Centre for
Public Impact
A BCG FOUNDATION

A brief introduction to...
Policy experimentation

Policy experimentation

Randomised controlled trials and quasi-experimentation

What is Policy experimentation?

Policy experimentation is the systematic experimentation by government that robustly assesses the distinct impact of policies and the cost-effectiveness of their implementation. Experimentation allows policies to be piloted, and subsequently adjusted or discontinued, before they are rolled out at scale, which can dramatically reduce the costs of mostly avoidable failures.

Experimentation facilitates government innovation by informing decision-making with new evidence, which departs from the status quo. Clearly, this calls for openness to the possibility of failure. However, such a risk inevitably exists in any policy, which is essentially an experiment conducted on citizens, whether the government treats it as such – via a deliberate process of evaluation and learning – or not.

What does it seek to achieve?

Experimentation aims to ground policy decisions and deployment in observable results - rather than beliefs or feelings - through deliberate and methodologically valid testing. It is underpinned by an acceptance of the inherent uncertainty of the actual outcomes of existing and prospective policies.

Even when there is a solid basis for a policy's merits, experimentation enables the necessary comparison of the cost-effectiveness of different implementation options, which allows for the optimal use of scarce resources.

Keys to success

Photography by SVYATOSLAV ROMANOV

What are the key success factors?

☑ Experiments that build on and complement each other

☑ Political and institutional willingness to put policies on the line and follow wherever results indicate

☑ An experimental setting in which the experiment's effect on stakeholders' behaviours does not undermine its validity

☑ An experimentation timeframe that is sufficient to avoid misleading conclusions about long-term effects

☑ An organisational culture which acknowledges that a pilot that reveals a policy to be flawed or ineffective is essentially a success, as it serves to avoid potentially greater political or economic costs

☑ An experimental design which enables the identification of impact variations across different subgroups or contexts, which are otherwise often masked by the effects of averaging

Keys to success

Things to look out for

The following table displays four specific challenges related to policy experimentation, as well as potential mitigating factors.

Issue	Main challenges to address	Ways to mitigate
Unforgiving politics	The high price of failure in public sector innovation encourages government to stick to the status quo.	Holding small trials to limit the political and economic costs of failure. Fostering a culture where the failure of experiments is acceptable.
Generalisability	Lessons from an experiment conducted in a particular time and place may not hold good for a different context.	Replicating trials in various locations, or at different points in time in the same location.
Unequal access	A robust trial involves withholding a potentially beneficial treatment from the comparison group.	Avoiding obvious treatment disparities between experiment subjects that are geographically close or otherwise related.
Validity	The setting may decrease an experiment's validity, e.g. when affected parties know of the experiment and its desired outcome.	Replicating trials in various locations, or in different points in time in the same location. Adaptingcommunications in order not to jeopardise an experiment's validity.

Approaches and tools

Photography by CRISTIANN EWMAN

Randomised controlled trials (RCTs)

Randomly assigning units of interest (people, households, schools, etc) to control and treatment groups, to compare the distinct causal effect of an intervention with making no changes at all.

Real-world RCTs: experimenting with interventions that entail the allocation of personnel, equipment or other physical resources.

Example:
To streamline the administration of the Malagasy education system, in 2005 the Madagascan Ministry of Education partnered with J-PAL, an anti-poverty NGO, to trial interventions at different levels of educational management using a two-year randomised control trial (RCT). The trial allowed for the comparison of the positive effects of district level versus school-level interventions and significantly improved school attendance, reduced repetition of school years and raised test scores. See: Centre for Public Impact, *Primary education management in Madagascar*

Digital RCTs: internet-based government interactions with citizens, businesses and others allows for comparatively quick, low-cost trials, including testing the effectiveness of multiple implementation variants.

Example:
Prospera Digital is a digital RCT launched in Mexico in 2015. It uses behavioural science to test the impact of personalised and timely information on pregnancy and maternity, and is delivered directly to Mexican women through free SMS messages. See Centre for Public Impact, *Building trust and legitimacy through innovation in Mexico*

Approaches and tools

Quasi-experimental designs: creating the most statistically valid comparison group for an intervention, when practical or ethical considerations prohibit random assignment between control and treatment groups.

Regression discontinuity design: when there is a clear quantitative eligibility threshold, a comparison can be made between subjects just above and below the threshold, assuming that it cannot be manipulated and the two groups are sufficiently similar.

Example: The 2012 low-income broadband programme in the US used regression discontinuity designs to assess the validity of their 14 pilot projects. See: Centre for Public Impact, *Low-income broadband in United States*

Difference-in-differences: measuring differences in the development of outcomes between control and treatment groups, assuming they would otherwise experience similar trends, by verifying parallel time trends pre-intervention and the absence of other disparities which might confound the comparison. This can be done prospectively through gradual roll-out of a policy, or retrospectively, as is more often the case.

Example: a retrospective difference-in-differences estimation of Indonesia's 2006 Programme for Community Empowerment evaluated the programme's impact on six indicators, including poverty status and unemployment rates. The evaluation's findings were subsequently used to inform adaptations of the policy. See: Centre for Public Impact, *Indonesia's Programme for Community Empowerment (PNPM)*

Approaches
and tools

Natural experiments: a setting in which individuals find themselves in different conditions determined by factors outside the government's control, but sufficiently resembling random assignment so as to emulate a trial with naturally occurring control and treatment groups.

Example: The UK's devolution of powers, giving more responsibility to city-regions and counties, has presented the opportunity for various "natural experiments": the possibility of comparing and contrasting naturally occurring differences in policy interventions across areas and jurisdictions. See: Centre for Public Impact, *U-turns: why it's sometimes better to reverse course*
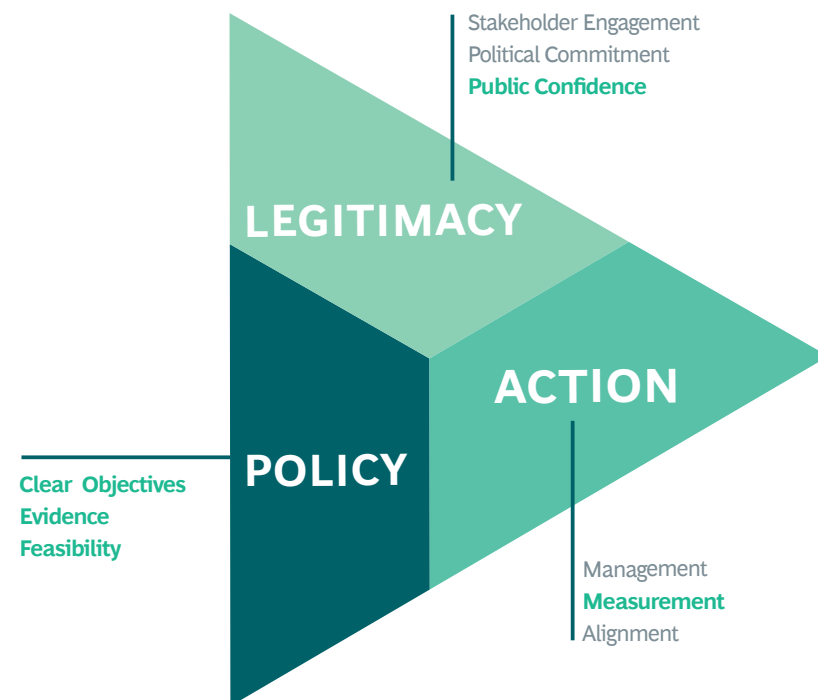
Fundamentals

The

How does Policy experimentation help achieve greater public impact?

CPI's Public Impact Fundamentals are a systematic attempt to understand what makes a successful policy outcome and describe what can be done to maximise the chances of achieving public impact. Below, we have highlighted the elements of the Fundamentals that are most likely to be positively influenced by Policy experimentation.

THE PUBLIC IMPACT FUNDAMENTALS

Stakeholder Engagement
Political Commitment
Public Confidence

LEGITIMACY

ACTION

POLICY

Clear Objectives
Evidence
Feasibility

Management
Measurement
Alignment

Further reading

The Centre for Public Impact,
The Public Impact Fundamentals, 2016

Jonathan Breckon
Better Public Services through Experimental
Government, Alliance for Useful Evidence, 2015

Laura Haynes, Owain Service,
Ben Goldacre and David Torgerson,
Adapt: Developing Public Policy
with Randomised Controlled Trials, Cabinet
Office, Behavioural Insights Team, 2012

HM Treasury,
The Magenta Book:
Guidance for Evaluation, 2011

European Commission,
Testing Social Policy Innovation, 2014

Centre for
Public Impact
A BCG FOUNDATION

Join the conversation

info@centreforpublicimpact.org
Follow us @CPI_foundation