



ACADEMY for
JUSTICE COMMISSIONING

Learn, Network & Share

Evening Seminar

Impact Measurement in the justice sector and beyond: Approaches and latest developments

13 April, 2016



@academyjustice



Academy for Justice Commissioning



ACADEMY for JUSTICE COMMISSIONING

Learn, Network & Share

Background to the Academy

- The Academy's mission is to bring people together to share knowledge and best practice and to promote excellence in social justice commissioning
- The Academy was created in 2007 and now has over 3000 cross sector members
- Services are designed to support the development of social justice commissioning and include nationwide events, elearning, commissioning themed learning groups and a website offering commissioning information

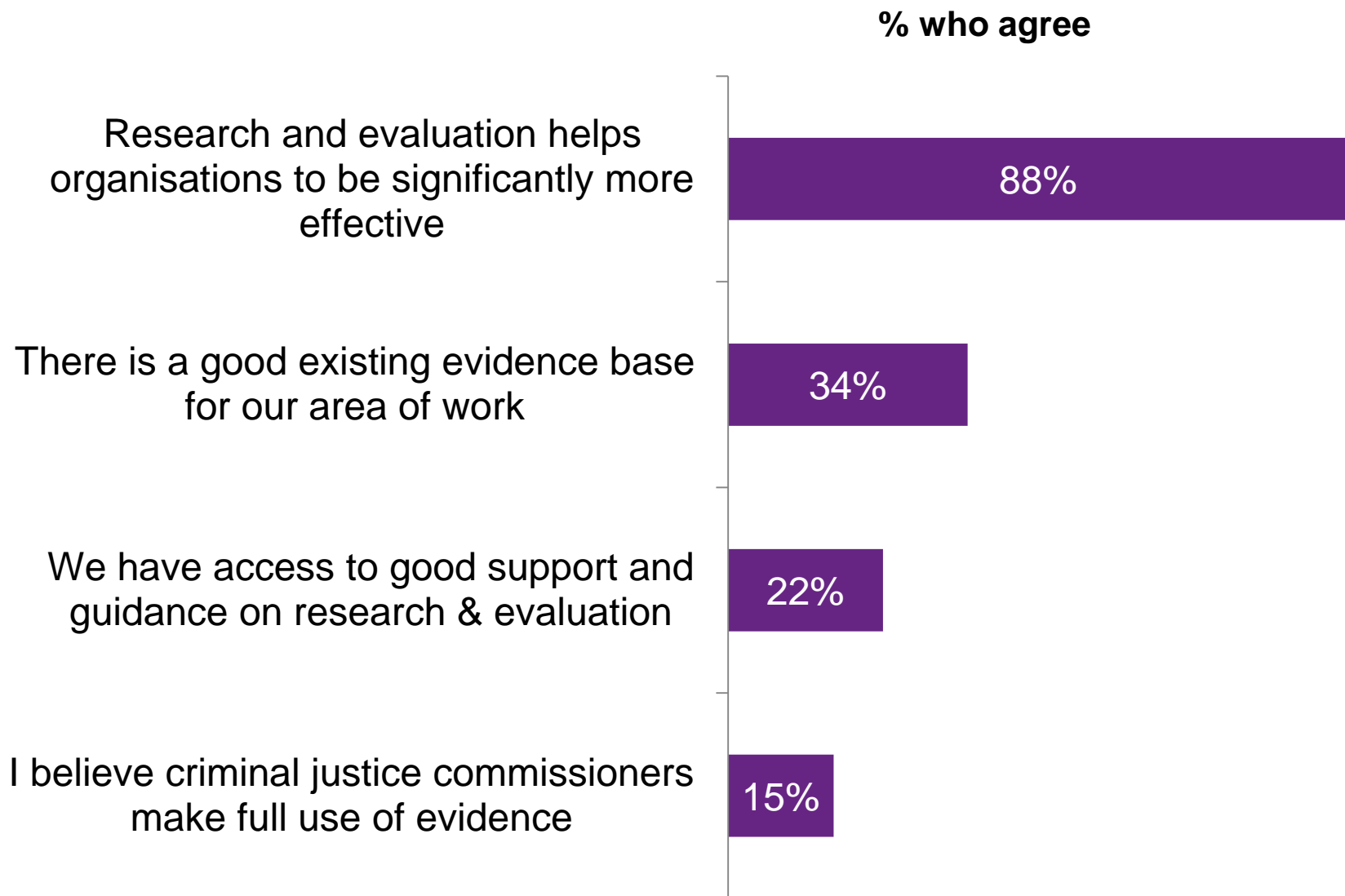
IMPACT EVALUATION AND THE JUSTICE SECTOR

James Noble, New Philanthropy Capital

James.noble@thinknpc.org

London, April 2016

ATTITUDES TO RESEARCH AND EVALUATION



WHY IS MEASURING EFFECTIVENESS AND IMPACT IMPORTANT?



TWO MAIN QUESTIONS

What is the evidence for the thing that we do?

Do we deliver the thing effectively?

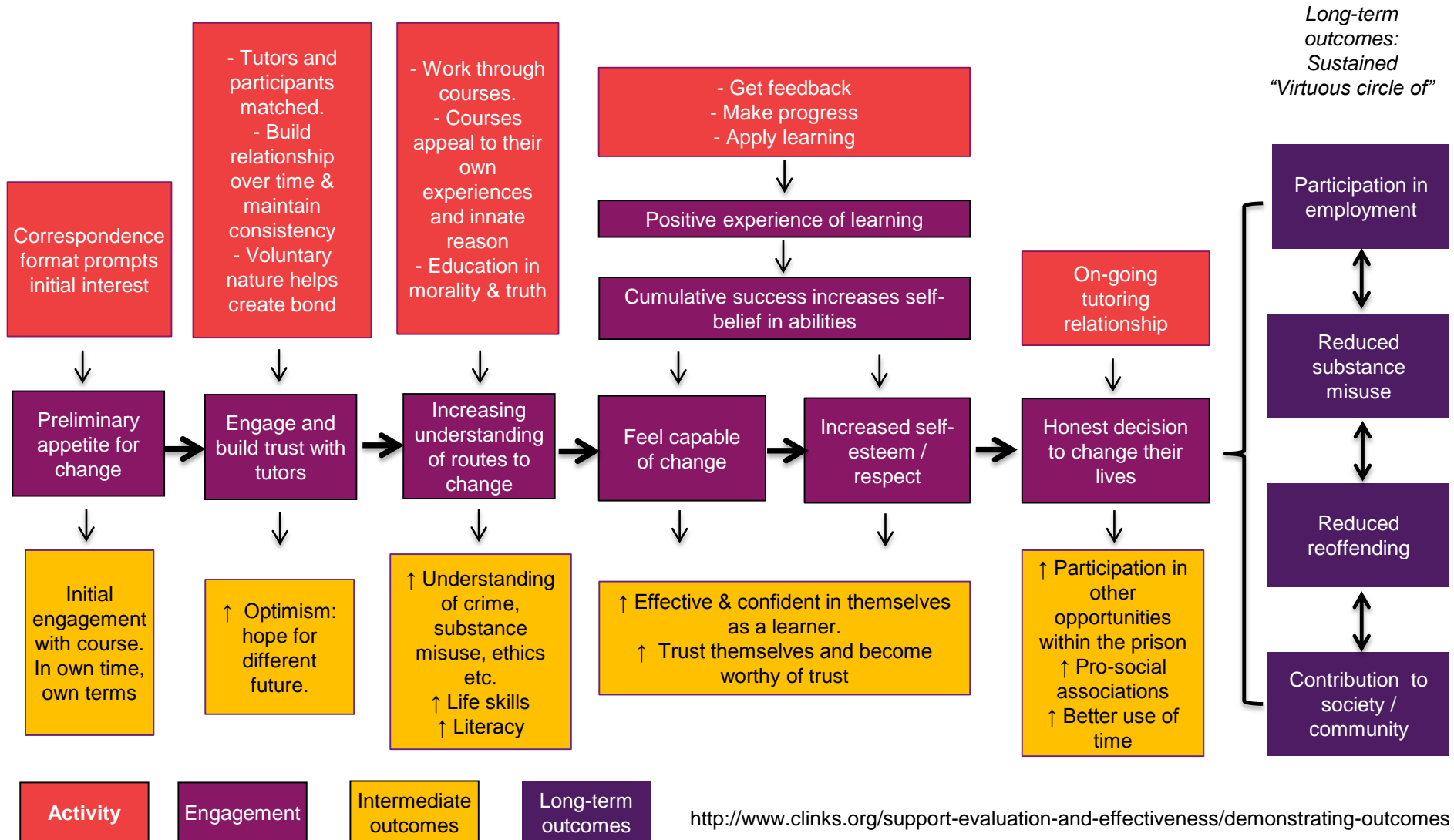
TWO MAIN QUESTIONS

What is the
evidence for
the thing that
we do?

Do we deliver
the thing
effectively?

HIGHER LEVEL OF THEORY OF CHANGE FOR CORRESPONDENCE COURSE

Courses are based on the assumption that those with a history of crime may have missed-out on education in some or all of the following; morals and ethics; literacy; difficulties adapting to conventional learning environments; problems with drug addiction; difficulties with family relationships and an underlying lack of positive identity- which can all contribute to a continuation of the offending cycle.



SOME THEMES FROM THE EXISTING EVIDENCE BASE

What?

- Holistic interventions that address multiple criminogenic needs
- Pro-social recreational activities
- Structured cognitive behavioural programmes
- Practical support
- Individual structured counselling and behaviour programmes

How?

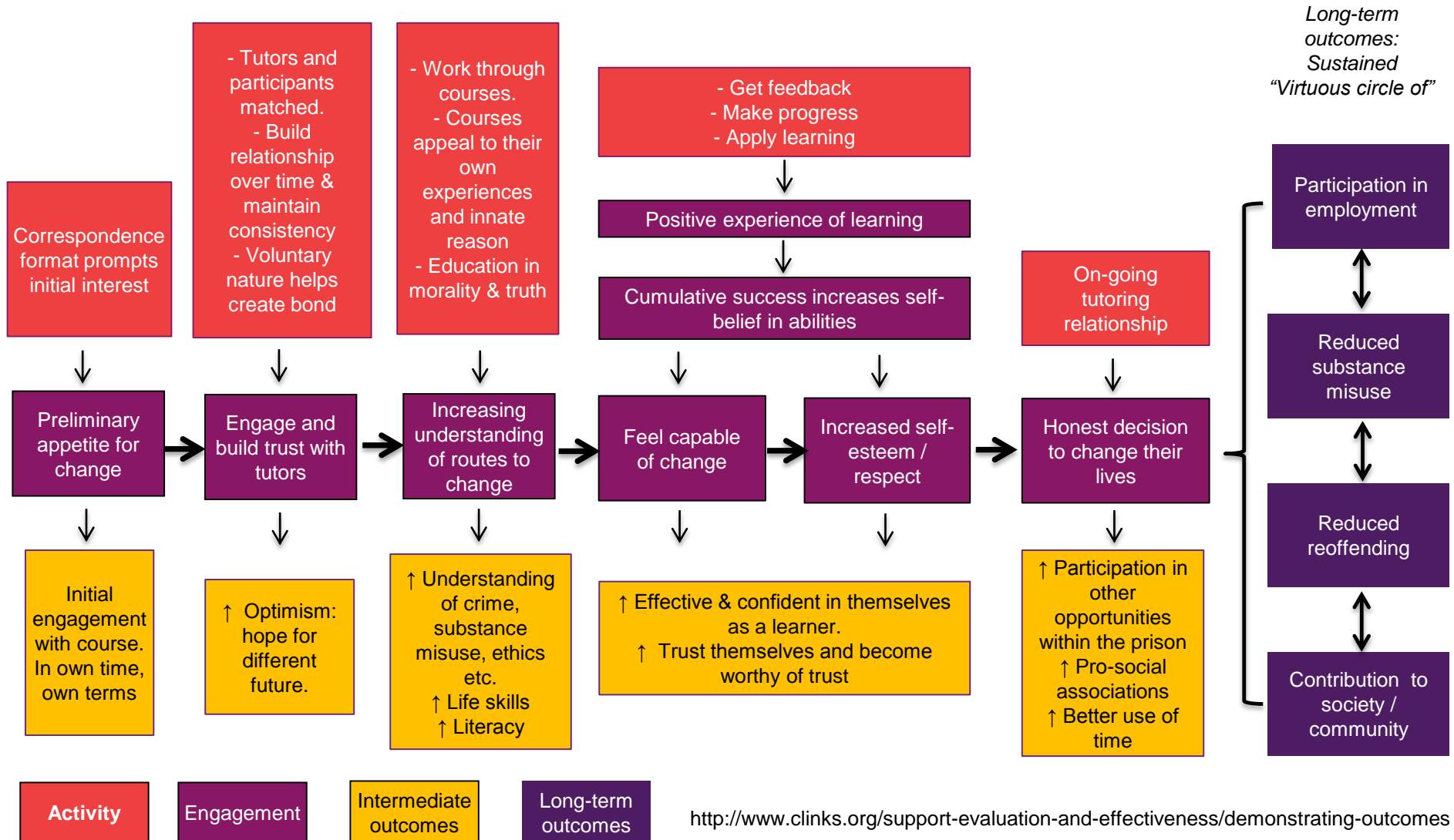
- Risks, needs, responsivity
- Appropriately sequenced services
- Trusting relationship with an offender manager
- Motivational and engaging, rather than control oriented
- Engaging offenders in setting goals for supervision

Intermediate outcomes

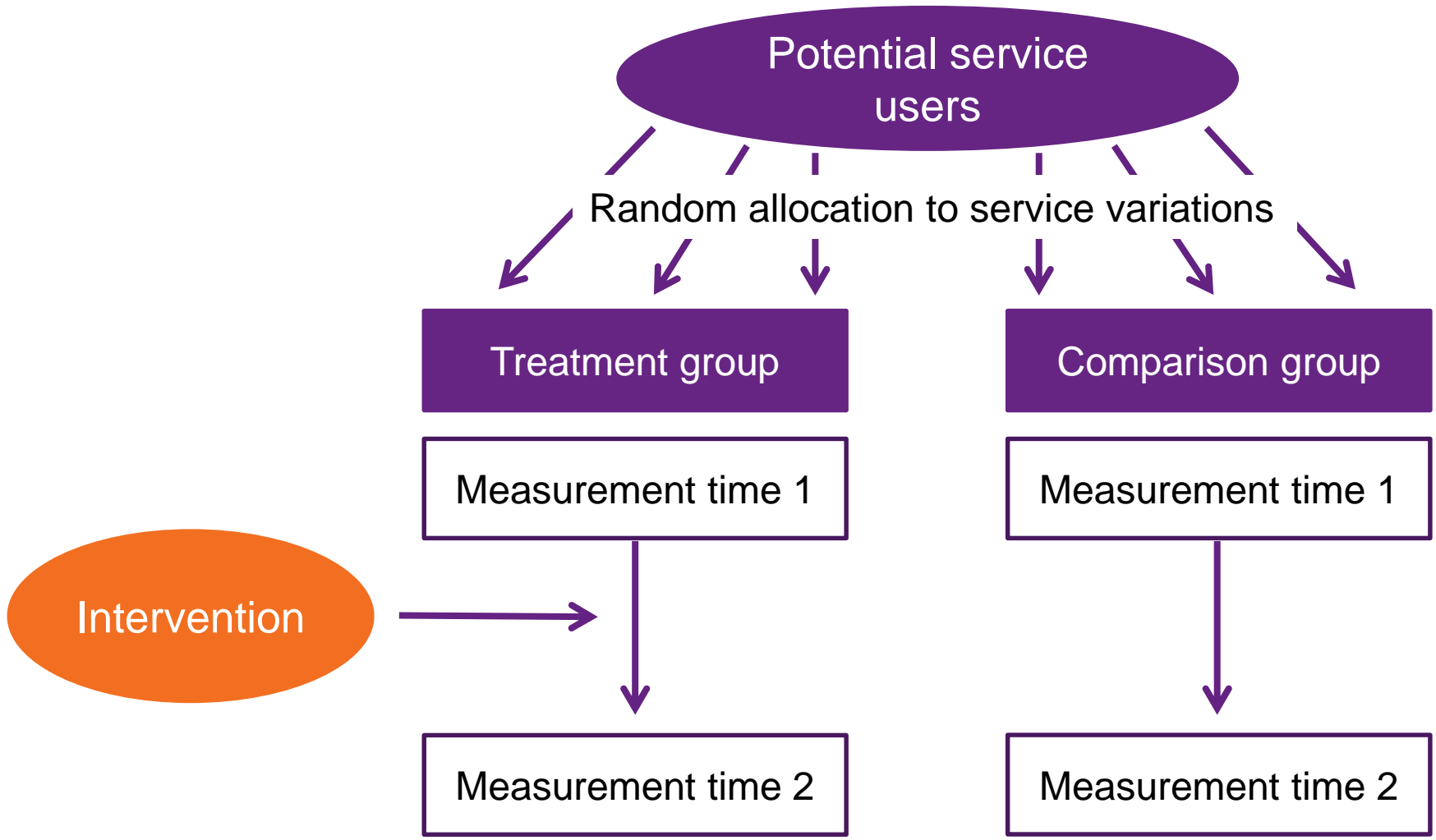
- Personal choice – decision to desist
- Hopeful about giving up crime
- Pro-social problem solving skills
- Reduced alcohol use
- Increased employability
- Improved social/family ties
- Non-criminal peer groups
- Housing stability
- Developing non-criminal identity
- Pro-social lifestyle

HIGHER LEVEL OF THEORY OF CHANGE FOR CORRESPONDENCE COURSE

Courses are based on the assumption that those with a history of crime may have missed-out on education in some or all of the following; morals and ethics; literacy; difficulties adapting to conventional learning environments; problems with drug addiction; difficulties with family relationships and an underlying lack of positive identity- which can all contribute to a continuation of the offending cycle.



THE LOGIC OF EXPERIMENTAL DESIGNS



FOUR KEY POINTS FOR EVIDENCING THE 'THING THAT YOU DO'



1. Develop a good theory of change
2. Use the existing evidence as much as possible
3. Be clear about intermediate outcomes, and look for opportunities to measure them consistently
4. Look for **opportunities** to measure long-term impact – while recognising that this is hard to do

TWO MAIN QUESTIONS

What is the
evidence for
the thing that
we do?

Do we deliver
the thing
effectively?

DEFINE AND MONITOR THE PROCESS OF CHANGE AMONGST SERVICE USERS



1) Stuck / Denial / Despair

• Leave me alone, I'm ok as I am OR I can't change

2) Just managing

• Someone has to help me most of the time. I need prompts and advice to manage

3) Safe and stabilised

• I'm starting to sort things out (benefits, probation appointments, rent, substance use not out of hand)

4) Understand self and possibilities for change

• I can change and I know it's up to me to take responsibility, but I still need some help

5) Raised aspirations and made plans

• I've started to change my behaviour & lifestyle, and take control of my own life

6) Work towards plans with support

• I'm really learning how to do this, and I'll keep on doing it

7) Working Independently

• I have a plan and I'm working towards it without support. Maybe learning / other opportunities. Pro social behaviours.

8) Self-reliant

• I can manage without help. I might even be able to help others

ONLINE DATA COLLECTION



QUALITATIVE RESEARCH



FINAL THOUGHTS

Charities should:

- Use the Justice Data Lab.
- Use the existing evidence base on criminal justice as much as possible. Link their theories of change to what is already known.
- Use Clink’s free resources to inform evaluation design. Be strategic about data collection; collect the minimum from all service users and generalise from more in-depth research.
- Collaborate with other organisations doing similar work. Coordinate data collection, research and analysis.
- Publish and share results. Be honest about failure.

Funders / commissioners should:

- Work together to get more consistency in your requirements. Reduce bureaucracy.
- Reward organisations that do good quality evaluation and are honest about what has / hasn’t worked.
- Fund evaluation projects and synthesis of evidence from charities.

Academics

- Conduct research and write-up results with service providers in mind.
- Make work accessible to the charity sector by publishing it online.

Government

- Be more open to different types of evidence (as long as it is good quality). ‘Clinching’ evidence from counterfactual studies is not all that matters.
- Fund the synthesis of evidence from charities.
- Keep opening-up administrative data for charities to better understand their effectiveness and impact.



SO YOU'VE DONE YOUR EVALUATION – WHAT NEXT? SOME THOUGHTS ON RESULTS FROM THE JUSTICE DATA LAB

Stephen Morris, PERU, Manchester Metropolitan University
s.morris@mmu.ac.uk

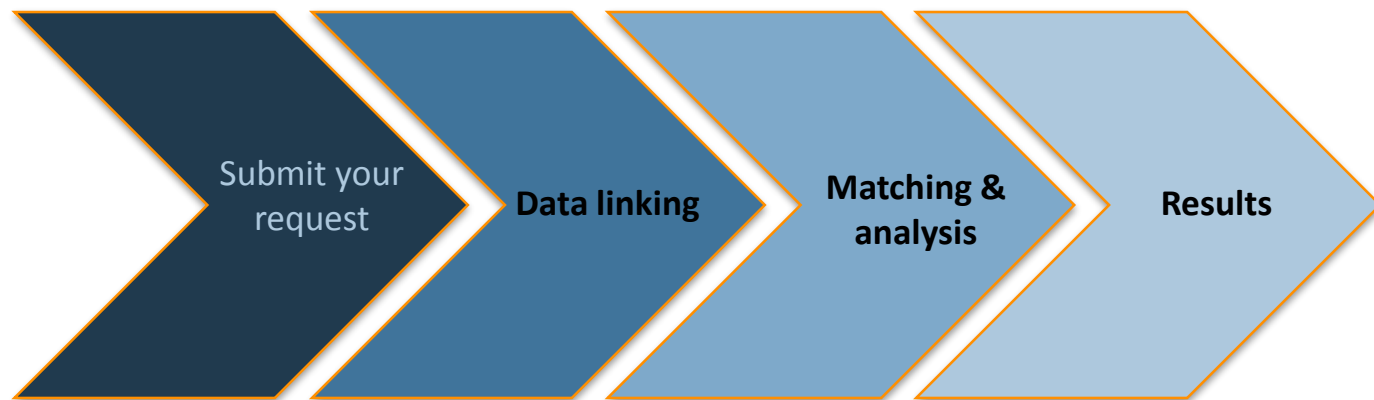
WHAT IS THE JUSTICE DATA LAB?

- Free service targeted at VCS/other providers working with offenders provided by MoJ
- Access to aggregate re-offending rates for the offenders providers have worked with
- Enables providers to understand their impact
- Historic evidence of effectiveness

Some thoughts on what results from this service might mean from the perspective of a provider/commissioner



HOW DOES IT WORK?



- Minimum of 60 cases
- Links your cases to data sources (2002): police national computer, employment & benefits data, OASys, etc.
- Analysis of effectiveness
- Results!



IN BRIEF, WHAT DOES THE DATA LAB DO FOR YOU?

- Provides an estimate of an intervention's effect on the 'one year re-conviction rate' (other related outcomes too!)
- Compares the average reconviction rate for people in an intervention to that of a **similar group** not in the intervention (a 'matched' **control group**)
- Control group represent the reconviction rate for the people in the intervention 'as if' they had not taken part – known as the **counterfactual** – it is an **estimate** only! And it is an **average!**
- Technically the ***Average Effect of Treatment on those Treated?***



MATCHED CONTROL GROUPS & MARYLAND SCALE

([HTTP://WWW.WHATWORKSGROWTH.ORG/PUBLIC/FILES/SCORING-GUIDE.PDF](http://www.whatworksgrowth.org/public/files/scoring-guide.pdf))

Robustness scores (based on an adjusted Maryland Scientific Methods Scale)

Level 1: Either (a) a cross-sectional comparison of treated groups with untreated groups, or (b) a before-and-after comparison of treated group, without an untreated comparison group.

Level 2: Use of adequate control variables and either (a) a cross-sectional comparison of treated groups with untreated groups, or (b) a before-and-after comparison of treated group

Level 3: Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention, and a comparison group used to provide a counterfactual

Better

Level 4: Quasi-randomness in treatment is exploited, e.g. instrumental variables, regression discontinuity, natural experiments

Level 5: Randomisation into treatment and control groups, with Randomised Control Trials (RCTs) providing the definitive example.



MATCHED CONTROL GROUPS & MARYLAND SCALE

([HTTP://WWW.WHATWORKSGROWTH.ORG/PUBLIC/FILES/SCORING-GUIDE.PDF](http://www.whatworksgrowth.org/public/files/scoring-guide.pdf))

Robustness scores (based on an adjusted Maryland Scientific Methods Scale)

Level 1: Either (a) a cross-sectional comparison of treated groups with untreated groups, or (b) a before-and-after comparison of treated group, without an untreated comparison group.

Level 2: Use of adequate control variables and either (a) a cross-sectional comparison of treated groups with untreated groups, or (b) a before-and-after comparison of treated group

Level 3: Comparison of outcomes in treated group after an intervention, with outcomes in the treated group before the intervention, and a comparison group used to provide a counterfactual

Better

Level 4: Quasi-randomness in treatment is exploited, e.g. instrumental variables, regression discontinuity, natural experiments

Level 5: Randomisation into treatment and control groups, with Randomised Control Trials (RCTs) providing the definitive example.

WHAT DO THE RESULTS LOOK LIKE?

- Phoenix Futures' prison based Therapeutic Communities (TC) are self help communities that assist to rebuild the lives of those with substance misuse issues

Analysis	Treatment group	Control group	1 year re-offending rate		Effect	Measures of uncertainty		
			Treatment Group	Control Group		Upper estimate	Lower estimate	P-value
1	93	67,502	39	43	-4	+6%	-14%	0.43
2	93	43,650	39	45	-6	+4%	-16%	0.23
3	93	43,580	39	44	-5	+5%	-15%	0.32

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/506330/phoenix-futures-report.pdf

WHAT DO THE RESULTS LOOK LIKE?

- Phoenix Futures' prison based Therapeutic Communities (TC) are self help communities that assist to rebuild the lives of those with substance misuse issues

Analysis	Treatment group	Control group	1 year re-offending rate		Effect	Measures of uncertainty		
			Treatment Group	Control Group		Upper estimate	Lower estimate	P-value
1	93	67,502	39	43	-4	+6%	-14%	0.43
2	93	43,650	39	45	-6	+4%	-16%	0.23
3	93	43,580	39	44	-5	+5%	-15%	0.32

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/506330/phoenix-futures-report.pdf

WHAT DO THE RESULTS LOOK LIKE?

- Phoenix Futures' prison based Therapeutic Communities (TC) are self help communities that assist to rebuild the lives of those with substance misuse issues

Analysis	Treatment group	Control group	1 year re-offending rate		Effect	Measures of uncertainty		
			Treatment Group	Control Group		Upper estimate	Lower estimate	P-value
1	93	67,502	39	43	-4	+6%	-14%	0.43
2	93	43,650	39	45	-6	+4%	-16%	0.23
3	93	43,580	39	44	-5	+5%	-15%	0.32

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/506330/phoenix-futures-report.pdf

WHAT DO THE RESULTS LOOK LIKE?

- Phoenix Futures' prison based Therapeutic Communities (TC) are self help communities that assist to rebuild the lives of those with substance misuse issues

Analysis	Treatment group	Control group	1 year re-offending rate		Effect	Measures of uncertainty		
			Treatment Group	Control Group		Upper estimate	Lower estimate	P-value
1	93	67,502	39	43	-4	+6%	-14%	0.43
2	93	43,650	39	45	-6	+4%	-16%	0.23
3	93	43,580	39	44	-5	+5%	-15%	0.32

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/506330/phoenix-futures-report.pdf

WHAT DO THE RESULTS LOOK LIKE?

- Phoenix Futures' prison based Therapeutic Communities (TC) are self help communities that assist to rebuild the lives of those with substance misuse issues

Analysis	Treatment group	Control group	1 year re-offending rate		Effect	Measures of uncertainty		
			Treatment Group	Control Group		Upper estimate	Lower estimate	P-value
1	93	67,502	39	43	-4	+6%	-14%	0.43
2	93	43,650	39	45	-6	+4%		0.23
3	93	43,580	39	44	-5	+5%		0.32

Looking for a value < 0.05

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/506330/phoenix-futures-report.pdf

HOW USEFUL ARE THESE TYPES OF ANALYSIS?

- Genuine advance in helping us understand what works
- The methods it uses are technically robust and scientifically credible
- But like all research there are challenges and limitations



WHAT MIGHT BE THE RESULTS OF YOUR JUSTICE DATA LAB EVALUATION?

Assume the goal of the intervention is to prevent re-conviction

- Your intervention *reduces* re-convictions

*Reconviction rates in the intervention group are **lower** than those in the control group, and this difference is statistically significant*

- Your intervention *increases* re-conviction rates

*Reconviction rates in the intervention group are **higher** than those in the control group, and this difference is statistically significant*

- Results are inconclusive

*Either positive or negative diffs in reconviction rates are too small for the analysis to distinguish them from **chance variations***



WHAT DOES IT ALL MEAN? I

Your intervention reduces re-conviction rates – effect is greater than chance - Good news! 😊

But always use evidence with **caution**:

- An average effect
- There is still the issue of bias
- Although we limit statistical uncertainty it is still there
- Good, scientifically credible, **process/qualitative research** is essential:
 - What is it about our intervention that **causes the effect**? Can we **explain** it? What is the ‘active ingredient’?
 - Our intervention ‘works’ but **compared to what**?
 - Would our intervention work **elsewhere**? In different settings? With a different set of clients? At different points in time, Etc.

Commissioners should generally avoid relying on single studies



WHAT DOES IT ALL MEAN? II

**Your intervention increases re-conviction rates - Oh dear!
Evidence of harm 😞**

- We shouldn't too quickly assume the worst
 - Right outcome?
 - Bias & uncertainty
 - Average effect
 - Again process/qualitative research vital:
 - Can we explain why? Is your intervention a good idea poorly **implemented**? Or not such a good idea (theory failure)?
 - Was there some crucial '**supporting factor(s)**' missing
 - **Formative assessment** – what could be done to improve the intervention

Commissioners – look for other studies of similar interventions – what have these shown?



WHAT DOES IT ALL MEAN? III

Your results are inconclusive - In some senses this is the most difficult position

- The average effect is the best estimate
- **Bias** and/or lack of **statistical power**
 - Boost the sample size & re-run the analysis
 - Consider the range of variables used & assumptions made
- ***Small N methods*** might be more appropriate
http://www.3ieimpact.org/media/filer_public/2012/06/29/working_paper_15.pdf
- Magnitude of the intervention effect may not be economically or substantively important – **process/qualitative research** important again!
 - Intervention design (theory of change is it plausible)
 - Implementation failure
 - Supporting or contextual factors



CONCLUDING REMARKS

- Justice Data Lab is a very welcome development
- Use of additional controls in OASys welcome
- But it is only a piece in the evaluation jigsaw
 - Commission a process evaluation along side – qualitative research vital but must be rigorous
- Technical skill and experience required
- No evidence comes without qualification
- Developing effective interventions is some times a slow, incremental process – test, learn, adapt!





**ACADEMY for
JUSTICE COMMISSIONING**

Learn, Network & Share

Any Questions ?

James Noble, Deputy Head of Measurement & Evaluation

New Philanthropy Capital

James.noble@thinknpc.org

Stephen Morris, Policy Evaluation Research Unit

Manchester Metropolitan University

s.morris@mmu.ac.uk



@academyjustice



Academy for Justice Commissioning