General Elimination Methodology Case Study

Impact Evaluation with Small Cohorts: Methodological Guidance (42–44)

Methodological Steps

Scriven, M. (2008). A Summative Evaluation of RCT Methodology: & An Alternative Approach to Causal Research. *Journal of Multi-Disciplinary Evaluation*, 5(9) 11–24. Available at: https://journals.sfu.ca/jmde/index.php/jmde_1/article/view/160
(Open Access)

White, H. and Phillips, D. (2012). Addressing Attribution of Cause and Effect in Small n Impact Evaluations: Towards an Integrated Framework, Working Paper 15, International Initiative for Impact Evaluation
Available at: https://www.3ieimpact.org/file/7911/download?token=unt2c-N2
(Open Access)

Case Study

Salazar, B., Mills, M. and Verissimo, D. (2019). Qualitative impact evaluation of a social marketing campaign for conservation. *Conservation Biology*, 33(3) 364–644. Available at https://doi.org/10.1111/cobi.13218
http://www.diogoverissimo.com/wp-content/uploads/2019/12/Salazar-Conservation-Biology-2019-Final-Conservation-Biology-cobi.13218.pdf
(Open Access)

Fabricated Widening Participation Example

Barkat, S. (2019). Evaluating the impact of the Academic Enrichment Programme on widening access to selective universities: Application of the Theory of Change framework. *British Educational Research Journal*, 45(6) 1160–1185.
Available at: https://doi.org/10.1002/berj.3556
(No Open Access version is currently available)

The fabricated example draws on a Theory of Change developed by Barkat (2019) to document an academic enrichment programme for Y12 students. All the details below, however, are fabricated and do not refer either to the intervention or its evaluation as described in the paper.

In the table below, the 'Case Study' column breaks down the case study evaluation into a series of methodological steps, as described in the Methodological Guidance. In the 'Fabricated WP Example' column, we apply the logic of these steps to a hypothetical evaluation of a fabricated widening participation intervention, to suggest how a General Elimination Theory approach to evaluation might unfold when applied to an intervention of this type. The nature of this 'Small *n*' approach means that there may be no single 'correct' way of applying this methodology. The example given should be considered illustrative rather than a definitive model.

Case Study	**Fabricated WP Example**
Salazar et al. 2019 Outline of paper: The article describes the application of a GEM to the evaluation of the long-term impacts of a social media campaign to increase the population of the Lora, an endangered parrot, on the Caribbean island of Bonaire. It outlines the different stages of the evaluation, from developing an initial theory of change and modelling potential causes for increases in the Lora population to gathering and testing additional causal theories with expert stakeholders. Key evaluation question: Did a social media campaign have a positive impact on the increase in the Lora parrot population? Page references below are to the published version of the article.	There are few published examples of GEM being applied to the evaluation of WP-focused interventions. The example below is a hypothetical model to suggest how this approach could be used in the evaluation of a WP intervention. The starting point for this discussion draws on a Theory of Change documented in Barkat 2019, but the discussion below is based on an entirely fabricated example case study. Outline of paper: The article outlines the development of a Theory of Change approach to evaluating the impact of an academic enrichment programme for disadvantaged young people in Y12. Key evaluation question: Does the academic enrichment programme improve the academic attainment of participants?
Step 1 - Establish a list of possible causes	
The researchers established an initial Theory of Change for possible causes of the increase in the Lora population. They then considered the necessary conditions required for each cause to lead to population change. The initial Theory of Change was constructed through a literature review of relevant campaign and environmental reports and grey literature concerning the Lora and the Lora population (636).	The evaluators established an initial list of possible causes for the lower attainment rates of disadvantaged students compared to their more advantaged peers. This list of causes was constructed through: i) interviews with key delivery staff, including programme leads ii) a review of programme documentation

This process led to the identification of 29 factors hypothesised to have an impact on the parrot population. These hypotheses were explored, expanded and tested – and rival explanations sought – through eight interviews with stakeholder groups, which included 33 participants, including 'veterinarians, biologists, tourism professionals, educators, government officials, media representatives, local residents, and conservation professionals (including forestry game wardens and conservation organization employees)'. Stakeholders were selected through both purposive and snowball sampling (636). Interviews gathered information on participant backgrounds and involvement with the Lora population, tested the TOC by asking for rival explanations for the increase in the Lora population, and sought responses on the social media campaign. The evaluation team also conducted a card-sorting exercise, in which participants were asked: i) whether each of the 29 causal factors in the TOC was understood to have affected the Lora population, not affected the Lora population, or whether they were unsure ii) to rank those cards which indicated a causal relationship on a scale of the strength of influence iii) to indicate any perceived causal relationship between the different causal factors (638–9). To be included in the Theory of Change, a factor had to be identified as causal by at least two-thirds of the interviewees. At the end of this process, stakeholders reached a broad agreement on 18 factors that had an impact on the bird population, of which 12 were viewed as having a positive impact. The remaining 11 factors were eliminated.	iii) a general literature review of the research exploring causes for reduced Level 3 attainment outcomes for disadvantaged students. This initial phase aimed to identify the desired short, intermediate and long-term intended impacts of the intervention and to understand how change was expected to occur. This process generated the following possible causes for increased attainment rates in the target participant group: a) The academic criteria for participation in the programme pre-selected those students most likely to succeed. b) Participation in the programme increased participants' feelings of self-esteem as a consequence of having been selected for the programme, which in turn bolstered their confidence and ambition. c) Bringing students together in a programme intensified peer support and reinforced each participant's confidence in their own ambitions.

Step 2 - List the modus operandi for each cause	
During the card-sorting exercise, participants were asked to explain how they considered each factor to have affected the Lora population. Having identified 12 potential causes for the population increase, and a description of how each was understood to have caused this change, a further round of research was conducted to assess whether the conditions required for each of these explanatory narratives to have an impact were present in the intervention context (639–640). The interview process also generated new sources of evidence (additional reports).	The evaluators then unpacked each of these three possible causes to explore assumptions about how they caused the intended outcomes. A further round of stakeholder interviews was conducted. The range of stakeholders was increased to include programme participants, their teachers and a selection of education experts. As part of this process, stakeholders were asked to explain how they understood each possible causal chain to have functioned in delivering increased attainment and HE application outcomes. They were also asked to indicate their perception of the strength of influence of each of these factors by ranking them on a 10-point scale. Causal Chain A: Interviewees suggested the academic requirements of the programme meant that participants were more likely to experience academic success than their less well-qualified peers. They also suggested, however, that academic potential was increased by the targeting of support provision for core academic skills. This was understood to help students focus on their own academic potential and increase their motivation, and was ranked as a relatively strong cause of improved outcomes. Stakeholders also rated academic skills sessions as having a positive impact on attainment outcomes by creating the space for students to focus and reflect on the importance of 'playing the game' when approaching assessments. Training students to understand and focus on assessment criteria was believed to support good exam technique and cause improved outcomes, and was ranked as a strong cause of improved outcomes. During the interviews, stakeholders were also asked about their understanding of the context and additional factors required for participants to develop their

	academic skills. This additional data included the suggestion that participants began the programme with certain pre-existing knowledge and core knowledge about academic expectations. Stakeholders also suggested that to effectively engage with, and benefit from, coaching in maximising assessment outcomes, participants needed to already have a threshold level of motivation and belief in their self-efficacy.
Step 3 - Assess each case against the evidence available	
Evidence from government reports and grey literature supported the effect of nine of these positive factors and refuted one, which was therefore eliminated. The lack of evidence to confirm or reject the influence of two further factors resulted in these also being eliminated (639–640). The revised Theory of Change incorporated the nine remaining causal factors, which were triangulated by interviews and supported by evidence. These were grouped into three causal chains: changes in environmental laws, social marketing campaigns and environmental education in schools. The TOC also indicated interactions between individual causal factors (Figure 3, 641) Example of the interaction of causal factors: Changes in environmental laws led to a pet parrot amnesty which *Interacted with* social media campaigns about the plight of the Lora. *In parallel* increased environmental education in primary schools focused on the plight of the Lara, again increasing awareness. All of these factors were assumed to reduce the demand for Loras as pets, and therefore reduce the poaching and trade of these endangered birds (641).	The stakeholder interviews also suggested areas for a more focused review of the literature on the relationship between peer support/peer-assisted learning, academic development and individual ambition development. Causal Chain B: Based on stakeholder interviews and literature reviews, insufficient evidence was found to support a clear causal relationship between feelings of mattering and academic self-confidence. Stakeholders suggested that this context was too complex to support a direct relationship. In some cases, participants were unsure why they had been selected for the programme – mitigating any sense of mattering to the hosting organisation. There was insufficient evidence in the literature of a strong relationship between feelings of mattering, belonging and academic confidence. This possible cause was eliminated. Causal Chain C: Stakeholders also rejected the idea that the time that students spent together would be sufficient to generate the levels of peer support and validation required to impact participants' academic confidence. The literature review found evidence of the positive impact of peer-support and peer-assisted learning in degree-level students, albeit in the context of more intensive programmes with established class cohorts, but no evidence of efficacy in the intervention target group. Stakeholders tended to agree that the peer-group effect was weak and that participants did not have sufficient opportunity to build peer support within the

The conclusion of the evaluation suggested that 'social marketing campaigns can change human behaviour by supporting environmental laws and regulations'. In addition, they 'influenced the Lora population by shifting social norms, changing behaviours towards Loras, and increasing compliance with environmental laws' (642).	structure of the programme. No evidence was found in the research literature to counter this. This possible cause was eliminated. Emergence of a Rival Hypothesis: Teaching stakeholders felt that the selection process encouraged participation among students who already had relatively high levels of knowledge, confidence and understanding of higher education, sufficient to apply for selective institutions and achieve high grades in Level 3 assessments. This was viewed as a rival hypothesis for increases in attainment and the role of the programme itself on increased attainment outcomes. The evaluation concluded that the programme had a positive impact on attainment outcomes. It demonstrated the close relationship between three core components of the programme: selecting young people with a baseline threshold of academic potential, creating an opportunity for participants to focus on the skills required for academic success, and developing strategies for identifying and focusing on attainment criteria. These factors interacted with the broader metacognitive focus of the skills development sessions.
Conclusion	
The report's authors suggest that the GEM approach enabled them to build an evidence-informed model capable of demonstrating the impact of the social media campaign and how it interacted with other possible causes. It also enabled them to test and eliminate other possible causes for which there was no evidence and no stakeholder consensus (642). Furthermore, they conclude that this evaluation approach i) eliminates similar-person bias – or the risk of seeking out respondents already likely to agree with hypotheses ii) eliminates courtesy bias – where respondents tell evaluators what they think evaluators want to hear iii) exposes gaps in knowledge or evidence of possible causal factors (642).	The hypothetical result of this fabricated evaluation case study served to focus the initially broad Theory of Change on existing evidence and stakeholder-informed causal models. Some initial hypotheses about impact were eliminated due to insufficient evidence. The result is a clear model of the factors and contexts in which this kind of programme can have a positive impact on Level 3 attainment outcomes.