



**GOVERNMENT  
OUTCOMES  
LAB**

# Setting and Measuring Outcomes Academy for Social Justice Commissioning

10<sup>th</sup> June 2019



@ukgolab  
@AcademyJustice

# Introducing us



**Andreea Anastasiu, Senior Policy  
Engagement Officer, GO Lab**



**Neil Stanworth, GO Lab Fellow of Practice  
& Director, ATQ Consultants**

# About the GO Lab



Joint partnership  
between UK  
Government &  
Oxford University

Established in 2016

Based at the  
Blavatnik School of  
Government, in  
Oxford

Centre of academic research and practice with a mission to improve the provision of public services to tackle complex social issues, with a focus on outcome based models

# Understanding what works



# Support for commissioners

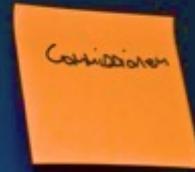
Knowledge Hub for  
SIBs & OBC

[golab.bsg.ox.ac.uk](http://golab.bsg.ox.ac.uk)

Guides & resources



SIB  
Readiness  
Framework



SIB  
projects  
database

HOW TO  
GUIDE  
CONTRACTING  
AND GOVERNANCE

Advice  
surgeries



Events &  
workshops



Webinars



Regional  
SIB  
Knowledge  
Clubs



Fellows of  
Practice

# Session overview

- 12.00 *Welcome remarks*
- 12.10 **Session 1 – why measure outcomes?**
- 12.45 **Session 2 – identifying outcomes**
- 13.45 *Lunch*
- 14.30 **Session 3 – measuring outcomes**
- 15.30 *Break*
- 15.45 **Session 4 – setting outcomes targets**
- 17.00 *Close*



**GOVERNMENT  
OUTCOMES  
LAB**

# 1: Why measure outcomes?



# Why focus on outcomes?

...understand wider benefits

... measure / evaluate success

... engage citizens

... galvanise collaboration

... provide a 'north star' for staff

... make a business case

... understand root causes

... kick-start innovation

# Case study – West London Zone



## Outcomes Framework

		Emotional and mental wellbeing	Positive relationships	Confidence and aspiration	Progress at school
		<i>'I feel good about myself'</i>	<i>'I get on fine with others'</i>	<i>'I am confident and want to do well'</i>	<i>'I have progressed well at school'</i>
Primary school	End of Two Year Programme	Strengths and Difficulties Questionnaire (SDQ)	SDQ peer relationship sub-scale	School attendance; SDQ emotional sub-scale; WLZ measure based on Brofenbrenner	On track to meet age-related expectations by the end of primary school
	End of primary school			School attendance; SDQ emotional sub-scale; Link Worker or teacher assessment re. confident for transition	KS2 results 'expected standard' in Reading, Writing and Maths
Secondary school	End of Two Year Programme	SDQ; Warwick Edinburgh Mental Wellbeing Scale	SDQ peer relationship sub-scale; Communities That Care	School attendance; Communities That Care	On track to achieve age-related expectations by the end of secondary school
	End of secondary school			School attendance; Link Worker and teacher progression planning survey for KS5	Level 4 GCSE in English and Maths; enrolled in full-time Education, Employment or Training

# Case study – West London Zone



## Outcomes Framework

Emotional and mental wellbeing	Positive relationships	Confidence and aspiration	Progress at school
<i>'I feel good about myself'</i>	<i>'I get on fine with others'</i>	<i>'I am confident and want to do well'</i>	<i>'I have progressed well at school'</i>

## Outcomes focus enables:

- Better joint understanding of child's needs
- Child's voice in their own support plan
- Multiple partners across sectors to galvanise around desired outcomes
- A framework for measuring and evaluating success
- A way for the Local Authority to target its spending

Primary school	End of Two Year Programme	<p>On track to meet age-related expectations by the end of primary school</p> <hr/> <p>KS2 results 'expected standard' in Reading, Writing and Maths</p> <hr/> <p>On track to achieve age-related expectations by the end of secondary school</p>					
	End of primary school						
Secondary school	End of Two Year Programme				SDQ; Warwick Edinburgh Mental Wellbeing Scale	SDQ peer relationship sub-scale; Communities That Care	Communities That Care
	End of secondary school						School attendance; Link Worker and teacher progression planning survey for KS5
		Level 4 GCSE in English and Maths; enrolled in full-time Education, Employment or Training					



## Scotland's National Performance Framework

- 11 national outcomes
- 81 indicators (range of economic, social & environmental)
- Performance reported openly and transparently <https://nationalperformance.gov.scot>

# Outcomes for customers – adult social care

- Emerging focus on outcome-based practice
- ‘Outcomes for customers’ – different emphasis and meaning across councils (and therefore ways in which they are used)

*1. Those outcomes that are demonstrated by the overall satisfaction of the customer*

*2. Outcomes that are defined by the Department of Health (England) in their Adult Social Care Outcomes Framework (ASCOF)*

*3. User-defined outcomes – personalisation; ‘nothing about me without me’*

*4. Outcomes that are defined by the way in which a person has been helped to gain or regain their independence.*

# DWP Rate Card

DWP pays for one or more outcomes per participant which can be linked to improved employability. A definitive list of outcomes and maximum prices DWP was willing to pay for Round 2 is as follows:

<b>Nature of Outcome</b>	<b>Maximum Price of Outcome</b>
Improved attitude towards school	£700
Improved behaviour	£1300
Improved attendance	£1400
Entry Level Qualification	£900
NVQ level 1 or equivalent	£1100
NVQ level 2 or equivalent	£3300
NVQ level 3 or equivalent	£5100
Entry into employment	£3500
Sustained Employment	£2000

The Maximum amount payable per individual is £11,700. The figure is based on 3 years of Annually Managed Expenditure (AME) savings.

DWP rate card for Social  
Impact Bond (SIB) projects

# How not to use an outcomes focus



# Outcomes vs Impact

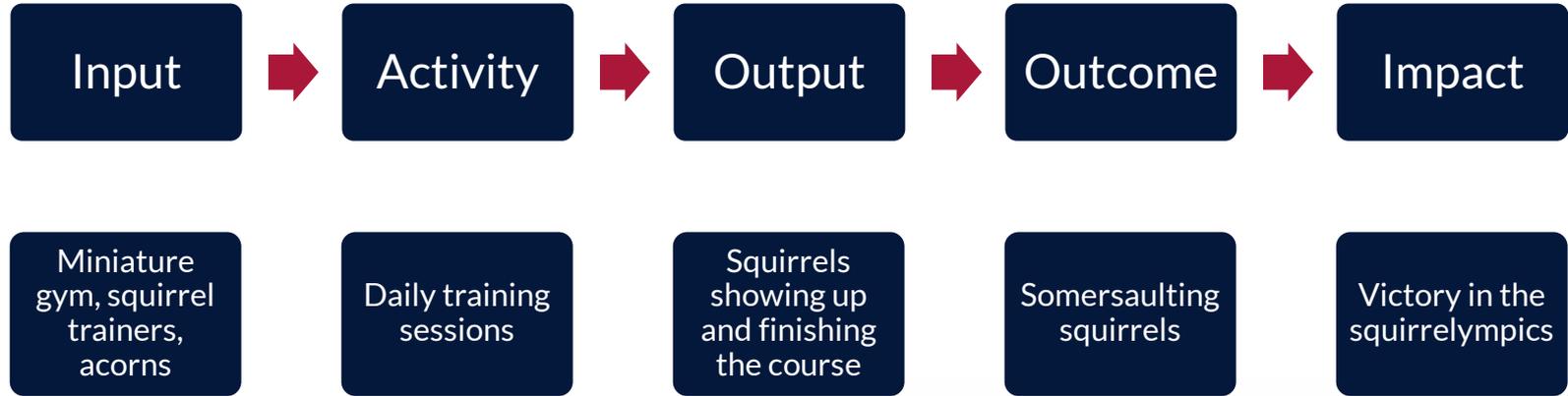


Impact measurement answers that “what would have happened anyway?” question

SSIR - Ten Reasons Not to Measure Impact—and What to Do Instead

[https://ssir.org/articles/entry/ten\\_reasons\\_not\\_to\\_measure\\_impact\\_and\\_what\\_to\\_do\\_instead](https://ssir.org/articles/entry/ten_reasons_not_to_measure_impact_and_what_to_do_instead)

# Logic model / theory of change



# Key concepts

<b>Outcome</b>	What changes for an individual (or other defined unit, such as a family) as the result of a service or intervention (e.g. improved learning outcomes)
<b>Measure / Indicator</b>	The specific way the commissioner chooses to determine whether that outcome has been achieved (e.g. a test score)
<b>Target / Metric / Trigger</b>	The specific value attached to the measure for the purposes of determining whether satisfactory performance has been achieved (e.g. a test score of 95 out of 100 or improvement of 30 points in a test score over a 5 month period).

# Outcomes, measures, targets

## Example: Youth unemployment reduction

<b>Outcome</b>	The young person is in employment
<b>Outcome measure</b> <i>(also termed an indicator)</i>	Confirmation from the employer that the person is employed by them
<b>Outcome target</b> <i>(also termed metrics or triggers)</i>	The young person is in continuous employment of a minimum of 16 hours per week for a defined period or That 20% of the total cohort are in continuous employment for a defined period on average

# Outcomes, measures, targets

## Example: Emotional wellbeing

<b>Outcome</b>	Improved emotional wellbeing of young people
<b>Outcome measure</b> <i>(also termed an indicator)</i>	An identifiable improvement in young people's resilience and ability to deal with challenges using the Strengths and Difficulties Questionnaire (SDQ)
<b>Outcome target</b> <i>(also termed metrics or triggers)</i>	The young person reduces their total SDQ score by a defined number over a specified period or that there is a mean reduction in the average score across the cohort as a whole

# Developing an outcomes framework

An outcomes framework should include:

- ✓ the outcomes to be used;
- ✓ the measures to be applied to each outcome;
- ✓ the specific targets to be applied to each measure, that determine the desired level of success; and
- ✓ when measurement takes place.

An outcomes framework involves:

- ✓ An iterative and progressive process
- ✓ Negotiation & consensus - building agreement among the contracting parties
- ✓ Review and revision (if necessary)



# GOVERNMENT OUTCOMES LAB

## Questions





**GOVERNMENT  
OUTCOMES  
LAB**

## 2: Identifying outcomes



@ukgolab

# Identifying outcomes

## Some general considerations

- Should you use ‘hard’ (objective) or ‘soft’ (subjective) outcomes and measures’?
- Do you need to use proxy measures for outcomes that cannot be measured directly
- How do you decide which measures are appropriate for your service?

# Hard & soft outcomes and measures

Type of outcome	Definition	Advantages	Disadvantages
Hard	Can be measured objectively	<ul style="list-style-type: none"><li>▪ Simpler to measure</li><li>▪ No risk of disagreement about achievement</li></ul>	<ul style="list-style-type: none"><li>▪ Not always available</li><li>▪ May not capture sustained impact</li><li>▪ May not reflect what matters to service users</li></ul>
Soft	Requires subjective assessment	<ul style="list-style-type: none"><li>▪ Useful when no hard outcome is available</li><li>▪ Can be used to test progress made</li><li>▪ Measures whether the service meets user expectations</li></ul>	<ul style="list-style-type: none"><li>▪ Consistency of measurement can be difficult</li><li>▪ Potential for disagreement about achievement</li></ul>

# Proxy outcomes & measures

An indirect measure of the desired outcome strongly correlated to that outcome, used when direct measures of the outcome are unavailable or cannot be measured.

# Proxy outcomes & measures

## Proxy measure

- Reduced hospital admissions
- Off benefit
- Reduced reconvictions

## True outcome

- Improved health
- In employment
- Reduced reoffending

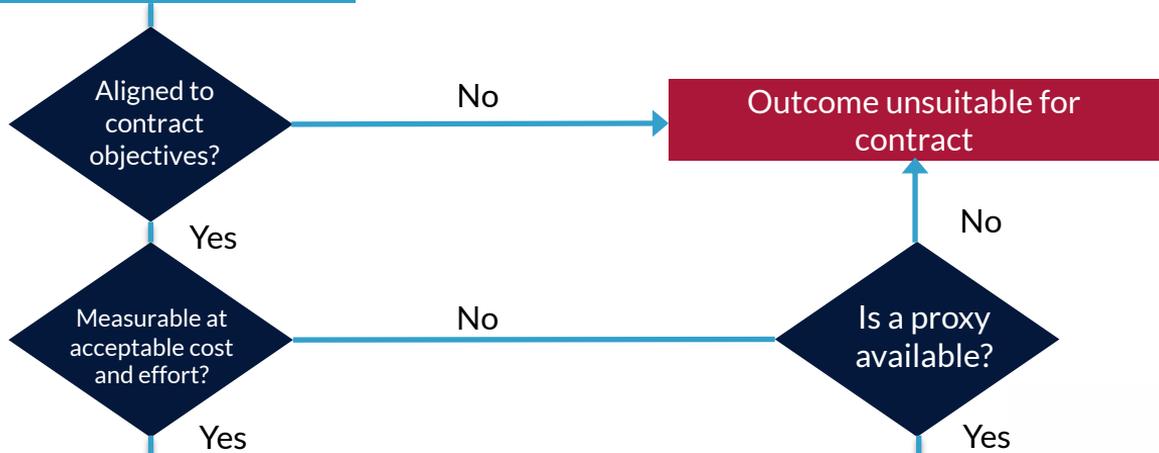
## Risks

- Reduced attendance might be due to other factors
- Person may cease claiming benefits without finding work
- Many offences go undetected

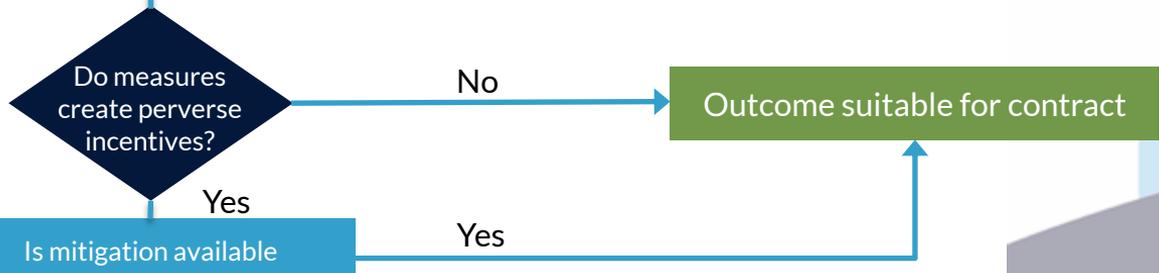
# Identifying outcomes

## A simplified decision-making tree

Identify possible outcome



Identify requirements for measurement and consider targets that can be used





**GOVERNMENT  
OUTCOMES  
LAB**

## Group exercise



# Exercise 1a – A possible answer

Outcome	For	Against	Top 4
Improved user well - being	<ul style="list-style-type: none"> <li>Directly reflects user experience</li> <li>Measures distance travelled</li> </ul>	<ul style="list-style-type: none"> <li>Requires subjective assessment</li> <li>Harder to agree a specific target/metric</li> </ul>	
Fewer GP appointments	<ul style="list-style-type: none"> <li>Good measure of impact on primary care</li> <li>Good proxy for improved self-management</li> </ul>	<ul style="list-style-type: none"> <li>No impact on primary costs</li> <li>Some data collection challenges</li> </ul>	
Reduced A&E attendance	<ul style="list-style-type: none"> <li>OK measure of impact on secondary care</li> <li>Reasonable proxy for improved SM</li> </ul>	<ul style="list-style-type: none"> <li>Low/no impact on secondary costs/demand</li> <li>Low attribution to intervention</li> </ul>	
Fewer OP appointments	<ul style="list-style-type: none"> <li>Reasonable proxy for improved SM</li> <li>Some impact on secondary costs/demand</li> </ul>	<ul style="list-style-type: none"> <li>Low/no impact on secondary costs/demand</li> <li>Low attribution to intervention</li> </ul>	
Reduced hospital admissions	<ul style="list-style-type: none"> <li>Good measure of impact on 2ndary care</li> <li>More impact on secondary costs/demand</li> </ul>	<ul style="list-style-type: none"> <li>May not work for all conditions</li> <li>Susceptible to variations in clinical practice</li> </ul>	
Reduced cost of admissions	<ul style="list-style-type: none"> <li>Good measure of impact on 2ndary care</li> <li>Direct measure of impact on 2ndary costs</li> </ul>	<ul style="list-style-type: none"> <li>Variable impact by condition/practice</li> <li>Costs linked to NHS tariff not intervention</li> </ul>	
Reduced prescription of SSRIs	<ul style="list-style-type: none"> <li>Measures impact on primary care</li> <li>OK proxy for improved well-being</li> </ul>	<ul style="list-style-type: none"> <li>Susceptible to variations in GP practice</li> <li>Low prescription costs so likely low impact</li> </ul>	
Reduced use of adult social care	<ul style="list-style-type: none"> <li>Only measure of impact on social care</li> <li>Only measure relevant to LA commissioner</li> </ul>	<ul style="list-style-type: none"> <li>Difficult to set an appropriate metric</li> <li>Data collection challenges</li> </ul>	

# Exercise 1b – A possible answer

Outcome	For	Against	Top 4
Improved user well – being and resilience	<ul style="list-style-type: none"> <li>Directly reflects user experience</li> <li>Measures distance travelled</li> </ul>	<ul style="list-style-type: none"> <li>Requires subjective assessment</li> <li>Harder to agree a specific target/metric</li> </ul>	
No offences committed	<ul style="list-style-type: none"> <li>Easy to measure (but not directly)</li> <li>Directly impacts crime and offending</li> </ul>	<ul style="list-style-type: none"> <li>Perverse incentive if any offence committed</li> <li>No value for those only at risk of offending</li> </ul>	
Fewer offences committed	<ul style="list-style-type: none"> <li>Easy to measure/direct impact as above</li> <li>No perverse incentive if offence committed</li> </ul>	<ul style="list-style-type: none"> <li>Low/no impact on secondary costs/demand</li> <li>Potentially low attribution to intervention</li> </ul>	
Improved in – school behaviour	<ul style="list-style-type: none"> <li>Good lead indicator of propensity to crime</li> <li>Potential proxy for improved well-being</li> </ul>	<ul style="list-style-type: none"> <li>Hard to define behaviour objectively</li> <li>Requires school cooperation to measure well</li> </ul>	
Improved school attendance	<ul style="list-style-type: none"> <li>Potential proxy for improved well-being</li> <li>Easy to measure and collect data</li> </ul>	<ul style="list-style-type: none"> <li>May not be attributable to intervention</li> <li>Needs careful target setting</li> </ul>	
Progression to post-16 E and E	<ul style="list-style-type: none"> <li>Good measure of long – term impact</li> <li>Direct impact on societal costs</li> </ul>	<ul style="list-style-type: none"> <li>May not be attributable to intervention</li> <li>Hard for provider to influence</li> </ul>	
Fewer exclusions from school	<ul style="list-style-type: none"> <li>Easy to measure</li> <li>OK proxy for improved behaviour</li> </ul>	<ul style="list-style-type: none"> <li>Perverse incentive if exclusion occurs</li> <li>No value for those not at risk of exclusion</li> </ul>	
Improved educational attainment	<ul style="list-style-type: none"> <li>Good proxy for long – term impact</li> <li>Easy to measure using some indicators</li> </ul>	<ul style="list-style-type: none"> <li>Relatively narrow measure of success</li> <li>Limited value for some service users</li> </ul>	



GOVERNMENT  
OUTCOMES  
LAB

## Questions





**GOVERNMENT  
OUTCOMES  
LAB**

## **3: Measuring outcomes**



# Measuring outcomes

## Some general considerations

- ✓ which measures (or combination of measures) are appropriate
- ✓ whether the data needed to measure the outcome are already collected for another purpose, for example government statistical returns or internal performance management;
- ✓ if not, whether the data collection requires significant investment in new collection processes and systems;
- ✓ who will be responsible for collecting the data and whether they have the capacity to do so; and
- ✓ whether the data need to be independently checked and validated

# Key issues and questions

- Hard vs. soft measures
- Binary vs. continuous measures
- Use of lead / progression measures
- Data availability

# Hard & soft outcome measures

## Example: reducing entry to local authority care

The Turning the Tide Social Impact Bond (SIB), commissioned by North Somerset Council, measures reductions in the number of older children entering local authority care as a result of family dysfunction and stress.

It uses a combination of:

- **A soft measure** of improved family / parental capacity and resilience, measured through [Triangle Consulting's Family Plus™ Star](#); and
- **A hard outcome measure** relating to the length of time children at risk of entering care are able to remain with their families.

# Binary and continuous measures

Type of outcome	Employment	Offending	Child Protection
Binary	In work/not in work	Not convicted/convicted	Looked after/not looked after
Continuous	Length of time in work Quality of job Progression to full time work	Reduced frequency Reduced severity	Length of time not in care Escalation/de-escalation to formal Child Protection

# Lead and final outcomes

## Lead/indicative outcome

- Improved school attendance and/or behaviour
- Engagement in part time or voluntary work
- Family attendance at a parenting support programme.

May lead to



## True final outcome

- Improved attainment and reduced risk of exclusion
- Full time employment
- Reduced risk of a child becoming 'in need' or on a Child Protection Plan.

# Data collection options

Data type	Pros	Cons
Administrative data	<ul style="list-style-type: none"><li>• Highly accurate</li><li>• Low cost</li></ul>	<ul style="list-style-type: none"><li>• May not exist</li><li>• May not cover population of interest</li><li>• May not directly address question of interest</li></ul>
Primary data	<ul style="list-style-type: none"><li>• Directly addresses question of interest</li></ul>	<ul style="list-style-type: none"><li>• High cost</li><li>• Possibility of bias</li></ul>

- ✓ Find out whether the required data is already collected for other purposes
- ✓ Do not to make assumptions about the availability of data from other parties or the ability of those parties to collect data on your behalf.



GOVERNMENT  
OUTCOMES  
LAB

## Questions





**GOVERNMENT  
OUTCOMES  
LAB**

## Group exercise



@ukgolab

# Exercise 2a – A possible answer

Outcome	Possible metrics	Best and why
Improved user well - being	▪ Improvement in WEMW scale	
	▪ Achievement of pre-defined WEMW score	
	▪ User self-reported improvement	
Fewer GP appointments	▪ Reduction in number of appointments	
	▪ Appointments below pre-defined number	
	▪ No GP appointments in period	
Reduced hospital admissions	▪ Reduction in number of planned admissions	
	▪ Reduction in number of un-planned admissions	
	▪ Admissions below pre-defined frequency	
Reduced use of adult social care	▪ Reduction in number of referrals to ASC	
	▪ Reduction in number of visits by ASC	
	▪ No visits by ASC in period	

# Exercise 2b – A possible answer

Outcome	Possible metrics	Best and why
Improved user well – being and resilience	▪ Improvement in WMEW scale	
	▪ Achievement of pre-defined WMEW score	
	▪ User self-reported improvement	
Improved in – school behaviour	▪ Behaviour improves to level agreed by school	
	▪ No exclusions or other disciplinary action	
	▪ Reduction in disciplinary incidents	
Improved school attendance	▪ Attendance improves to school average	
	▪ Attendance improves against current level	
	▪ No unauthorised absences	
Improved educational attainment	▪ Achievement of GCSE grades 4-9	
	▪ Achievement of NVQ level 3 or equivalent	
	▪ Improvement in expected attainment	



**GOVERNMENT  
OUTCOMES  
LAB**

## **4: Setting outcomes targets and outcomes-based payment**



# Setting targets/payments

## Some general considerations

- Should the target/payment be across the cohort as a whole or linked to individual outcome achievement. This depends on:
  - whether you know what would happen without the intervention – the ‘counterfactual’;
  - whether you know and can adjust for the ‘deadweight’?
- Which measures should you set targets for (or link to payment);
- How do you decide what is achievable; and
- How do you avoid perverse incentives for all parties to the contract?

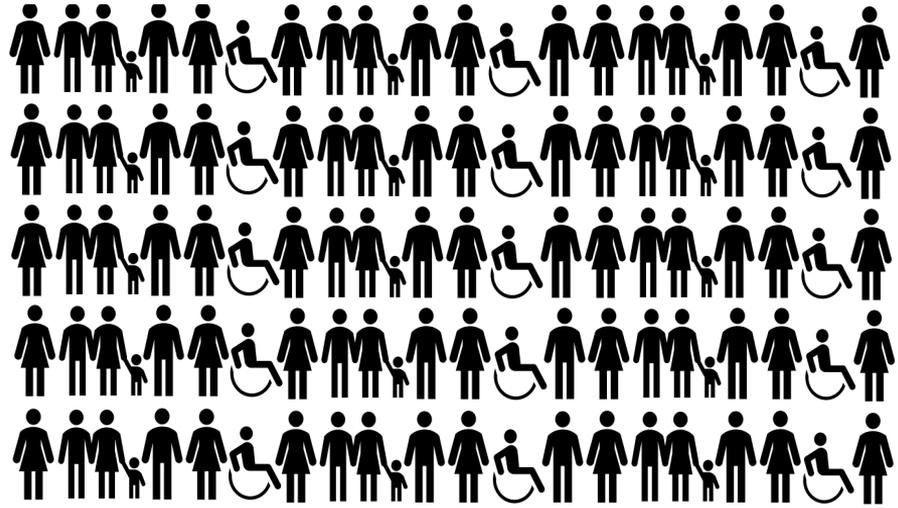
# Cohort v individual measurement

Cohort outcome measurement	Individual outcome measurement
<ul style="list-style-type: none"><li>▪ Works best when the current adverse outcomes vary across the cohort</li><li>▪ Usually requires comparison with a group who did not receive the intervention (the counter-factual)</li><li>▪ Does not normally require a separate calculation of deadweight</li></ul>	<ul style="list-style-type: none"><li>▪ Works best when the cohort are experiencing similar adverse outcomes</li><li>▪ Does not usually involve a comparison group or other baseline</li><li>▪ Requires good evidence of the likely level of deadweight</li></ul>

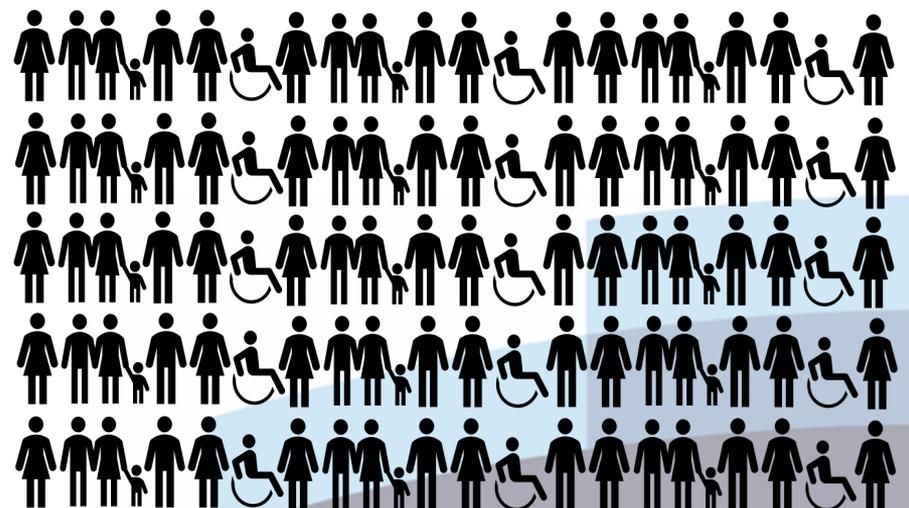
# Cohort measurement

What difference did the intervention make?

## Treatment group



## Comparison group



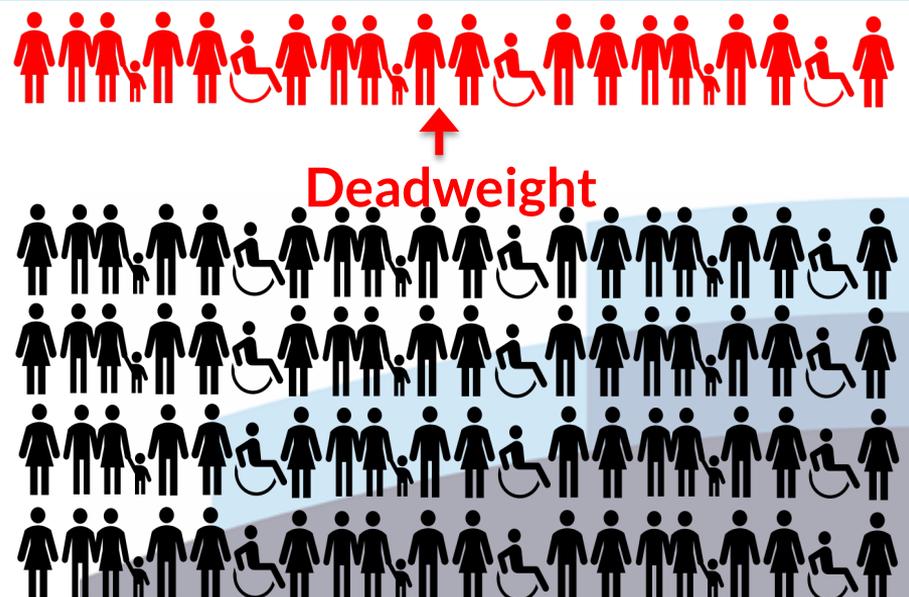
# Cohort measurement

What difference did the intervention make?

Treatment group – 40% difference



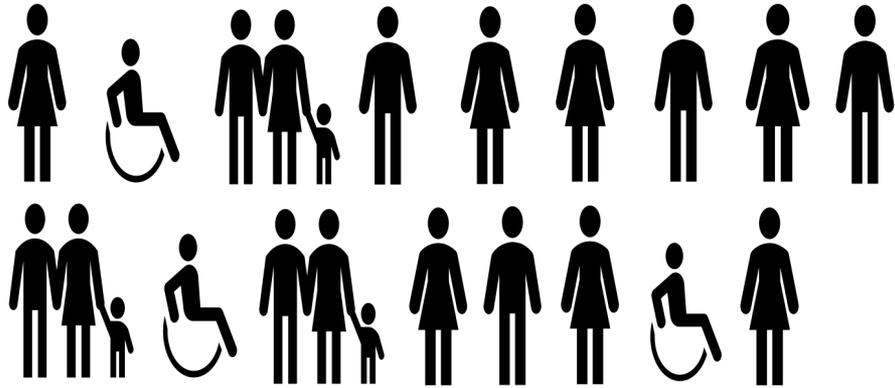
Comparison group – 20% difference



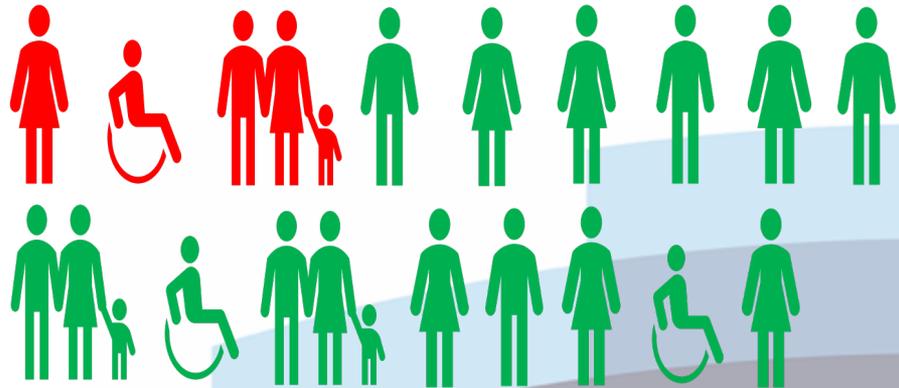
# Individual measurement

How many people achieved the outcome?

20 beneficiaries achieved the outcome



16 outcomes is the impact net of deadweight of 20%



# Cohort v individual measurement

## Examples

### HMP Peterborough SIB

- 7.5% reduction in reoffending across all SIB cohorts, against a national comparison group compiled using Propensity Score Matching
- 10% reduction in individual annual cohorts



### DWP Innovation Fund

NATURE OF OUTCOME	MAXIMUM PRICE OF OUTCOME
Improved attitude towards school	£700
Improved behaviour	£1,300
Improved attendance	£1,400
Entry level qualification	£900
NVQ Level 1 or equivalent	£1,100
NVQ Level 2 or equivalent	£3,300
NVQ Level 3 or equivalent	£5,100
Entry into employment	£3,500
Sustained employment	£2,000

The maximum amount payable per individual is £11,700. The figure is based on 3 years of Annually Managed Expenditure (AME) savings.

# Setting outcome targets

- Specify the level of against which success will be judged / outcome payments will be made
- Must align with the social objectives of the contract (what does good look like?)
- Should take account of existing evidence around the effectiveness of the intervention and the effect of deadweight
- Could be as simple as the achievement of a hard binary outcome
- But should usually also consider:
  - Progression towards the main outcome
  - The sustainment of the outcome and/or further improvement

# Achievability

- Find out how successful others have been in the past when attempting to achieve the same (or similar outcome)
- Ask providers what they think is achievable
- Use a combination of lead / progression targets which are easier to achieve, and stretch targets which are harder

# Perverse incentives

**Perverse incentives are incentives that encourage contract stakeholders to behave in a way that is detrimental to contractual goals even if some outcome metrics improve**

## Some examples:

- Measures and metrics that link to absence of or reduction in referrals to a statutory agency
- A simplistic binary outcome measure, which can easily be 'failed' by a high proportion of the cohort – this can lead to 'parking' of those who can no longer achieve the outcome
- Setting a metric that is achieved with varying degrees of effort for different members of the cohort – this can lead to creaming (focussing the attention on those easier to help)
- A single time-related milestone at which the person receiving the information is deemed to have achieved a positive outcome – sometimes termed cliff edge

# How we have learnt to set better outcomes

## Example: children on the edge of care

Metrics relating to the prevention of entry to care have evolved from:

- Cohort-wide measurement of the total number of days in care compared to a baseline or comparison group (complex and time-consuming to measure)
  - *via*
- Bullet payments paid as an individual tariff at intervals (e.g. 6 months) reflecting the length of time out of care (simpler but prone to distortion or perverse incentives) *to*
  - *to*
- A payment per individual for each care day avoided, totted up and paid at intervals (combining the advantages of a tariff with flexibility and avoidance of distortion)
- A soft measure is often used alongside this e.g. the Family Star



# GOVERNMENT OUTCOMES LAB

## Questions





# GOVERNMENT OUTCOMES LAB

## Group exercise



# Exercise 3a – A possible answer

Outcome	Cohort measure	Individual measure	Stretch/sustainment measures
Improved user well - being	%age improvement in average score on Warwick-Edinburgh Mental Well-being scale	%age improvement in individual score on Warwick-Edinburgh Mental Well-being scale	Improvement targets set and measured at 6 and 12 months
			Improvement targets set and measured at 12 and 24 months
Fewer GP appointments	%age reduction in total number of GP appointments compared to previous number or a cohort of similar users elsewhere	Service user reduces the number of appointments made by X% compared to previous year	Reduction level maintained year on year
			Further reductions in second and subsequent years
Reduced hospital admissions	Total number of unplanned admissions reduces by an average of X per person over 2 years compared to previous average or average for cohort of similar users	Service user has X fewer unplanned admissions per year	Number of planned admissions also reduce
			Unplanned admissions continue to reduce in subsequent years
Reduced use of adult social care	%age reduction in total referrals requiring an assessment by adult social care over two years compared to previous level	Each service user has X fewer referrals requiring an assessment by adult social care each year	Reduction in referrals maintained in third year
			Further reductions in third and subsequent years

# Exercise 3b – A possible answer

Outcome	Cohort measure	Individual measure	Stretch/sustainment measures
Improved user well – being and resilience	%age improvement in average score on Warwick-Edinburgh Mental Well-being scale	%age improvement in individual score on Warwick-Edinburgh Mental Well-being scale	Improvement targets set and measured at 6 and 12 months
			Improvement targets set and measured at 12 and 24 months
Improved in - school behaviour	%age of all pupils achieving acceptable behaviour level across one year	Pupil improves behaviour to acceptable level within one year	Improvement in behaviour sustained in second and subsequent years
			Behaviour improves further to level defined by school as 'good'
Improved school attendance	%age of all pupils reducing unauthorised absence by defined amount	Pupil reduces unauthorised absence by defined amount	Further reduction in unauthorised absence in subsequent years
			Attendance improves to average level of all pupils in school
Improved educational attainment	%age of all pupils achieving or exceeding attainment targets set in plans	Pupil achieves or exceeds attainment targets set in plans	Pupils exceed attainment plan
			Pupils achieve defined number of GCSE Grades 4-9

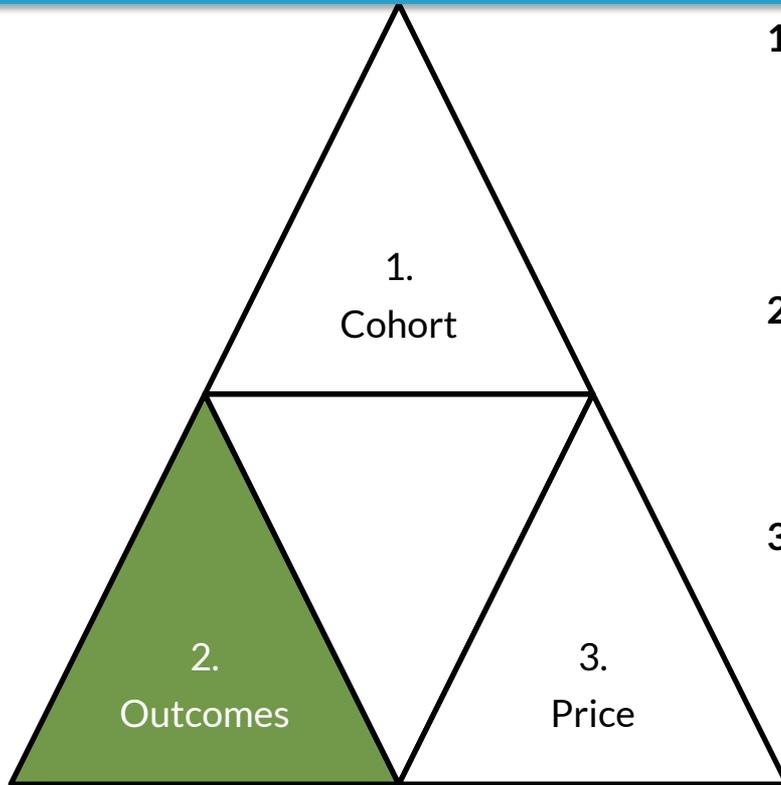
# Three top tips for commissioners (and others)

- ✓ Simple is not always best
- ✓ Avoid superficially attractive hard and binary metrics that can cause major problems

- ✓ Be flexible
- ✓ Measures and metrics will change as you develop your framework and talk to stakeholders

- ✓ Learn from others' experience
- ✓ Use existing metrics where they are available – they save time & effort

# Designing a robust framework



- 1. Tightly defined eligible cohort**
  - Clear, objective criteria
  - Understanding of how far participants are from the desired outcomes
  - Independent referral / identification mechanism
- 2. Alignment between payable outcomes and policy objectives**
  - Logical link between activity, outputs and outcomes
  - Adequate period of time for tracking
  - A way to tell if the effect has 'stuck'
- 3. Accurate price-setting of outcomes**
  - Robust estimate of likely level of benefit vs what would happen anyway ("deadweight")
  - A way to get confidence that any outcomes are caused by the intervention ("attribution")
  - Suitably long outcome tracking-period

Whilst it is not practical for these aspects to be perfect, commissioners should focus on them to avoid perverse incentives

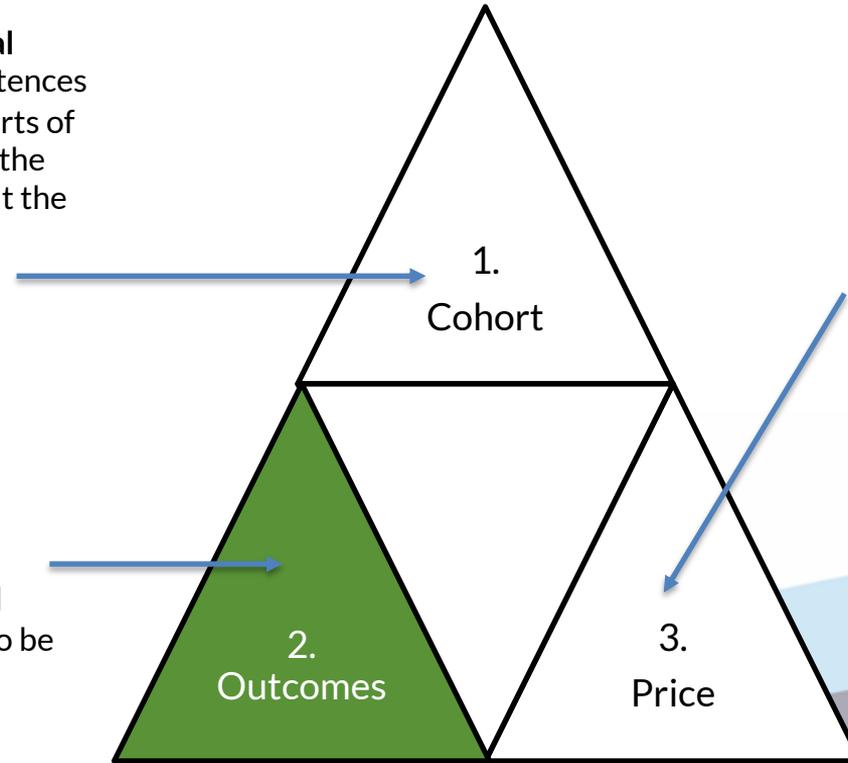
# What does 'good' look like?

## Example: HMP Peterborough SIB

1. Eligibility **criteria clear and impartial** (offenders in Peterborough prison sentences of less than 12 months over three cohorts of 1,000) and **referral was independent** (the service was offered to everyone who fit the eligibility criteria, participation was voluntary).

2. Policy **intent to reduce reoffending** and the **project paid on a cohort-based reduction in reconviction**, a valid proxy measure tracked over a suitably long period to be durable.

3. Includes **real-time comparator group in the payment mechanism**: commissioners can say with confidence that the **outcomes they paid for are attributable to the ONE service, and additional to business as usual**



# Resources for commissioners

Setting and measuring outcomes 2 chapters ▾

## Setting and measuring outcomes

A guide to identifying, defining and measuring outcomes for the purposes of outcome based commissioning

Download as PDF ↕



Chapter 1  
**General guidance**

- Introduction and overview
- Understanding outcomes, measures and metrics
- Different types of outcome
- Individual versus cohort measurement
- Developing an outcomes framework
- Identifying the right outcomes

### About the guide

This guide provides advice on the processes of identifying, defining and measuring outcomes when putting in place an outcomes-based contract or social impact bond. It is aimed mainly at local commissioners, but should also be useful to providers of services and interventions who are involved in the development of new outcomes-based contracts – especially for the sector in transition.



GOVERNMENT  
OUTCOMES  
LAB

# Advice Surgeries

How to guides:  
[golab.bsg.ox.ac.uk/guidance](http://golab.bsg.ox.ac.uk/guidance)

Book online:  
[golab.bsg.ox.ac.uk/guidance/advice-surgeries](http://golab.bsg.ox.ac.uk/guidance/advice-surgeries)

# Get in touch



@ukgolab



[golab.bsg.ox.ac.uk](http://golab.bsg.ox.ac.uk)



[golab@bsg.ox.ac.uk](mailto:golab@bsg.ox.ac.uk)