Using HEAT's tracking data in a counterfactual analysis

Part of a mixed methods approach to evaluating outreach activities
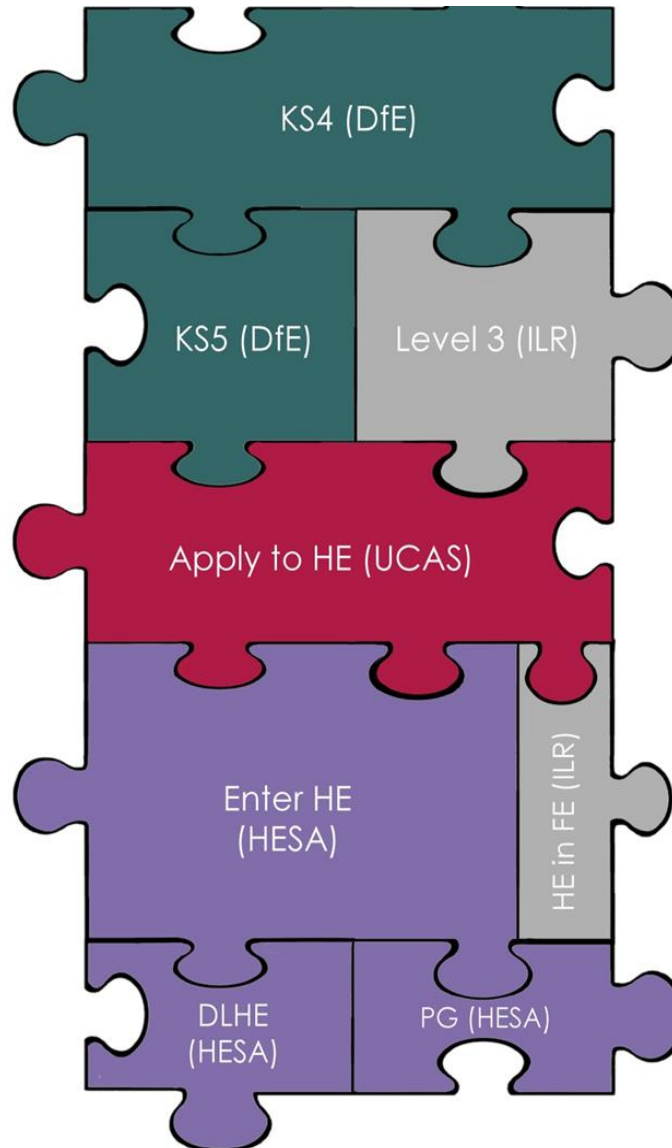
Anna Anthony, Senior Data Analyst

Presentation Content

- An overview of HEAT's tracking data
- Using tracking data in a counterfactual analysis
- An example of counterfactual analysis using the HEAT aggregate set
- Limitations
- How HEAT can help plan a robust research design

HEAT
Higher Education Access Tracker

The HEAT Track Jigsaw

Outreach participants added to HEAT by 84 member organisations

KS4 (DfE)

KS5 (DfE) | Level 3 (ILR)

Apply to HE (UCAS)

Enter HE (HESA) | HE in FE (ILR)

DLHE (HESA) | PG (HESA)

Report 1: Level 2
(underlying dataset not available to members)
- KS4 attainment of participants engaged pre 16

Report 2: Level 3
(underlying dataset not available to members)
- Transition from Level 2 to Level 3
- Post 16 trajectories
- KS5 attainment of participants engaged post 16
- HEI choice (contextualised with KS5 attainment)

Report 3: Level 4 +
(pseudonymised HESA dataset available to members)

- Application and conversion to HE
- Progression to HE (contextualised with KS4 attainment)
- Post 18 trajectories (HE, HE in FE, Apprenticeships)
- Achievement and success in HE
- Progression to PG
- Graduate outcomes

Tracking data in a counterfactual analysis

What is counterfactual analysis?

A comparison between *what happened* and *what would have happened* in the absence of the intervention.

"For strong Type 2 and for all Type 3 evaluations, you want to have a counterfactual or comparator to establish the impact of your intervention or activity above what might otherwise have occurred". (OfS, 2019)

Office for Students | OfS

Contact us   About   Search site

Home | Advice and guidance | For students | News, blog and events | Publications | Data and analysis | Get involved

Home > Publications > Standards of evidence and evaluating impact of outreach

Independent research

Standards of evidence and evaluating impact of outreach

It is important that higher education providers prioritise generating, sharing and learning from evidence about what is working to make higher education more equitable. Compelling evidence should underpin every provider's strategy to improve access and participation. Providers also need to evaluate their activities robustly to make sure they continue to be effective and contribute to national evidence.

These documents give advice on improving evaluation practice and strengthening evidence to support providers to make informed decisions about investment in access and participation activities.

Related publications

Degree apprenticeships: Motivations research

Evaluation of NMMLGP

LG pilot projects evaluation

Catalyst fund projects evaluation
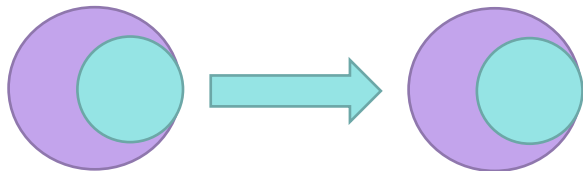
Date:
28 February 2019

Source: OfS website

Slide 4

Counterfactual Analysis & Tracking Data

Non-experimental designs

- Measures outcomes
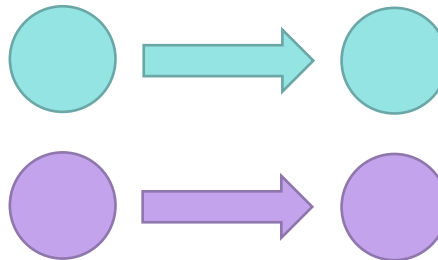- Comparator group does not consider selection bias

Quasi-experimental designs

- Identifies a comparator group
- Matching techniques e.g. PSM
- Success will depend on the quality of matched variables
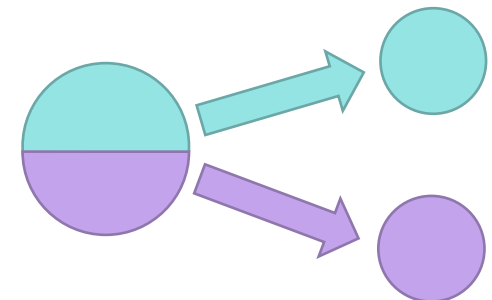- Feasible with retrospective tracking data

Experimental designs

- RCTs – championed by 'What Works' Centres
- Identifies a 'control' group through randomisation
- May not be feasible in an outreach context

Non-experimental design

*Quasi-experimental design*

Experimental design

Counterfactual Analysis from HEAT

An example of a quasi-experimental counterfactual analysis using HEAT's aggregate dataset

Research Question: How can we use tracking data from HEAT's aggregate set to evidence the impact of outreach on Key Stage 4 (GCSE) attainment?

223,725 secondary school students tracked through HEAT

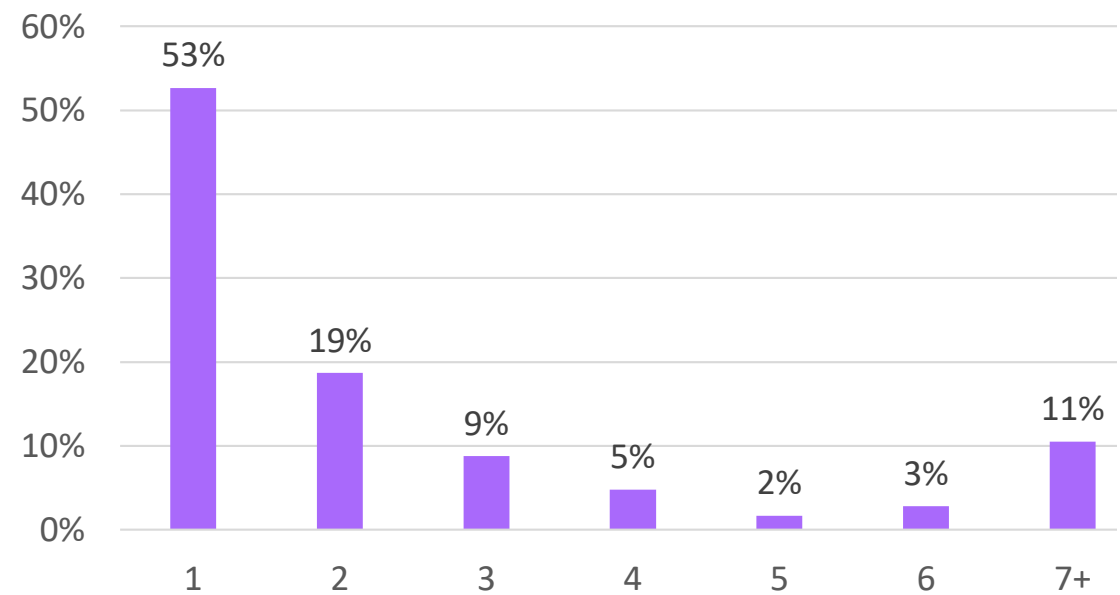Outreach → Motivate students to engage in class → Increased Attainment
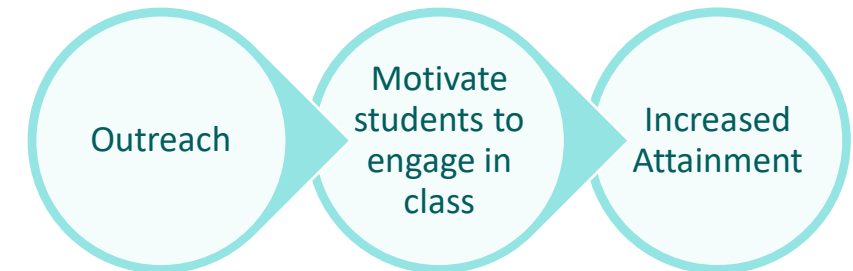
Three criteria:

✓ Engaged in outreach before sitting their GCSEs

✓ Were not 'on track' to achieve 5 good GCSEs including English & Maths

✓ Were disadvantaged according to our set of proxies
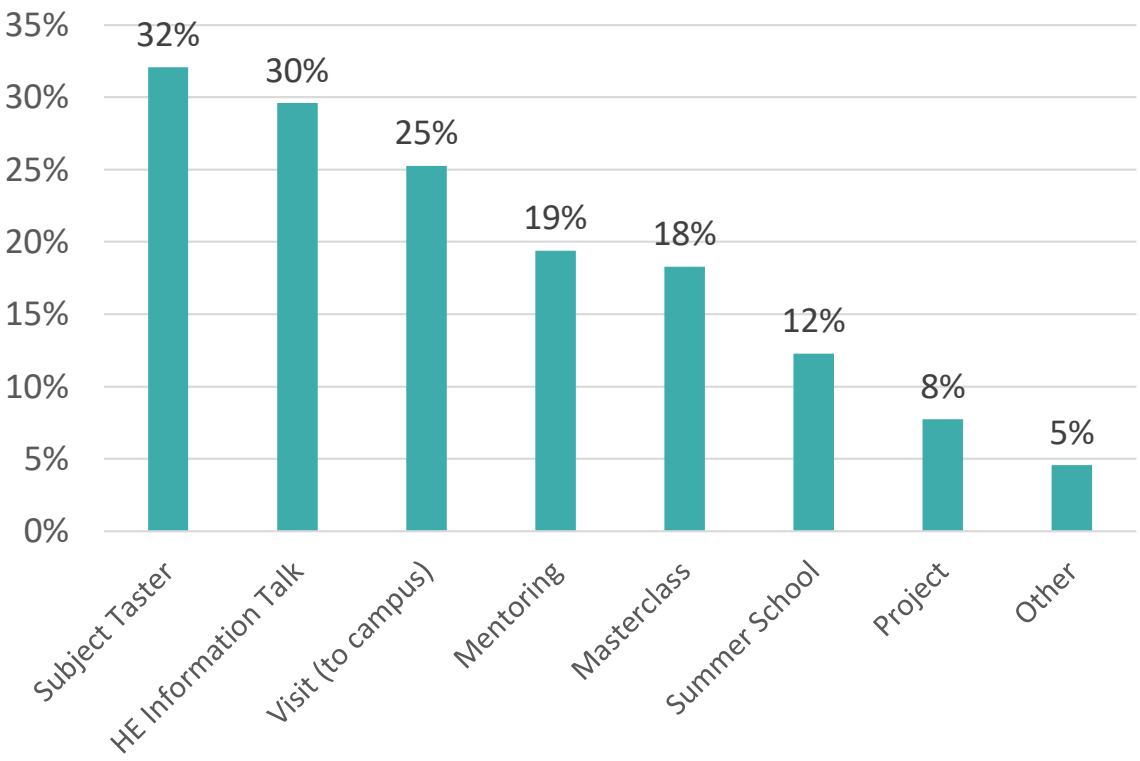
17% (n=17,305) participants met all three criteria

Counterfactual Analysis from HEAT

HEAT
Higher Education Access Tracker

Activity Participation of the Sample Population

Number of Activities

Number	Percentage
1	53%
2	19%
3	9%
4	5%
5	2%
6	3%
7+	11%

Type of Activities

Type	Percentage
Subject Taster	32%
HE Information Talk	30%
Visit (to campus)	25%
Mentoring	19%
Masterclass	18%
Summer School	12%
Project	8%
Other	5%

Counterfactual Analysis from HEAT

A matched comparison group design

Propensity Score Matching (Type 3)

Participant group
3+ Activities

Comparator group
Single HE Talk
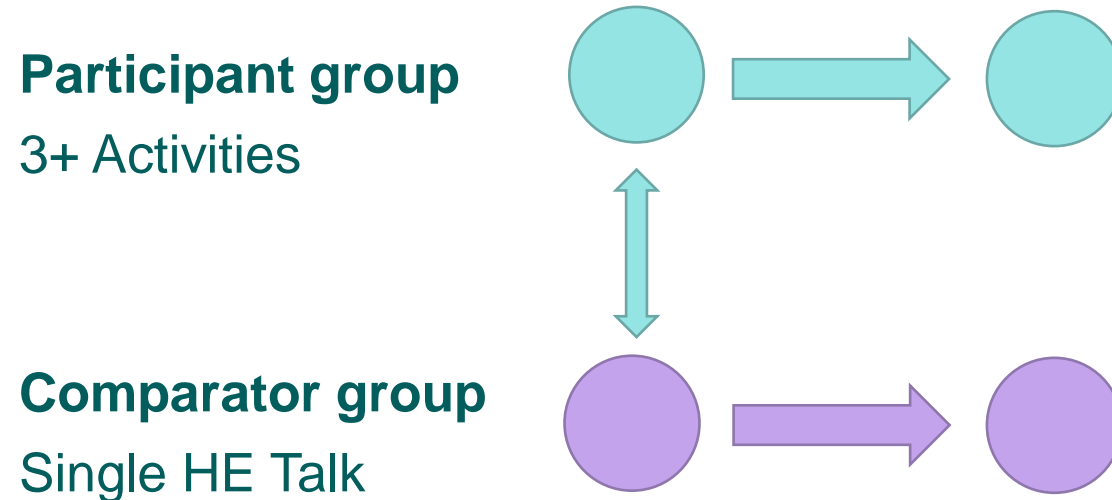
Factors used in matching are critical

- Gender
- Ethnicity
- IDACI Quintile
- KS2 attainment level
- School performance at KS4
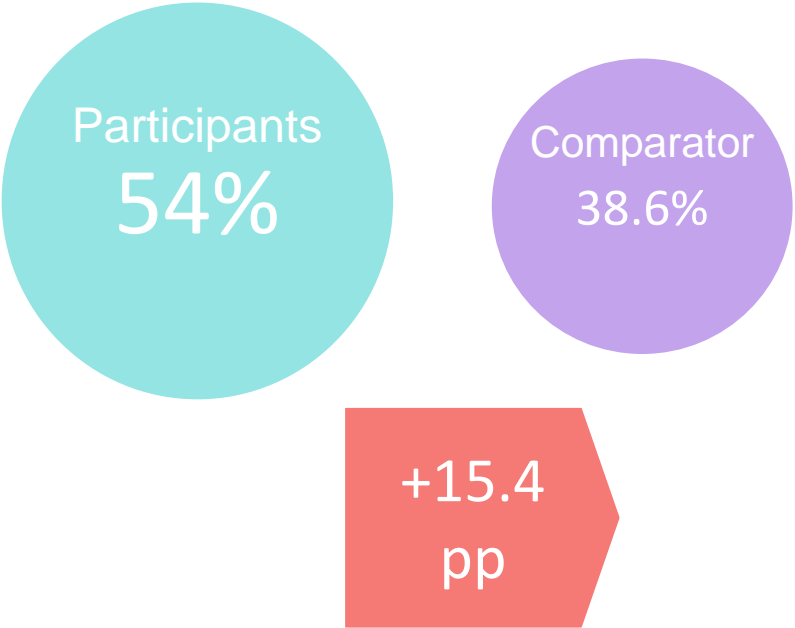- KS4 exam year

N = 1,634 or 817 pairs
High tolerance = Groups very similar in composition
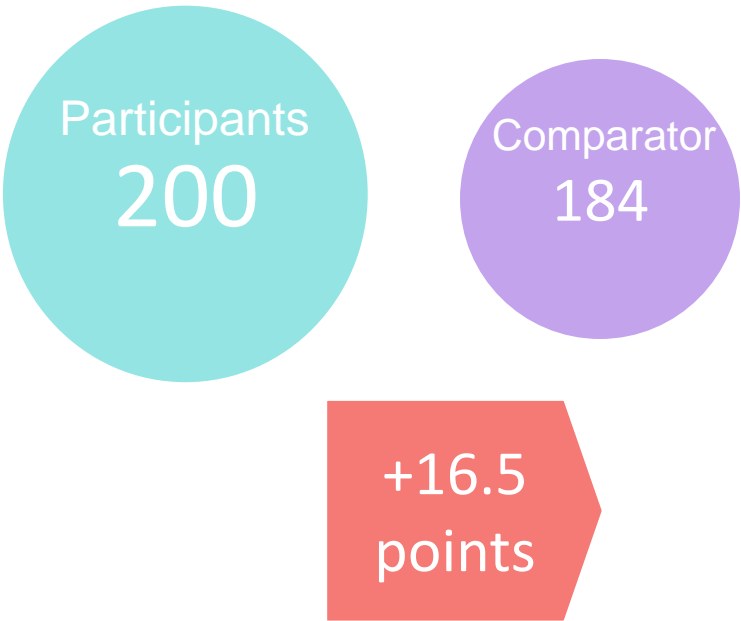
Counterfactual Analysis from HEAT

Match Variable	Pre Matching		Post Matching	
	Participant Group	**Comparison Group**	**Participant Group**	**Comparison Group**
Gender				
Female	58%	55%	57%	57%
Male	42%	45%	43%	43%
Ethnicity				
White	79%	85%	79%	79%
Asian / Asian British	10%	7%	10%	10%
Black / African / Caribbean / Black Briti	6%	4%	6%	6%
Mixed / Multiple ethnic groups	3%	3%	3%	3%
Other ethnic group	1%	1%	1%	1%
KS2 attainment				
High (>Level 4)	0%	0%	0%	0%
Medium (Level 4)	26%	45%	26%	26%
Low (<Level 4)	74%	55%	74%	74%
IDACI Quintile				
Quintile 1 (Most deprived)	41%	36%	40%	40%
Quintile 2	48%	55%	49%	49%
Quintile 3	7%	5%	7%	7%
Quintile 4	3%	3%	2%	2%
Quintile 5 (Least deprived)	1%	1%	1%	1%
School performance at KS4 Quintile				
Quintile 1 (Most deprived)	35%	33%	34%	34%
Quintile 2	40%	45%	41%	41%
Quintile 3	15%	12%	14%	14%
Quintile 4	5%	5%	5%	5%
Quintile 5 (Least deprived)	5%	5%	5%	5%

Counterfactual Analysis from HEAT

HEAT
Higher Education Access Tracker

The Results

Percentage who achieved five GCSEs at A*-C including English and Maths

Participants
54%

Comparator
38.6%

+15.4 pp

Average capped point score of the 'best eight' GCSEs

Participants
200

Comparator
184

**+16.5 points**

Counterfactual Analysis from HEAT

Limitations

- Unobserved variables
  - Motivation of students not included in matching
  - Selection bias remains even after matching?
- No knowledge of activity content
  - Trade off between sample size and quality
  - Don't know why/how outreach raises attainment
  - May have more value at local level

Mitigating the Limitations

- Guidance states that:
  - Activities should be set in a theoretically sound framework setting out causal mechanisms
  - Triangulate findings using other methods (mixed methods)

Subsequent Outcomes from HEAT
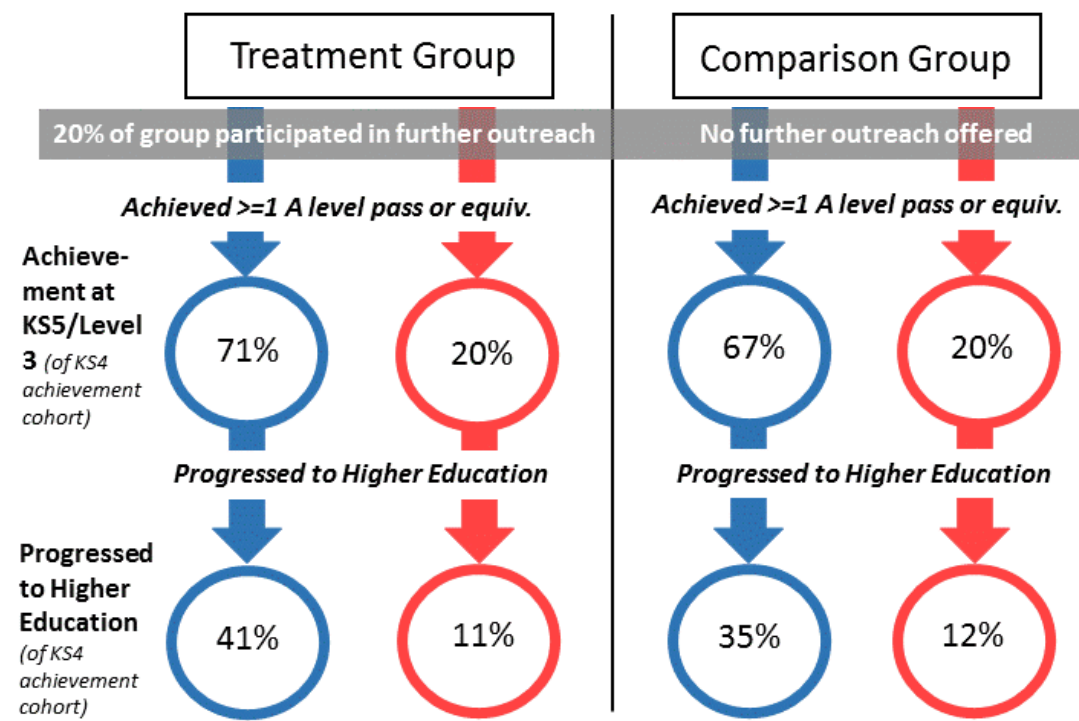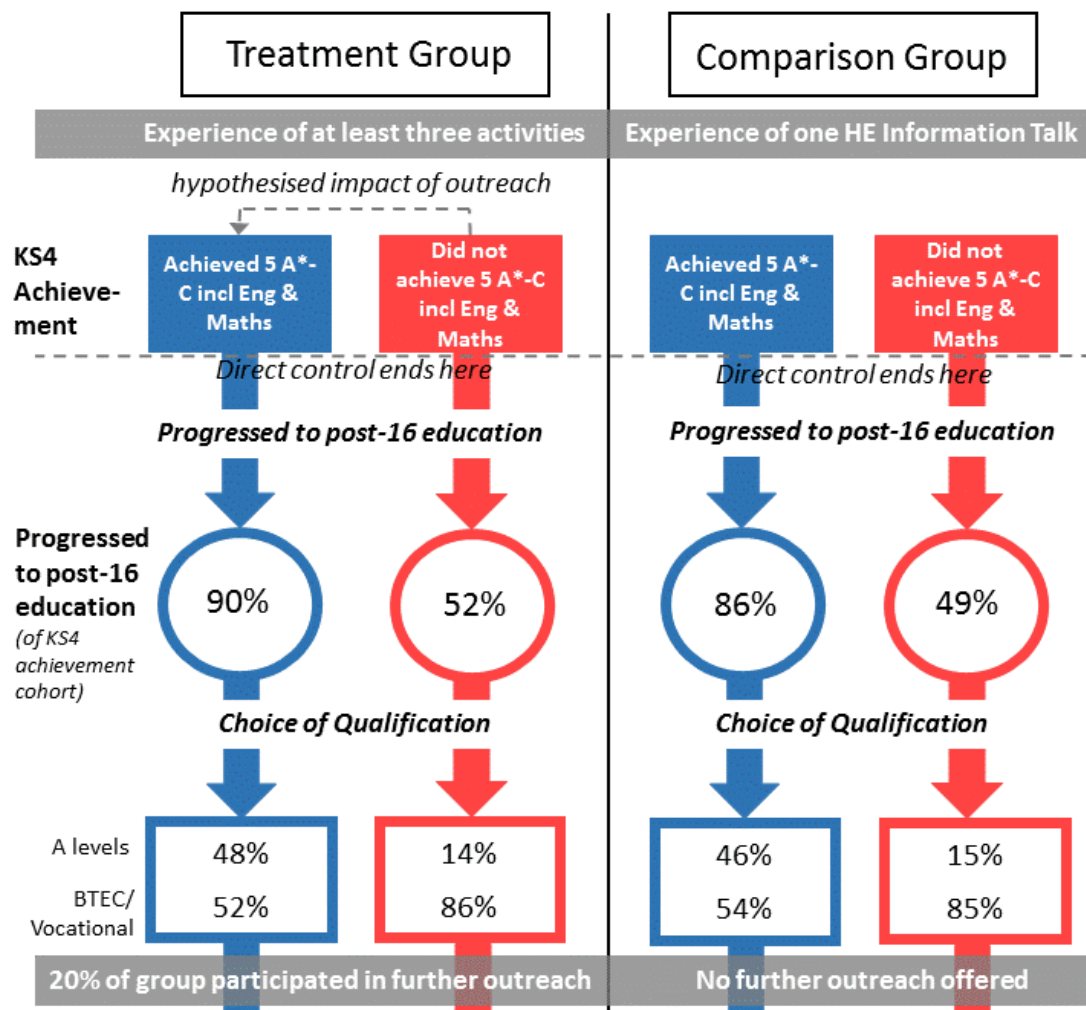
HEAT's Evaluation Planning Tool

HEAT
Higher Education Access Tracker

| General | Theory of Change | Target Students | Tar... | ...ies | Stakeholders | Audit |

Outcomes

Outcome 1: Increased Self-efficacy

Please indicate how you will measure this by ticking the boxes be...

SURVEY	
INTERVIEWS	
OBSERVATIONS	
FOCUS GROUPS	
ATTAINMENT SCORES	None ...
PROGRESSION TO HE	None ...
OTHER	None ...

Survey

- [ ] Data collected post intervention only
- [ ] Data collected pre intervention only
- [ ] Data collected pre and post intervention
- [ ] Using a comparison group
- [ ] Using an experimental or quasi-experimental research design

CLOSE

Slide 13

Questions?

- Has anyone tried PSM with their local tracking data?
- Are we missing any covariates?

References

HEAT website
https://www.heat.ac.uk

Requests for support (HEAT Helpdesk)
support@heat.ac.uk

Requests for training
comms@heat.ac.uk

Anna Anthony, Senior Data Analyst
Anna.anthony@heat.ac.uk