders to develop an Innovation Partnership for the end-to-end development of vehicles and the ongoing management of the fleet. It is open to both single-source bidders and a consortium or partnership of providers.

Next Steps

Following the bid process, a consortium of a major vehicle manufacturer and a fleet management firm is selected to adapt the existing fleet, develop new models and deliver ongoing maintenance services. Using Agile methodology, vehicles and management processes will be formally tested at several key milestones in the delivery plan.

TEST AND LEARN

Challenge

Following its earlier Policy Trial for distributing grants, the Department has determined customers' preference to use prepaid credit cards to receive payments. This presents the Department with a number of challenges including security and fraud prevention, ensuring funds are used as intended, the use of both physical and "virtual" cards and the underlying technology and processes to manage the grants.

Approach

Because of the wide range of potential service providers, the Department runs a Restricted procurement to select a provider to run a Test and Learn programme with customers in the West Midlands, an area that is fairly typical of the target market. The Department knows that with many open questions, the final service is unlikely to be identical to that being tested and suppliers are made aware that a subsequent competitive procurement will be required for the full programme.

Next Steps

A supplier is selected to run the Test and Learn over a 6 month trial period. The outcomes from this programme are analysed, the specifications and desired outcomes confirmed and a second Restricted procurement is used to select the provider for the full implementation.

PILOT

Challenge

Following a Policy Trial and a subsequent tender using the Competitive Procedure with Negotiation, a Department has selected a provider to manage a complex health education communications programme that is due to roll out across England and Wales over an 18 month period. This will involve managing communications and training programmes with GPs, NHS Trusts and schools, ensuring consistent messages are received by all parties.

Approach

The selected provider has suggested a gradual rollout of services, starting with two pilots in selected NHS Trust areas, one of which has low incidence of the area of concern, while the other has an above average incidence. The pilots will run for a 6 week period, after which results will be reviewed with stakeholder representatives to determine any variations to the approach that will be required. A "break clause" is included in the supplier contract in the unlikely event that the pilots fail to meet SLA target outcomes.

Next Steps

The outcomes of the pilot are used to "tweak" the final service requirements that are implemented nationwide. As these have no impact on the commercial terms of the contract, the rollout is able to continue as planned in the procurement