Using the Research Project Lifecycle to Design and Run Small Scale RCTs

YEF Virtual Learning Cafe: 16th July 2024

Tori Olphin, M.B.E. – Head of Data Science, Research and Evaluation, Thames Valley Violence Prevention Partnership

A little bit of background

- Chief Data Scientist and Head of Research with a background in policing
- Has conducted or been involved in the research design and analysis of between 200 and 250 trials conducted live in the public sector
- Including multiple randomised controlled trials
- Experience of conducting large scale change management projects, value stream mapping, target design and monitoring, and resource allocation projects
- Designed with a project manager with experience gained at Amazon and startup tech companies, as well as other larger private sector organisations

- The views and conclusions that have led to the design and presentation of this approach have arisen due to experience that has made it clear that there are repeated pitfalls that public sector agencies fall into when trying to produce good evidence of what works, and when making decisions about how to identify what money and resource should be spent on

Motivations for searching for this approach:

Some Issues with Public Sector Interventions and Research

Common issues:

- Short funding deadlines lead to less planning
- Analysis not planned before implementation
- Lack of tracking of delivery
- Fear of running randomised trials
- Control groups are usually not pre-planned
- Counterfactual not identified
- No baseline measurements taken
- Eligibility criteria not easily identified
- Group sizes not based on prior evidence

Motivations for searching for this approach:

Some Issues with Public Sector Interventions and Research

Leads to:

- Not being able to say what works, or what effect has been had for the money/resource
- Asking for someone to evaluate something after it has already been finished
- Issues with interventions not being picked up until later
- Clashes between interventions
- Rushed implementation
- Inefficient or insufficient levels of resourcing

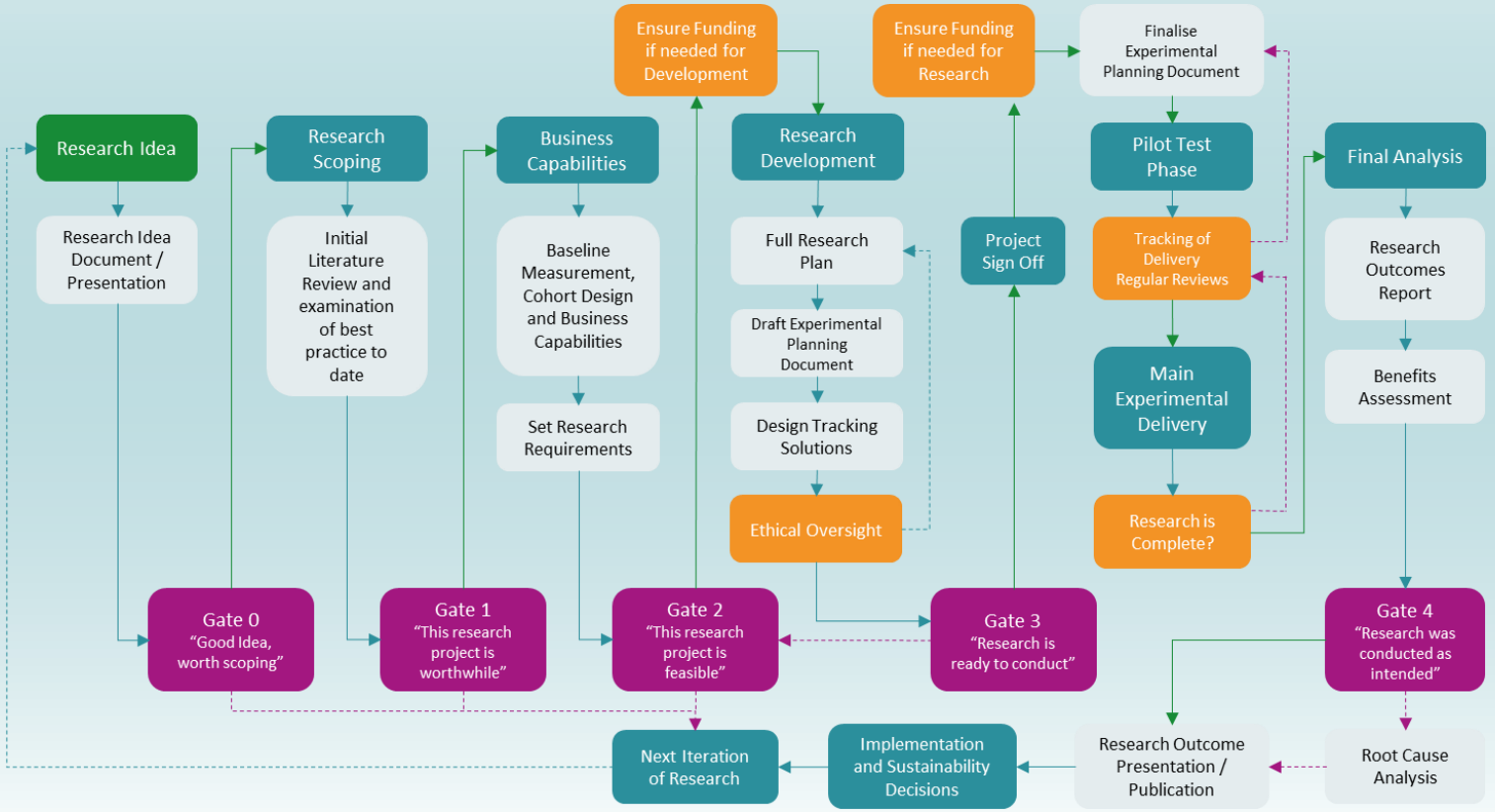
Principles the Lifecycle is built upon

- We need strong evidence to be able to argue for spending money or resource in one way, over another
- Some things either don't work and would be a waste of resources, or backfire and cause harm, therefore we should test to ensure we are not doing either
- The fundamentals of research are the most important part: Spending time on baseline measurements and planning saves far more time and effort later
- Analysis and Evaluation happens in the planning phase, if you have not planned the trial based on how you will evaluate it, the evaluation will be poor
- It is better to have the best test you can do of something that can actually be implemented, than a perfect experiment examining something that only works in laboratory conditions

Principles the Lifecycle is built upon

- One of the most valuable things we can do to start with is acknowledge where we know nothing
- Most of our risk assessment tools are inaccurate, so we cannot actually predict which people will suffer most harm, therefore randomisation may be fairer than we often think
- The harder you try and fail to disprove your findings, the more convincing those findings become. If you do disprove them, you learn lessons for the next implementation
- No trial survives contact with the real world unscathed, and we need to avoid the error of ignoring the weaknesses of evidence. It is better to say that we failed, and learn from it, than to draw unsafe conclusions and base future policy on that false evidence

Research Project Lifecycle

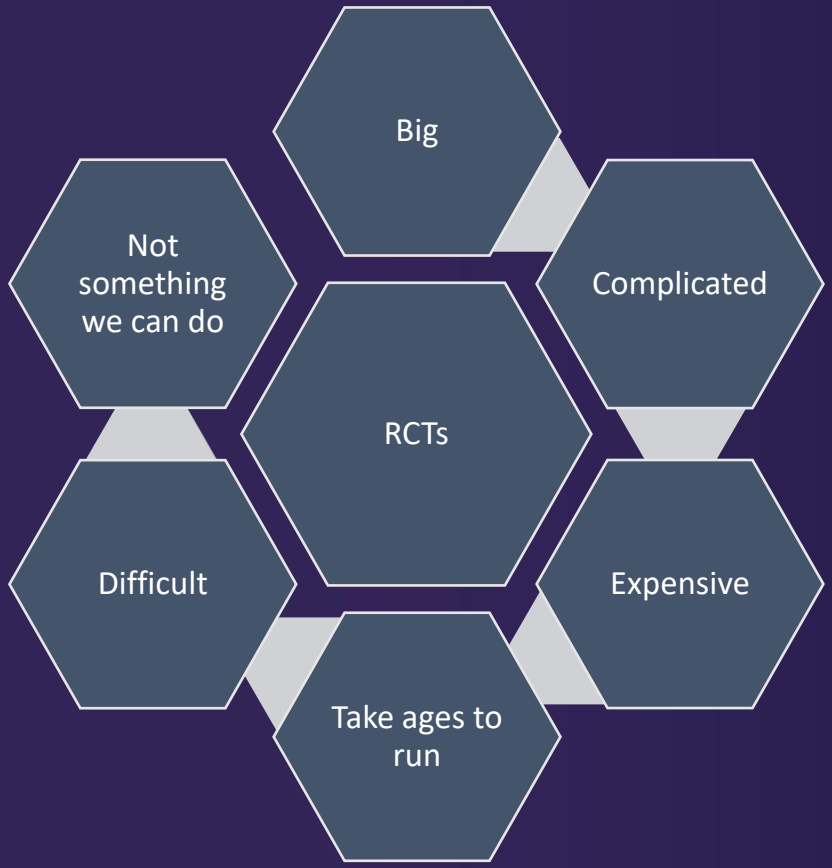


Reference: Olphin, T.P.A. (2023). *Research Project Lifecycle: A Structured Approach to Conducting Research in the Public Sector.* Reading, UK: Thames Valley Violence Reduction Unit.

To reference, please use:

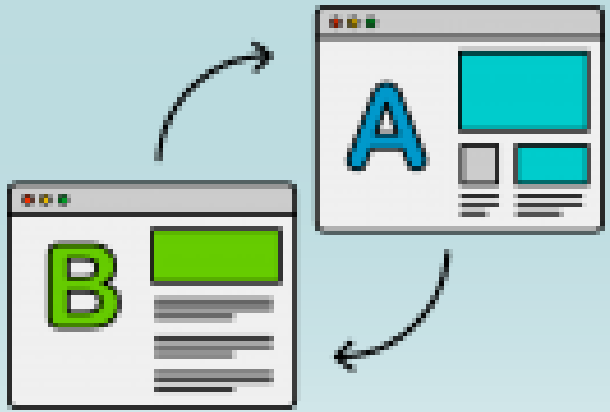
Olphin, T.P.A., (2023). *Research Project Lifecycle: A Structured Approach to Conducting Research in the Public Sector*, Reading, UK: Thames Valley Violence Reduction Unit

Debunking Myths:



Experiments can be cheap and simple to run

- Experiments run in all areas of our lives, and we are likely subject to them each and every day
  - Which colour button are you more likely to click on
  - Whether a price drop makes you more likely to buy a product
- Experiments can have a massive range of complexity
  - From simple nudge trials (e.g. paying taxes on time)
  - To trials of alternatives to prosecution that might require legislation changes (e.g. Project CARA)



Benefits of keeping trials compact

- Faster Results
- Easier to manage
- Less expensive
- More things can be tested
- Efficient Delivery is more likely to be implemented
- Minimises testing requirement for ethics conversations

Some Key Concepts for Keeping RCTs Small

Keep It Simple

More likely to have an impact = Smaller experiment needed to test

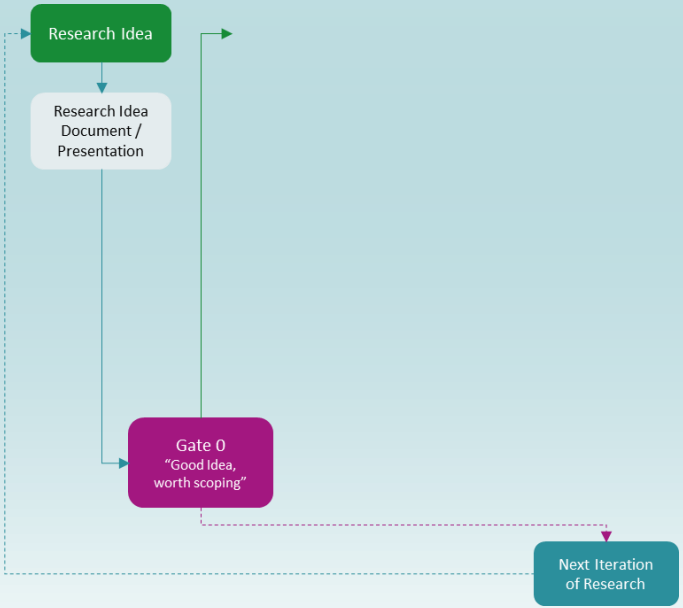
There are many different ways of solving a problem

Some things you need to ensure

- Strong design and proper planning
- Ability to make sure treatment and control are as desired
- Delivery is tracked
- Access to outcome data
- It is much easier to do this when the elements of the trial are within your zone of control – harder for external evaluators to keep costs down

- Be inventive… Consider ways to get same results from something smaller

Ideas and Gate 0



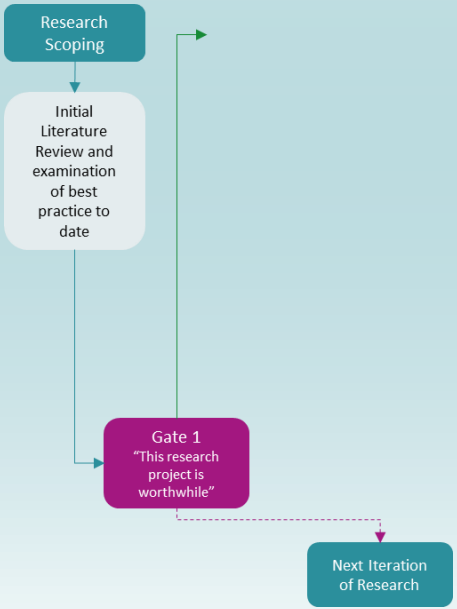
"One of the most valuable things we can do to start with is acknowledge where we know nothing"

Both exciting and terrifying… most things delivered in the public sector have not been tested well

- Encourage many ideas, some will naturally turn out to be smaller
- Find areas of boom and bust funding
- Using something that already exists, but too many cases to work with
- Untested interventions that partners are about to start anyway

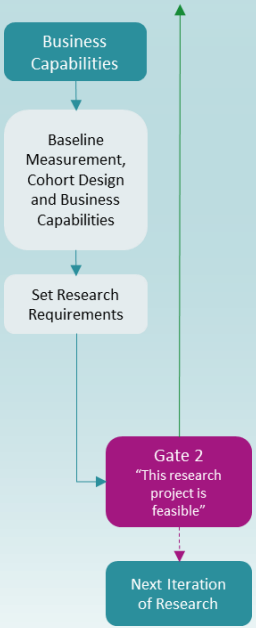
Research Scoping and Gate 1

"More likely to have an impact = Smaller experiment needed to test"



- Design the best intervention you can, based on prior evidence:
  - Use existing literature reviews as a base
  - Rapid evidence review is possible to do swiftly, make sure to summarise detail
  - Out of the evidence that is strongest:
    - What works best, what works less?
    - What power is likely
    - What cohorts have been most effective
    - How should it be designed?
    - What are the elements that would make the best possible intervention
    - Pay attention to context
      - *"Cambridge is not Chicago"*

Business Capabilities and Gate 2



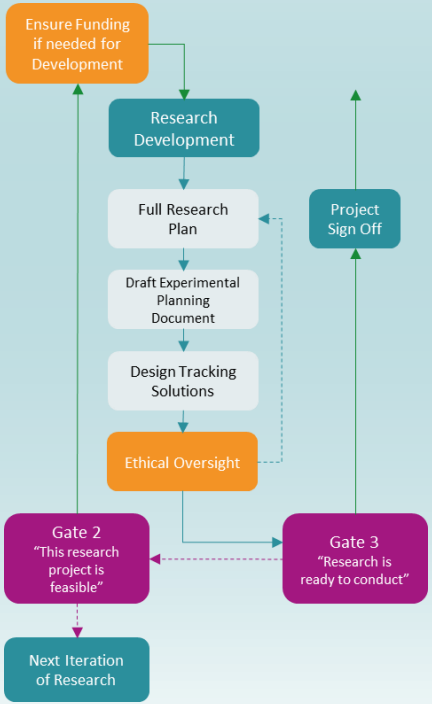
"More likely to have an impact = Smaller experiment needed to test"

"The fundamentals of research are the most important part: Spending time on baseline measurements and planning saves far more time and effort later"

Maximise Likely Effect Size

- Design the best cohort you can, based on prior evidence
  - Identify cohorts with higher likelihood of the unwanted outcome?
  - Ideally avoid expert judgment identification
  - Which groups are more likely to be paid for afterwards?
- Make sure you can actually deliver the intervention
  - Two small complete trials beats one unfinished trial
  - Partner up with other areas and share the load
  - What caseloads can be managed – Effective Staffing

Research Design and Gate 3



"There are many different ways of solving a problem"

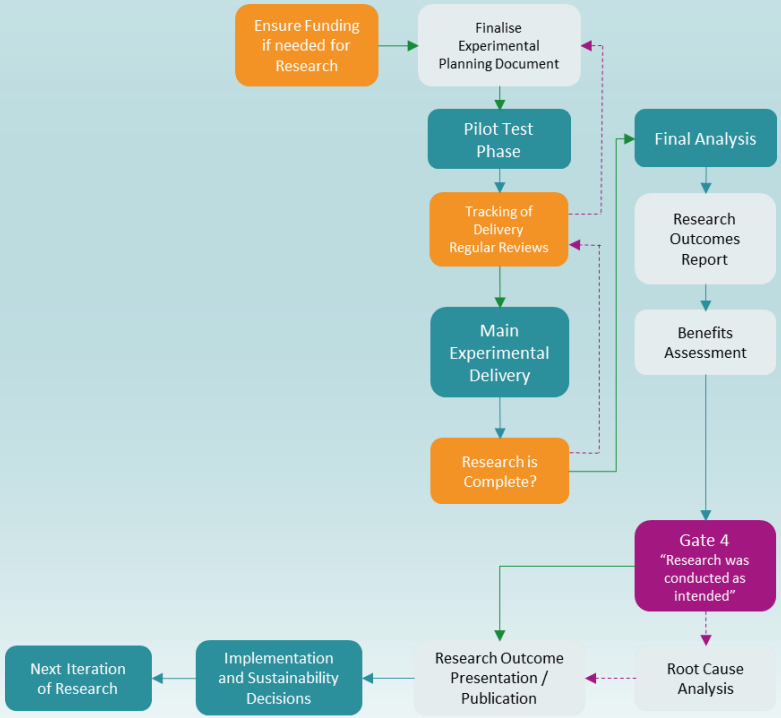
Be inventive

- Most efficient mechanism of delivery (maximise your reach for the money)
- Be clear what you are testing and keep it simple
- Choose something you can deliver (in your zone of control)
- Most ethical version of what can be delivered; often considering this helps to make it smaller
- Accurate tracking ensures you find out early if things are not going to plan
- More uptake (less dropout)
  - Design out barriers
- Realistic environment
  - More similar to business as usual, often more likely to be delivered
  - It can make it more tricky to find a good way, it can also allow for some interesting free experiments
- Motivated delivery
  - If the people delivering the intervention believe in it, they are more likely to keep to it

A Personal Aim

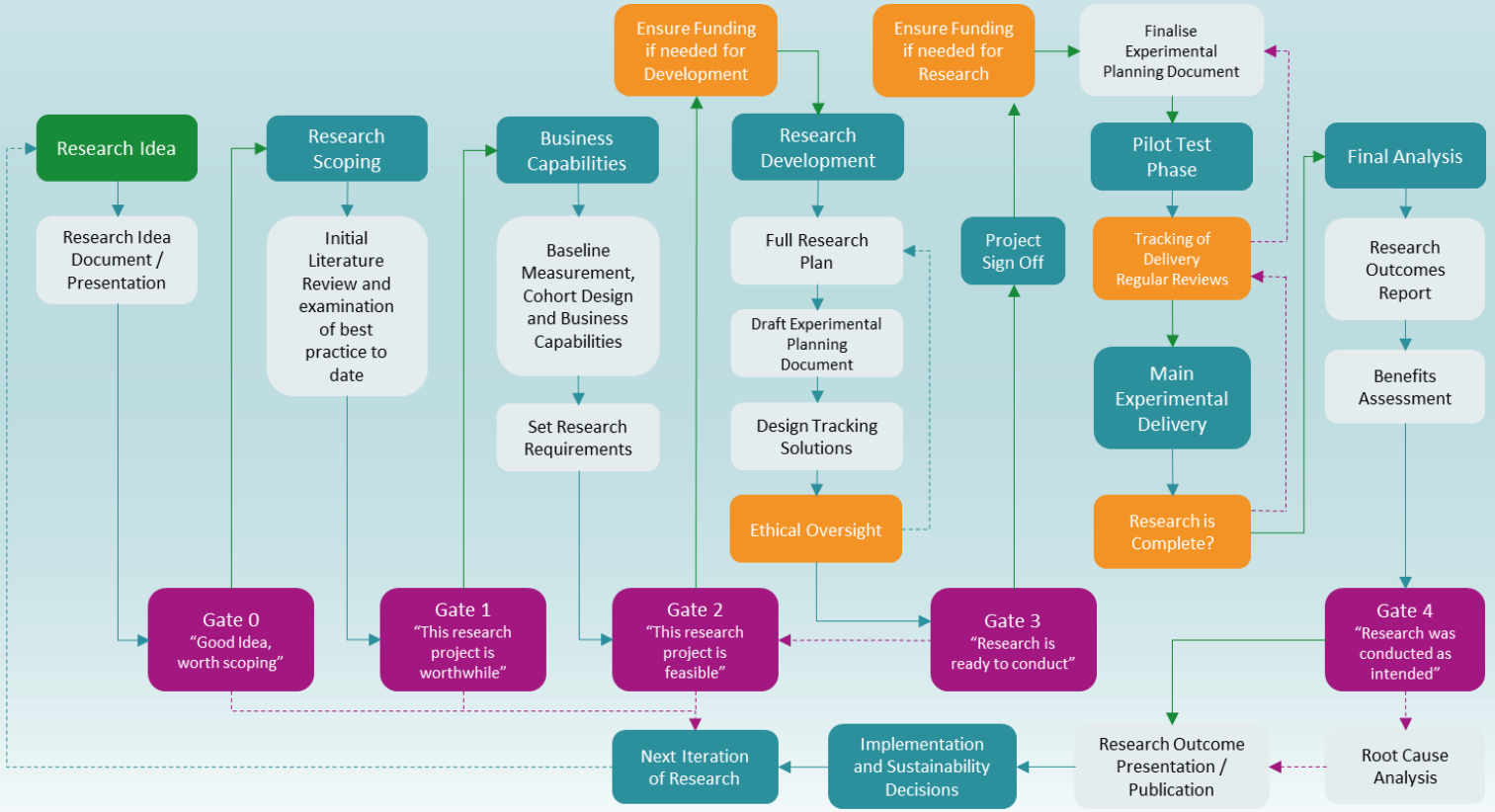
- Many ideas, scoped at gate 3
- Ready to go when we have
  - Funding
  - No overlap
  - Political Desire
- Avoid cohort overlap and trial overlap
  - The more tests you have, the more complications there are with tracking and ensuring no overlap of trials, but it is possible
  - *Pause here if not ready yet*
- Imagine as a funder always getting well scoped trials, with evidence behind choices that have been made, and confidence behind delivery
- Futures scanning can allow us to prioritise research areas that are likely to have funding released, and trials can be ready when this happens

Delivery and Gate 4



- Run Pilot Tests
  - Ensure realistic caseloads to be able to deliver a good service
  - Spot issues early
- Interim Analysis can allow for earlier results
- Avoid scope creep
- Learn from your mistakes, and allow others to as well
- Share what you learned, and make it more likely that more research will be wanted

Research Project Lifecycle



Reference: Olphin, T.P.A. (2023). *Research Project Lifecycle: A Structured Approach to Conducting Research in the Public Sector.* Reading, UK: Thames Valley Violence Reduction Unit.

Acknowledging Potential to be Underpowered

Don’t want to waste resources running things that would never show an effect, but…

No trials run means no knowledge gained

It is better to have some trials that give knowledge, and some that fail

- Acknowledge weakness
  - Failing to find an effect may not mean that it doesn’t work
- Don’t over claim negative findings

Small, simple or inexpensive does not always mean underpowered

- Nudge based campaigns can be run very inexpensively
- Taking the time to make sure the delivery is as efficient as possible often just reduces cost

Giving an effective intervention to a cohort with likely outcome can require very small numbers

Some additional thoughts…

- What are the opportunities you could exploit?
- How can you make it most efficient whilst being externally generalisable?
- Giving thought to it up front
- There is also an important place for qualitative research
  - When we use it meaningfully
  - Can help to get an idea of making something that will work better

THAMES VALLEY
VIOLENCE PREVENTION
PARTNERSHIP

A Focused Deterrence Example

A Focused Deterrence Example

Maybe needs renaming to long-term contextual problem solving, with procedurally just police support

- **Cohort**
  - Likely to reoffend
  - Motivated delivery
- **Intervention**
  - Likely to show an impact
  - Problem solving approach paired with procedural justice
- **Results at 6 months**
  - All measures in the right direction, with statistically significant decrease of 57% in knife offences as suspects

Contact

Tori Olphin, M.B.E.

Head of Data Science, Research and Evaluation
Thames Valley Violence Prevention Partnership

tori.olphin@thamesvalley.police.uk
+447921 808988