Knowledge Exchange funding: A review of novel evaluation methodologies

Final report

November 2019

SQW

Contents

Executive Summary

Context

1. Research England commissioned SQW, supported by City-REDI, to undertake a study examining the potential theory-based approaches that Research England could implement as part of the next overall evaluation of the Higher Education Innovation Fund (HEIF). Such theory-based approaches have not previously been implemented systematically or at a national level in a UK knowledge exchange (KE) context and were therefore considered novel.

2. The aim was to consider potential approach/approaches that can complement quantitative evidence on the outcomes and impacts HEIF funding is generating (i.e. the 'what'), with systematic and robust evidence on the mechanisms by which HEIF-funded activities lead to these outcomes/impacts (i.e. the 'how').

3. Specifically, the study sought to identify (including via a formal review of academic literature) and test potential theory-based approaches that would enable an evaluation to:

   - **Provide better explanation of 'how' HEIF generates impact**: exposing the relationship between inputs, activities, output, and outcomes, and considering its relative contribution alongside other factors and activities.

   - **Provide more detail and granularity on HEIF impact, beyond average return on investment (ROI) figures**: focusing on understanding the complete picture of value created by different KE functions, which can help to inform strategy.

Contact:	Peter Farrar	Tel:	07776 431421	email:	pfarrar@sqw.co.uk

Approved by:	Joe Duggett Director	Date:	26/11/19

Findings

4. A review of academic literature and wider evidence found that the application of theory-based methodologies to the evaluation of KE would be novel. We found no clear examples from elsewhere of such an evaluation being conducted, and therefore identified no methodologies that have been conclusively demonstrated to work when applied to a programme like HEIF.

5. We did, however, identify a range of methodologies with the (exploratory) potential to be applied to KE. Each methodology identified had its strengths and weaknesses but based on the priorities set out above, and the need to manage level of inputs required by those involved in the programme, two methodologies were deemed most appropriate potentially to meet the requirements for the evaluation of HEIF: Contribution Analysis and Realist Evaluation.

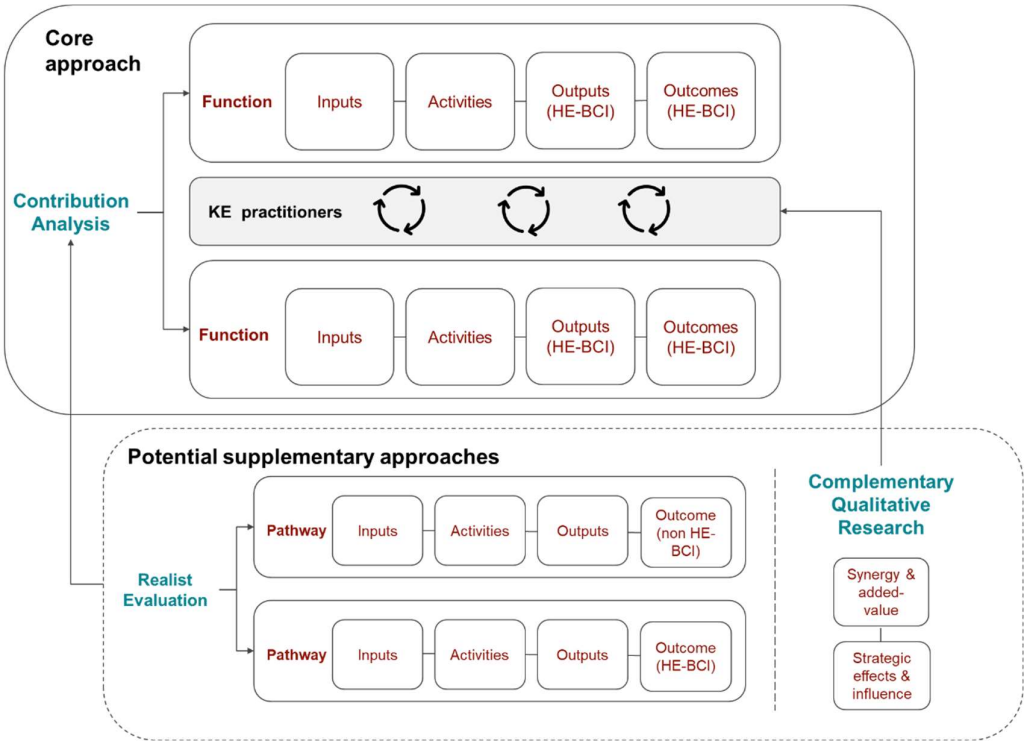
6. Both Contribution Analysis and Realist Evaluation rely on the development of a 'Theory of Change' (ToC) at the outset of the evaluation process. This ToC then serves as a hypothesis for the subsequent evaluation research to test. We focused on two KE functions as 'exemplars' and, with input from KE practitioners, developed a ToC for each to test the potential applicability and deliverability of the theory-based methods.

7. The process of developing the ToCs highlighted that the pathways to outcomes for HEIF are complex, although focusing specifically on those outputs and outcomes that are captured in the existing Higher Education Business and Community Interaction (HE-BCI) survey means that theory-based methods may be proportionate. The process of developing the ToCs also indicated that there will be important linkages and inter-dependencies between different KE functions, including the role of KE practitioners. Identifying and understanding these will be important for the subsequent evaluation.

Conceptual framework and implementation

8. The desk-review and findings from the development of the exemplar ToCs informed the development of a conceptual framework, that combines a 'core approach' focused on Contribution Analysis at the level of the functions (and focused specifically on those outputs and outcomes that are captured in the existing HE-BCI survey) with 'potential supplementary approaches' involving Realist Evaluation (considering individual outcomes in more detail) and complementary research on the 'connecting and translational' role of KE practitioners.

Figure 1: Conceptual framework



Source: SQW

9. Practical methods to implement the approaches have been set out. These would support the initial development of logic models and ToCs for each KE function, then the subsequent implementation of the 'six step' approach of Contribution Analysis. This would gather evidence for a 'contribution story', producing plausible evidence on the role that HEIF played in generating the KE outputs/outcomes in each function, and how important HEIF was relative to other factors in explaining how and why the outputs/outcomes were realised.

10. The approaches set out in the proposed framework are novel in a KE context. Their implementation therefore does pose a risk to Research England, and this is recognised explicitly and transparently by the study team. This should inform the next steps in taking forward the findings of the work, which may include piloting the approaches in advance of a full roll-out.

1. Introduction

Background and context

The Higher Education Innovation Fund (HEIF)

1.1 The Higher Education Innovation Fund (HEIF) is a well-established component of the UK's knowledge exchange (KE) landscape. Since 2001 it has distributed funding to higher education institutions (HEIs) in England to increase their capacity and capability for KE, the broad range of knowledge-based interactions between institutions and the wider world[1], which result in economic and social impact.

1.2 HEIF is currently administered by Research England and allocates funding to approximately 80% of eligible HEIs in England[2]. In 2019/20, HEIF allocated £213m to HEIs, calculated based on evidence from the Higher Education Business and Community Interaction (HE-BCI) survey for 2015-16 to 2017-18, and reflecting decisions made by Research England on the acceptability of institutions' five-year KE strategies. The Government recently announced that from 2020/21, the HEIF allocation will increase to £250m.

Evaluating HEIF

1.3 The last full-scale evaluation of HEIF was published in 2009[3]. This produced quantitative evidence on the value of HEIF, particularly in terms of its return on investment (ROI), which focused on the comparison of investment via HEIF (and other 'third stream' funding at the time) and income from KE activities (e.g. from collaborative and contract research). The 2009 evaluation was updated in 2015[4], with an evaluation focused on quantitative impacts. A separate qualitative evaluation of the fund also took place in 2015[5]. This produced a series of case studies that provided a narrative on the non-quantifiable value of the fund.

1.4 However, while these evaluations have provided good evidence on the ROI and benefits of HEIF support, they have generated high-level and average results. Crucially, they have not provided (and were not intended to provide) evidence of how different uses of HEIF drive different and specific impacts, and the relationships between activities and resulting outputs and outcomes across different categories of KE. This evidence gap means that although there is historic evidence on what HEIF has achieved in quantitative terms, and an established process to replicate this evidence in the future, '*how*' it has done this – which can both help to inform strategy and demonstrate further the value of the Fund – is uncertain.

1.5 This is in part due to challenges facing the evaluation of as broad and complex a programme of funding as HEIF. Key characteristics of HEIF that make understanding the pathways to outputs and outcomes complex and include:

---

[1] Including the exchange of ideas, evidence and expertise.
[2] UKRI, 2019. *The Higher Education Innovation Fund (HEIF)* (accessible here).
[3] PACEC & University of Cambridge Centre for Business Research, 2009. *Evaluation of the effectiveness and role of HEFCE/OSI third stream funding* (accessible here).
[4] Ulrichsen, 2015. *Assessing the economic impacts of the Higher Education Innovation Fund: A mixed-method quantitative assessment* (accessible here).
[5] PACEC, 2015. *Evaluating the non-monetised achievements of the Higher Education Innovation Fund* (accessible here).

- the diversity of HEIs funded, with the fund providing support to over 100 institutions in the latest (2019/20) funding round, which each has its own model and approach to the delivery of KE activity
- the flexible nature of HEIF, and consequentially the variety of KE activities it supports
- challenges delineating the relative impact of HEIF from the impact of HEIs' other funding sources; and
- identifying longer-term impacts generated by KE activity, and the likelihood that these impacts will have been driven by multiple inputs.

An introduction to theory-based evaluation

1.6 Theory-based evaluation is referred to by many names[6]. These different names reflect nuances of emphasis and practice, but a key feature is the use of a 'theory of change' (ToC) that specifies relationships between a programme's aims, activities and outcomes, plus the contextual factors influencing progress/success. Development and use of a ToC enables theory-based evaluations to both examine what has been achieved by a programme – and importantly for this study – to test assumptions about *how* programmes' activities and underlying mechanisms lead to expected/observed outcomes[7].

1.7 Theory-based methodologies generally take one of two approaches:

- **Define a programme's ToC and then gather evidence to test and substantiate whether that ToC played out in practice**. This might include exploring several alternative hypotheses outlining what might have occurred if assumptions made in the ToC (about contextual factors enabling activities to deliver anticipated outputs, for example) did/did not play out in practice. Evidence is gathered to assess each explanation, and establish causation beyond reasonable doubt by validating, invalidating and/or revising explanation explanations, documenting links between a programme's inputs, activities, outputs and outcomes/impacts.
- **Gather evidence that can be used to examine the steps taken between initial inputs** (e.g. funding provided) **and subsequent outcomes/impacts,** with a primary focus on assessing how a programme might be improved (as opposed to focusing on demonstrating a causal relationship or explaining how outcomes/impacts have been realised). This approach is normally implemented as part of a wider evaluation framework.

1.8 Theory-based methodologies are suited to a range of contexts, two of which are particularly relevant for HEIF and KE funding: contexts where activities are complicated[8] or pathways to

[6] Including programme-theory evaluation, theory-driven evaluation, theory-guided evaluation, theory of action, theory of change, programme logic, and logical frameworks.
[7] In contrast to quantitative evaluation methodologies, which attribute activities to outcomes by isolating relevant variables as exploring the correlation between them.
[8] e.g. where there are multiple components/partners involved in delivery or outcomes are influenced by a wide range of external factors

impact are unclear[9]; and contexts where the evaluation is required to understand how, why and under what conditions effects have taken place[10].

1.9 A range of research techniques are used in theory-based evaluation, and they are likely to involve mixed methods, including both quantitative and qualitative techniques. However, they differ from (and can complement) impact evaluation methodologies (which often involve the use of control groups and focus principally on quantitative analysis, including econometric techniques) by providing an overarching theory-based framework, and enabling different sources of evidence to be synthesised and triangulated.

1.10 A full glossary of the terminology and concepts discussed in this report is included in 6.11Annex A:.

This study

1.11 In this context, Research England commissioned SQW, supported by City-REDI, to undertake a study to consider the potential theory-based approaches that Research England could implement as part of the next overall evaluation of HEIF. Such theory-based approaches have not previously been implemented systematically or at a national level in a UK KE context, and were therefore considered novel.

1.12 The aim was to consider potential approach/approaches that can complement planned quantitative evidence on the outcomes and impacts HEIF funding is generating (i.e. the 'what'), with systematic and robust evidence on the mechanisms by which HEIF-funded activities lead to these outcomes/impacts (i.e. the 'how').

1.13 Specifically, the study sought to identify (including via a formal review of academic literature) and test potential theory-based approaches that would enable an evaluation to:

- **Provide better explanation of 'how' HEIF generates impact**: exposing the relationship between inputs, activities, output, and outcomes, and considering its relative contribution alongside other factors and activities
- **Provide more detail and granularity on HEIF impact, beyond average ROIs**: focusing on understanding the complete picture of value created by different KE functions, which can help to inform strategy.

1.14 Once potential theory-based methodologies were identified and tested, the study was tasked with developing a conceptual framework – based on exemplar logic models/theories of change – and providing recommendations of how this framework could be implemented in practice as part of the next overall evaluation of HEIF.

1.15 Four points are important to note in relation to the purpose and remit of the study:

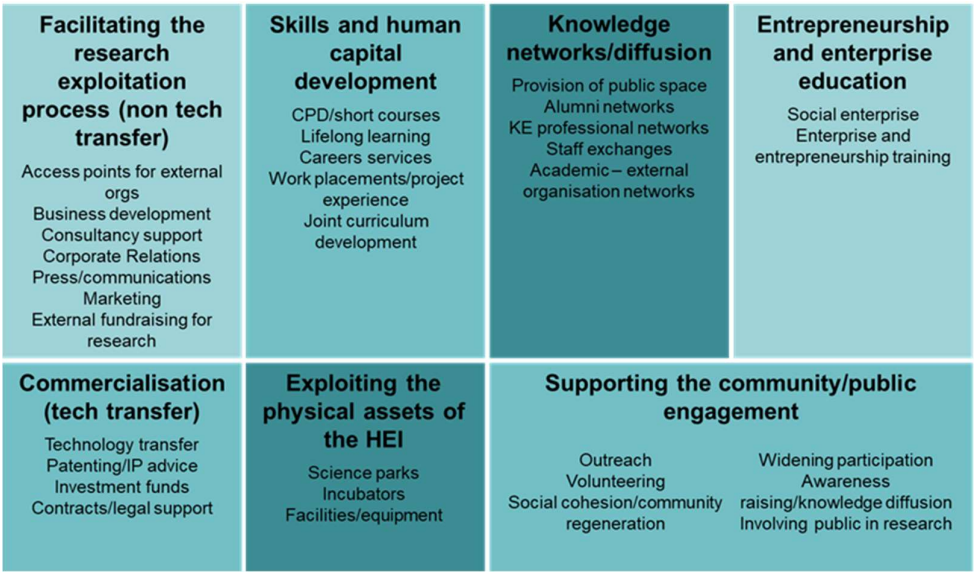
- The study was based around the existing depiction of KE 'functions' (summarised in Figure 1-1) that has been adopted by Research England and used, for example, in HEIF institutional strategy templates to record expenditure, and in discussions around the

[9] e.g. where there are emergent/uncertain outcomes or the intervention's cause and effect are not well defined
[10] Including where interventions have been applied in a variety of different contexts and had varying levels of success in each.

development of the KE framework. The study did not include seeking to develop an updated or alternative function typology, and the functions were identified as the key building block and framework around which potential theory-based methodologies for the evaluation of HEIF should be framed.

- Linked to the above point, the study was not tasked with undertaking an audit of KE activities, the development of detailed descriptions of every use of HEIF, or a comprehensive list of HEIF outputs and impacts. The process of developing exemplar logic models/ToCs to test the potential approaches did involve seeking to capture the relevant activities, outputs, and outcomes/impacts in order to test the evaluation approaches; however, this drew on existing evidence and descriptions, and feedback from practitioners. These logic models/ToCs are therefore illustrative only and would need to be revised and considered in more detail in the full implementation of the proposed methodologies (as discussed in more detail later in this report).
- The focus of the study was on testing and developing approaches for theory-based evaluation methodologies that can improve the evidence base available to Research England on how it/HEIF generates impacts only. Methodologies identified that focused principally on supporting continuous learning/improvement, or on development/capacity building within HEIs were not to be considered further. This focus on the 'accountability' purpose of evaluation was an important factor in the initial assessment and short-listing of potential methodologies, as discussed in Section 2.
- The study did not include considering whether theory-based evaluation methodologies could inform or provide further evidence (e.g. by filling gaps in evidence) for the quantitative methods used to assess the financial or other outcomes of the HEIF funding.

Figure 1-1: HEIF functions



Source: The state of the English university knowledge exchange landscape (RSM PACEC, 2017)

This report

1.16 This report presents the findings of our study and is structured as follows:

- Section 2: Identifying viable qualitative methodologies
- Section 3. Conceptual application of methodologies
- Section 4. Proposed evaluation framework
- Section 5: Implementation
- Section 6: Conclusions and recommendations

1.17 Appended after Section 6 are the following annexes:

- Annex A: a full glossary of the terminology and concepts discussed in this report.
- Annex B: detailed findings from the literature review.
- Annex C: Theory of Change diagrams setting out the process by which HEIF generates those outcomes documented in HE-BCI (for two functions).

2. Identifying potential novel methodologies

Study scope and priorities

2.1 The over-arching objective of this study was to explore the potential novel theory-based evaluation methodologies that could be applied as part of the next full-scale evaluation of HEIF, complementing quantitative evaluation approaches. The focus was on identifying methodologies that can better explain and test the causal relationships between KE activities and their outputs and outcomes – the 'how', alongside the 'what' that will be provided by quantitative approaches.

2.2 During an initial stage of the study, the research team met with Research England to discuss and refine the scope and priorities, including via interviews with those involved in the administration and/or evaluation of HEIF, as well as external researchers who led previous quantitative evaluation research on HEIF and KE on behalf of Research England.

2.3 The key messages and implications of this scoping phase were as follows:

- **Breadth and complexity of HEIF**: Within functions, HEIF can be used to support a broad range of activities. The balance and mix of these activities and their subsequent pathway to outcomes/impact will vary between universities, even in functions where the processes that lead to outcomes/impact are well recognised. Further, the way in which HEIF is allocated by universities will vary based on the range of other funding sources they have access to; for example, larger research-oriented universities may have more freedom to use HEIF funding to test experimental research concepts or approaches to external engagement. Internal and external contexts will also influence KE activity and outcomes. The study should consider the potential methodologies that can seek to consider and address this complexity as a priority for the work.

- **Maximising the existing evidence base**: Notwithstanding the level of complexity noted above, through the HE-BCI survey there is a strong evidence base available on the delivery of KE outputs and outcomes by institutions supported by HEIF. These are the key outputs and outcomes that the theory-based evaluation should focus on in terms of explaining the relationships between HEIF supported inputs/activities.

- **Focus on a number of KE functions**: Given the complexity within and across KE functions, it was agreed that the study should focus on developing a conceptual framework drawing on exemplar logic models/ToCs based on a number of prioritised functions, rather than all seven. It was recognised that the methodologies may also be able to explain causal links to outcomes/impact in other functions, and the review and framework needed to recognise this potential read across. However, developing exemplar logic models/ToCs to inform the proposed approach was not within the scope of this study. Drawing on the imperative to maximise and focus on the existing evidence base around outputs and outcomes, a key criterion for selection was for functions that contained activities where the links to HE-BCI outputs/outcomes were best established (e.g. CPD activity leading to the generation of KE income). The functions should also represent substantive uses of HEIF. Based on this, it was agreed that the study would focus on two functions:

- facilitating the research exploitation process (non-technology transfer)
- skills and human capital development (including enterprise education[11]).

- **Identifying applicable methodologies**: It was recognised that the study may consider that different theory-based methodologies are better suited to evaluating different functions, or to evaluating different levels of activity within a function (e.g. evaluating the impact of a function as a whole, versus examining linear pathways to individual outcomes). The outcome of this study might therefore be to identify a range of methodologies with different applications, rather than one single preferred methodology. As part of this, the study should consider the practical implications of the methodologies (e.g. value for money, proportionality), and consider the potential trade-offs between conceptual rigour/granularity of analysis and deliverability.

Evidence review process and findings

2.4 The identification of potential novel methodologies involved a literature review of international academic and grey literature, to examine the previous application of theory-based methodologies to evaluations and reviews of KE investment and activities. We also completed a parallel scoping review examining methodologies with the *potential* to be applied in a KE context, even if they did not yet appear in the literature,

Literature review

Search protocol and sifting

2.5 A protocol for the literature search was developed, informed by this study's original brief, scoping interviews with stakeholders and the research team's own knowledge of evaluation and KE activities. The protocol was developed detailing the key research questions of the review and the inclusion/exclusion parameters for documents identified by the search. This included publication dates, region/country of focus, and language of publication.

2.6 The protocol also detailed a set of search terms to be used, covering:

- **Different methodologies and methods** of likely interest for this study
- **KE activities and settings** these methodologies might have been used to evaluate
- **Impacts and benefits** likely to be associated with KE activities or funding programmes.

2.7 This protocol was then used by the City-REDI to search seven academic bibliographic databases[12]. This search returned an initial list of 7,065 documents that were potentially within scope.

2.8 These documents were first sifted according to their titles. Each title was reviewed and documents clearly out of scope for this study were excluded. Remaining documents were

[11] Enterprise education is an element of the 'entrepreneurship and enterprise education' function, but for the purposes of this study's modelling has been included as part of skills and human capital development.
[12] Econlit, Business Source Premier, Web of Science, Social Science Citation Index, Arts and Humanities Citation Index, Conference Proceeding Citation Index, Book Citation Index.

further sifted through a review of the abstracts. Similar to the title sift, the purpose of the abstract sift was to identify the documents that contained details of theory-based qualitative methodologies being used to evaluate KE. The abstract and title sift produced a final shortlist of 13 documents for detailed review.

2.9 The process of searching for and sifting evidence is detailed fully in Annex B.

Detailed review approach

2.10 A full text review was then undertaken of the 13 shortlisted documents. A structured review template was used to extract information from each document. This template contained a range of fields, each detailing the specific piece of information required (or the absence of that information noted clearly). Factors explored by the review and captured in this template included:

- **An overview of the approach**, detailing the process by which the methodology recommends a) collecting evidence and b) analysing that evidence.
- **An assessment of the methodology's novelty,** including its similarity/difference to other approaches and any methodologies from which it has been developed.
- **Examples of the methodology's previous application**, and its success in demonstrating pathways to impact when used to evaluate KE activities/funding.
- **The methodology's resource-/data-intensiveness** (i.e. the range of stakeholders that would need to be engaged in the research as researchers or as research participants, and the level of involvement required).
- **An assessment of the methodology's rigour**, including its robustness, reliability and replicability.

2.11 An overall assessment of the methodology's strength and weaknesses was also made, taking into account all of the factors set out in this list.

Documents review

2.12 The 13 shortlisted documents reviewed in detail are set out in Table 2-1. Full details of each document and the populated review template are provided in Annex B.

Table 2-1: Documents reviewed for literature review

Text	Description of content/focus
Abboud, 2016	An example of the use of **Outcome Harvesting** in community development, an area with similarly varying and non-linear pathways to impact as the commercialisation of university technologies.
Archibald, 2018; Banks, 2017; Salter, 2014	Three texts on **Realist Evaluation**, including: a published protocol for an evaluation of relationship between transdisciplinary research collaborations and knowledge translation; an evaluation of the development of diagnostic pathways; and a systematic review around the challenges in using Realist Evaluation in evaluating knowledge translation.
Baumgartner, 2017; Beach, 2018; Tho, 2013	Three texts on **Qualitative Comparative Analysis** (QCA), examining: how its application can be made more rigorous and transparent; how it can be combined

Text	Description of content/focus
	with process tracing for counterfactual analysis; and its use to test a 'programme theory' of how business school students can support knowledge transfer to SMEs.
Befani, 2017	A text on the use of **Process Tracing** building on **Contribution Analysis** to examine the impact of universal health campaign to influence policy priorities (which has parallels with KE interventions where universities support policy innovation).
Coryn, 2009	An adaption of **Success Case Method** that includes longitudinal follow up to examine casual associations where 'more scientifically rigorous' approaches such as RCTs are unethical or impractical.
Dart, 2003	A text describing the application of **Most Significant Change** to an agricultural knowledge exchange programme, where an HEI provided support in improving productivity of the dairy sector.
Gates, 2018	A paper discussing **Critical Systems Heuristics**; this is not an evaluation methodology but an approach to best define and demonstrate the value of complex interventions.
Kittel, 2013	A paper examining the methodological debates around the use of **Process Tracing**, now accepted to have good validity and considered by some second only to experimental methodologies in uncovering causal activity-outcome relationships.
Saunders, 2015	A text on evaluative research, combining three different evaluation methodologies (including theory-based evaluation) to provide a report that would meet the need of key stakeholders on the Quality Assessment Framework in Scotland.

Source: City-REDI literature review

2.13 The methodologies referenced in the above table (the bold text) are described in more detail later in this chapter (Table 2-2).

Summary of findings

2.14 The aim of the literature review was to identify existing evidence on the use of rigorous theory-based evaluation methodologies that could be applied potentially to HEIF to better explain and evidence the mechanisms by which HEIF-funded activities lead to outcomes/impacts.

2.15 Our overall finding was that the **formal literature contained very few reported instances of the application of novel theory-based methodologies to the evaluation of KE**. The review did not identify any published evaluations of closely comparable schemes or interventions to HEIF applying theory-based methodologies, indicating that they would be novel in this context.

2.16 The literature review therefore suggests that there will be a need to **translate methodologies that have gained acceptance in other fields of investigation and adapting them to the evaluation of KE**.

2.17 Realist Evaluation is an example of one such methodology: it has gained acceptance in the evaluation of health services in particular and, being 'methods neutral'[13], is beginning to be used to evaluate knowledge translation interventions in the health sector. The key insights in relation to Realist Evaluation from the literature reviewed include that it has the advantage of

[13] It does not prescribe a certain set of research methods that must be used to gather the evidence, meaning the researcher is free to choose the methods and data collection processes that best suit their purposes and sector.

being scalable (i.e. it can be trialled using a limited number of cases, and subsequently rolled out more widely if the results are found to provide valuable evidence), and that evaluations conducted to a good standard using this approach can be synthesised as part of a realist systematic review. The approach has also gained acceptance across the evaluation community (and in policy circles) as a means of understanding in what contexts specific interventions are more or less likely to result intended outcomes. However, the review also highlighted issues related to its proportionality and resource-intensive nature.

2.18 The review also identified the application and strengths/weaknesses of techniques that are potentially promising in relation to HEIF, such as QCA (which analyses different combinations of elements of an intervention, contextual variables and outcomes to determine which combinations of intervention elements and contextual variables lead to outcomes) and Outcome Harvesting (which can be used where there are multiple pathways and outcomes can be unknown and difficult to articulate and do not occur in linear fashion, which is of relevance to HEIF).

2.19 However, the review highlighted that these methodologies are likely to be very resource intensive – including requiring significant input from practitioners – which may have implications for proportionality and deliverability in this context. The level at which the evaluation is focused (i.e. at a function level, or specific outcomes) will influence this issue, with some methodologies more likely to be applicable when considering specific outcomes.

2.20 Other techniques such as Contribution Analysis and Process Tracing were also evident in the literature, which can be implemented as retrospective evaluation (i.e. after the event, rather than in programme development or in real-time during delivery), which can employ mixed-methods and draw on monitoring and secondary data alongside qualitative inputs. These methodologies can validate and test claims around causation and could support accountability, whilst supporting continuous learning and improvement in delivery. Other methodologies such as Most Significant Change are probably best suited to supporting learning and continuous improvement of programme delivery, rather than for accountability purposes.

Scoping review

2.21 Alongside the literature review, a parallel scoping review was undertaken considering the *potential* theory-based approaches that could potentially be considered, even where these may not yet have been applied in practice in a KE context. The focus of this review was on the analytical evaluation approaches themselves and their theoretical grounding/rationale, rather than specific research methods (e.g. case studies, interviews, surveys) or examples of methodologies' practical application.

2.22 Drawing on a range of sources including evaluation guidance materials (e.g. the UK Government Magenta Book and European Commission Evaluation Sourcebook), online toolkits and guides published by authoritative sources (e.g. Better Evaluation, Bond), previous studies undertaken by the study team, and other grey literature (e.g. other reports, online presentations), seven methodologies were identified that were considered to be of most relevance to the requirements of the study: Contribution Analysis; Process Tracing; Qualitative Comparative Analysis; Realist Evaluation; General Elimination Methodology; Most Significant Change; and Success Case Method.

2.23 For each methodology, data was collected on the approach and the key steps/stages involved to inform as assessment of its appropriateness for the evaluation of HEIF and the strengths and weaknesses of the approach in this context, considering issues including potential bias, robustness, the ability to extrapolate/generalise the findings outside cases/sample used for the analysis, proportionality, evidence requirements, and timing (with some methodologies more geared towards longitudinal research). This was used to inform the assessment of the methodology alongside the findings from the literature review, as summarised below.

Novel methodologies considered

2.24 The findings from the two reviews were collated and synthesised, suggesting that eight theory-based methodologies should be considered in more detail, and with the potential to be applied to a qualitative evaluation of HEIF. The eight methodologies, and their key characteristics (drawn from the findings of the literature and scoping review) are summarised in Table 2-2.

Table 2-2: Methodologies shortlisted following reviews

Methodology	Overview and key elements	Considerations
Contribution Analysis	A methodology for examining evidence against an intervention's ToC to develop a plausible 'contribution story', and relative contribution to other factors. Contribution Analysis tests and refines theoretical links between different elements of a ToC and assumptions about how they lead to outcomes. In doing so, it builds up evidence to demonstrate the contribution made by an intervention to the realisation of outcomes, whilst also establishing the relative importance of wider factors (e.g. market opportunities, business strategy, regulation, other interventions). This produces a 'contribution story' about the influence that the intervention itself (instead of other factors) has had on the realisation of observed outcomes. If an evaluator can validate a ToC with evidence, and account for the relative importance of major external influencing factors, it is thought to be reasonable to conclude that the intervention has made a difference to realisation of the outcome(s).	• Applicable in complex contexts where there are multiple components, partners, and where sole attribution is difficult • Factors external factors/conditions into analysis • Can be used in absence of a counterfactual to explain outcomes and the relative role of the intervention in the causal chain • Method-neutral and assumes mixed-methods, so highly adaptable to intervention context • Produces 'plausible' explanations rather than definitive evidence • Assumes a comprehensive and well-articulated theory of change • Few 'ground rules' to follow in application and approach to causal claims
General Elimination Methodology	A methodology that focuses on systematically eliminating competing causal explanations, to arrive at a causal explanation established with a high degree of certainty. General Elimination Methodology centres around systematically identifying and ruling out competing explanations of how observed outcomes have been realised. It does so by examining the facts of a case, establishing theories about different 'modus operandi' (MO) – single causes, or sequences of events – that might need to be present in order for outcomes to be realised, and sequentially eliminating theories for which data shows outcomes being realised without the presence of the hypothesised MO. In doing so, it provides a framework for evaluation which can establish causal claims beyond reasonable doubt.	• Examines multiple competing hypotheses • Can be used to arrive at causal claims of outcomes with a high degree of certainty • Suited to tracing 'backwards' from an outcome to its causes • Relies on researcher identifying all plausible hypotheses and variables of observed outcomes • Does not draw on a ToC • Limited evidence in the literature on usage and applicability
Most Significant Change	A methodology that involves selecting and analysing a representative sample of cases to examine 'significant changes/stories' produced by a programme/intervention. Most Significant Change is a form of participatory monitoring and evaluation that involves collecting, selecting and analysing a representative sample of personal accounts or stories of 'significant changes' that have occurred, relayed by programme stakeholders/beneficiaries. Its aim is to uncover the significant	• Reduces researcher bias by utilising participant/stakeholder insights • Can be used on an ongoing basis for continuous learning • Poor generalisability • Relies on stakeholder perceptions to identify and establish causality

Methodology	Overview and key elements	Considerations
	changes that have occurred for beneficiaries since the inception of the programme. MSC is intended as a tool for project management and continual improvement, applied throughout the lifecycle of a programme.	• Relies on stakeholders to identify impacts
Outcome Harvesting	A methodology for collecting evidence of change and then working backwards to assess whether/how a programme contributed to that change. Outcome Harvesting consists of six iterative steps that help uncover the changes in behaviours and actions that lead to outcomes. It actively engages a variety of important stakeholders in the process: *the change agent* (the individual or organization that influences an outcome), *the social actor* (the individual, group, community etc. that changes as a result of the social actor's activities), *the harvest user* (the individual(s) who will use the information gleaned from the OH for a variety of reasons) and *the harvester* (the person responsible for the OH process). Through engagement with these stakeholders, OH formulates, verifies, analyses and interprets 'outcomes' in contexts where relations of cause and effect are not fully understood.	• Well-suited to complex programmes lacking clear theory of change • Designed for outcomes influenced by numerous different factors • Starts with outcome and works backwards • Relies on outcomes being identified by research participants • Requires significant stakeholder involvement • Limited robustness, reliability and replicability
Process Tracing/ Contribution Tracing	Two similar methodologies for testing multiple hypotheses on the causal relationship between different inputs/activities and outcomes. Process tracing is a methodology for testing multiple hypotheses on the causal relationship between different independent variables (inputs/activities) and a dependent variable (output/outcome). It does so via arranging these variables into a chronological sequence, then assessing the extent to which qualitative evidence can be used to prove a hypothesis about a causal relationship between these variables. Contribution tracing undertakes these same initial steps, then develops them further by applying statistical techniques to analysis of the relationship between these variables to produce confidence scores (stating the statistical probability of a relationship between different inputs/activities and outputs/outcomes).	• Considers and tests multiple competing hypotheses • Suited to tracing 'backwards' from an outcome to its causes • Suited to examining in detail how individual outcomes realised • Cannot definitively demonstrate causality • Complex interventions make this approach increasingly resource-intensive • Requires very well-defined outcomes and anticipated pathways to impact and hypotheses
Qualitative Comparative Analysis	A case-based, comparative methodology for identifying the factors/conditions that lead to a given outcome, in a given context – applying complex modelling techniques. To deliver QCA researchers first identify the outcomes they wish to test and the potential influencing factors whose impact on realisation of outcomes they wish to	• Can establish causation • Identifies factors explaining outcomes • Considers impact of context/wider factors • Can be generalised where applied robustly

Methodology	Overview and key elements	Considerations
	test. They will then identify a set of cases (e.g. institutions, locations, businesses) that will be included in the analysis, including cases realising a mixture of positive and negative outcomes. Data on both outcomes and influencing factors is then collected and analysed, examining the combinations of influencing factors that are most heavily associated with realisation or non-realisation of outcomes. Findings of this analysis are then interpreted to develop or refine original hypotheses on the factors most important in enabling outcomes to be realised.	• Cannot be used with very small (<10) or large (50+) numbers of cases • Struggles in highly complex contexts • Relies on complete data-sets, with risks to analysis • Does not explain 'how or why' impacts realised, rather the conditions/factors that are associated with the outcome/impact
Realist Evaluation	A methodology that seeks to identify the mechanisms that enable interventions to achieve their results, incl. variation across different contexts – 'what works, for who, and how'. Realist Evaluation is based on a 'context-mechanism-outcome' (CMO) approach. This first identifies the outcomes, then works backwards to identify the mechanisms that enabled those outcomes to be realised, and then the contexts in which the mechanisms did or didn't enable this outcome to be realised. The findings of the CMO analysis are then used to determine which CMO configuration(s) offer the most robust and plausible explanations for the overall outcomes observed.	• Grounded in theory, and well suited to complex, multi-faceted interventions • Works backwards from outcome to explain its cause • Method-neutral, so highly adaptable to intervention context • Cannot establish a true counterfactual • Potentially very resource-intensive (at a KE function level)
Success Case Method	A methodology for taking the most and least successful cases within an intervention, to understand the factors that enhance or impede the intervention's realisation of impacts. Researchers create an 'impact model' that defines what 'success' should look like, implement a survey to search for the 'best' and 'worst' cases, and then interview and document the best cases as case studies (including qualitative learning from interviews on best practice and barriers to realising impacts).	• Can provide evidence of causality • Considers and is sensitive to impact of context/wider factors • Applicable where 'success' criteria are well-defined • Relies on impacts already being evidenced at 'case' level • Focuses on outlier cases only, does not explain wider pathways/outcomes

Source: SQW and City-REDI review of literature

Assessment of novel methodologies

Assessment criteria

2.25 A workshop attended by all study steering group members[14] was subsequently held to discuss these findings and their implications for the application of each methodology to a qualitative evaluation of HEIF. Each shortlisted methodology was assessed against a set of core criteria, including whether each was:

- **Capable of testing and demonstrating/explaining how outcomes/impacts are realised**, and relationships between inputs/activities and outcomes, including the role of an intervention in explaining these outcomes
- **Focused on the accountability purpose of evaluation**: as opposed to an emphasis on supporting programme development, learning and capacity development (which includes highly participatory approaches)
- **Capable (as far as practical) of accounting for the complex landscape and range of factors influencing outcomes**; central to the proposed role of the theory-based methodologies to complement the quantitative approach is that they are capable of explaining 'how' HEIF/supporting funding leads to impact in a complex environment
- **Deliverability, with implementation requiring modest resources and inputs from HEIF-funded institutions**; they should not involve *substantive* new data collection, or research activity (although they will necessarily rely on input from institutions and others e.g. in qualitative research via interviews, case studies etc.).

2.26 Preferred methodologies would be capable of producing valid and reliable evidence of the relationships between HEIF-funded activities and outcomes/impacts that the existing evidence indicates may be attributed to HEIF; they would also be able to produce this evidence without being overly resource-intensive, particularly without placing a significant burden of research participation on stakeholders in the KE landscape[15].

Findings

2.27 Our assessment of the eight identified methodologies against each of the above criteria is set out in Table 2-3.

[14] From Research England, SQW, City-REDI and PraxisAuril.
[15] Including staff and students at higher education institutions, business representatives and others delivering and/or benefiting from HEIF-funded KE activity.

Table 2-3: Assessment of theory-based methodologies

	Conceptual			Practical
	Capable of testing and explaining 'how' outcomes are realised	**Focused on 'accountability' purpose of evaluation**	**Applicable to complex programmes and contexts[16]**	**Deliverability, with modest resource requirements**
Contribution Analysis	✓	✓	✓	✓
General Elimination Methodology		✓	✓	✓
Most Significant Change	✓		(✓)	
Outcome Harvesting			✓	
Process Tracing / Contribution Tracing	✓	✓		✓
Qualitative Comparative Analysis		✓	(✓)	
Realist Evaluation	✓	✓	✓	
Success Case Method	✓			✓

Source: SQW and City-REDI analysis of literature

2.28 Two methodological approaches met all three of the core conceptual criteria set out above: Contribution Analysis, and Realist Evaluation. Both rely on well-defined logic models/ToCs, which would therefore be a particular focus of the next stage of this study.

2.29 The assessment of deliverability, and the extent to which the method could be applied with modest resource requirements is more nuanced, and somewhat subjective. As noted above, this does depend in part on the level at which the analysis is focused, and the method could be applied to different levels of rigour and alignment with core principles; indeed, as these are generally relative novel methodologies, and based principally on qualitative rather than quantitative techniques, there is scope for considerable variation in application.

2.30 This said, several methodologies have been discounted that would require very significant practitioner and stakeholder inputs throughout the process (such as Outcome Harvesting and Most Significant Change, which are also more suited to developmental and learning evaluation requirements). QCA is highly technical (applying complex modelling techniques combining qualitative and quantitative data) and is not considered likely to be deliverable in a highly complex context such as HEIF.

2.31 Of the two methodologies that meet the conceptual criteria, Contribution Analysis is considered to be relatively deliverable in terms of resource intensity. Inputs would be required to help develop the ToC, and the approach would involve primary research to test

[16] Where a tick is shown inside brackets, this indicates that our literature review identified at least one example of this methodology being applied within a KE context, or within a complex context similar to KE.

the ToC and gather evidence on other contributory factors explaining the relationships between activities and outputs/outcomes. However, the focus of Contribution Analysis is on producing a 'plausible contribution story' (rather than definitive evidence) that can help to manage the level of input required and does not follow a highly prescribed analytical approach.

2.32 As indicated in Table 2-3 (and noted previously), Realist Evaluation is potentially more resource intensive, in part owing to the required methodological approach, which required developing so-called CMOCs (context-mechanism-outcome configurations) for each outcome in question as the core analytical method. Further details of the two methodologies, and the implications of this for their potential application to inform the evaluation of HEIF, are set out in Section 3.

3. Conceptual application of methodologies

Short-listed evaluation approaches

3.1 The two approaches considered to be most appropriate for further investigation - based on the literature review and testing against key criteria and requirements of the evaluation of HEIF - are explained in more detail below.

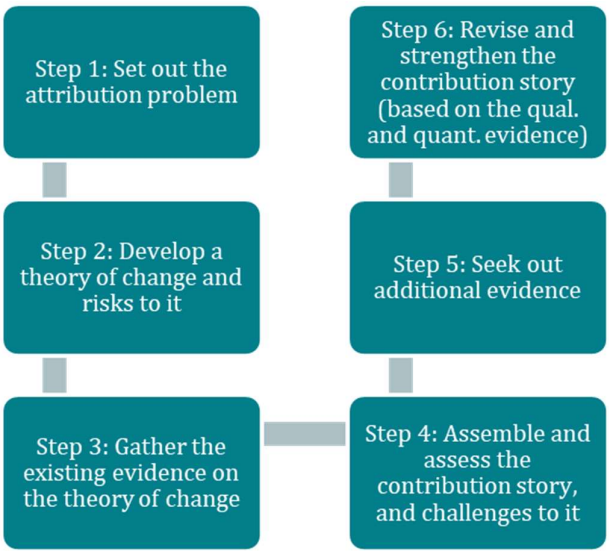
Contribution Analysis

3.2 Contribution Analysis is a theory-driven approach that "*aims to define the links between each element of a logic model, and test and refine these theoretical links between the programme and the expected impacts. It provides a framework for analysing not just whether the programme has had an impact, but how that impact materialised and whether any particular element of the programme or contextual factors were crucial to the impact*"[17].

3.3 In doing so, it builds up evidence to explain and demonstrate the contribution an intervention makes to subsequent outcomes, whilst also establishing the relative importance of wider factors. This produces a 'contribution story' about the influence that the intervention itself (instead of or alongside other factors) has had on the realisation of observed outcomes.

3.4 Formal Contribution Analysis uses an iterative six step process (set out in Figure 3-1) of evidence gathering and analysis to compare a postulated theory of change to the evidence of what happened in practice. Important in this process is testing the assumptions, barriers/risks, and other factors that may explain how outcomes have been realised. This process enables an evaluation to provide a plausible explanation based on the evidence as to how far an intervention programme has progressed in line with the logic model.

Figure 3-1: Six steps of contribution analysis



Source: Mayne, 2008, Contribution Analysis: An Approach to Exploring Cause and Effect, ILAC Brief 16

[17] Innovate UK (2018) Evaluation Framework. How we assess our impact on business and the economy.

3.5 Contribution Analysis does not prescribe specific methods and often involves a mixed-methods approach, so is highly adaptable to intervention context and can draw on qualitative and quantitative evidence (including for example monitoring data and HE-BCI survey data).

3.6 The characteristics of Contribution Analysis, which are relevant to KE, include that the approach is applicable in highly complex contexts e.g. where causality is 'recursive' (i.e. a realised outcome feeds back to and strengthens its own causes, generating further outcomes), and where outcomes are brought about as a result of lots of different relationships, meaning attribution of outcomes to a single cause is challenging. This is potentially particularly relevant for HEIF; Contribution Analysis does not seek to establish a single decisive causal factor or explanation for how an outcome has been realised, but acknowledges the interplay between various partial causes, and aims to assess their relative importance. This would enable an evaluation to comment on the role that HEIF has played alongside other factors, and its relative influence. The approach can also cover non-standard pathways and time-paths to impact, and where intended outcomes may be difficult to trace or quantify.

3.7 Contribution Analysis is predicated on a well-articulated ToC. It is therefore suited to non-experimental contexts (i.e. not pilots, or new interventions), interventions where the anticipated routes to outcomes and impacts can be identified (and subsequently tested), and where the scope for large variation in the nature of activity delivered is limited. Given the flexibility of the use of HEIF, this suggests that the approach would be most appropriate at a clearly-defined function-level, focused on the distinct strands of KE activity, thereby limiting the level of variance in implementation i.e. the evaluation is focused on the delivery of the same or similar activities across institutions, albeit delivered in different contexts.

3.8 Contribution Analysis explicitly seeks to account for the wide range of external factors and conditions into the analysis and assessment of how impact has been delivered, which is an important consideration given the range of factors that may influence KE outcomes (as discussed further below), and the diversity of HEIs and their operating and strategic contexts.

Realist Evaluation

3.9 Realist Evaluation is an approach that seeks to examine '*what works, for whom, to what extent, and in what contexts*'. Realist approaches – which are formally a way of thinking rather than a specific evaluation method – assume that nothing works everywhere or for everyone, and that context really does make a difference to outcomes[18]. Realist Evaluation approaches have a particular focus on understanding how causation works, and why programme outcomes work (or do not work) in different contexts.

3.10 Realist evaluation seeks to identify the 'generative mechanisms' that enable an intervention to achieve results, including those that influence its success in different contexts. In doing so, it can then examine the extent to which different contexts influence and/or cause activities to generate outcomes. Key to this is establishing how the 'mechanisms' interact with the 'context' (e.g. historical, cultural, location, economic etc.) to produce 'outcomes', i.e. mechanism + context = outcome.

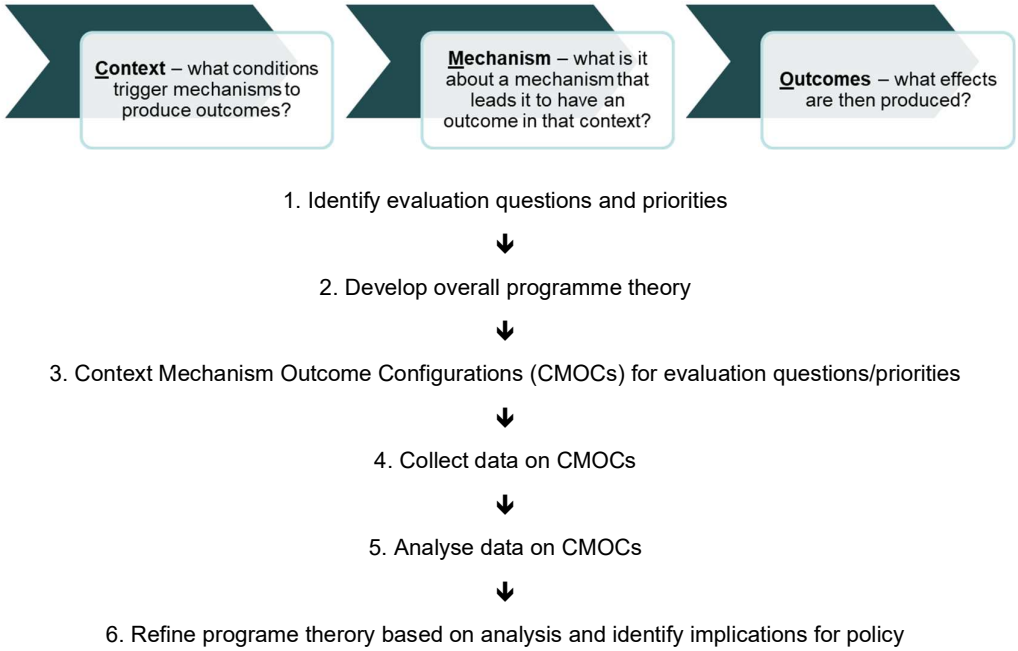
3.11 A ToC establishes the prevailing theory based on existing evidence/experience. Evaluation research and analysis then considers how the planned activities, their target population and

[18] Methods Lab, 2014, Realist Impact Evaluation, An Introduction

contexts will interact to produce a series of mini-theories called CMOCs, relating the various contexts to the multiple mechanisms to produce the various outcomes. It is often an iterative process of theory building, testing and refinement, which in turn allows causal statements about attribution to be made.

3.12 The CMOCs should be developed for those areas of interest identified for the evaluation, rather than all necessary elements of the intervention or programme subject to evaluation. Data is then collected and analysed against each of the hypothesised CMOCs to test and refine the underpinning ToC and identify how and in what contexts the intervention has generated outcomes. A broad summary of the overall approach and underpinning theory is set out in Figure 3-2.

Figure 3-2: Overview of realis evaluation approach



Source: Based on Pawson, R. and Tilley, N. (1997) Realistic Evaluation and Methods Lab, 2014, Realist Impact Evaluation, An Introduction

3.13 Consistent with Contribution Analysis, Realist Evaluation is method-neutral and can involve mixed-methods, so it is highly adaptable to intervention context and can draw on qualitative and quantitative evidence (including for example monitoring data and HE-BCI survey data). The characteristics of Contribution Analysis, which are relevant to KE, include that the approach is well-suited to complex, multi-faceted interventions. It is focused explicitly on testing programmes and interventions that appear to 'work' but where 'how and for who' in not yet understood fully, which is a key focus of the potential approaches for KE.

3.14 RE is also predicated on a well-articulated ToC. However, it is also recognised that the approach is potentially more resource intensive, including both the development of an overarching ToC (for example, for a KE function), and the individual CMOCs that are required for each relevant outcome to be tested in evaluation research: each CMOC must be able to be read "as a sentence" i.e. "In context X, mechanism Y, generates outcome Z". Given the range of contexts and mechanisms for KE functions and the multiple outcomes within each function, at the level of a function this is potentially a very significant exercise, which may require

significant inputs from practitioners. Linked to this, as illustrated in the overview, Realist Evaluation is a highly technical and complex method that requires significant technical and subject matter and technical expertise, which may have implications for accessibility and deliverability.

3.15 Taken together, the principles of Realist Evaluation are highly relevant for the purposes of the HEIF evaluation, but practically they may not be viable, particularly at the level of KE functions, which is the key level of interest identified by Research England for the evaluation,

Testing the application of evaluation options

Purpose

3.16 As described above, both Contribution Analysis and Realist Evaluation rely on the identification of a ToC for the intervention subject to evaluation. A ToC is based on a 'logic model' that sets out the key building blocks of an intervention (i.e. inputs, activities, outputs, and outcomes), and considers in more detail the links between these building blocks, and the assumptions, barriers and other factors that will influence the pathway from activity to outcomes.

3.17 To inform the assessment of potential methodologies for KE, the study team therefore developed two exemplar logic models and ToCs for the prioritised functions: 'Facilitating the research exploitation process', and 'Skills and human capital development, including enterprise education'.

3.18 The purpose of this process was:

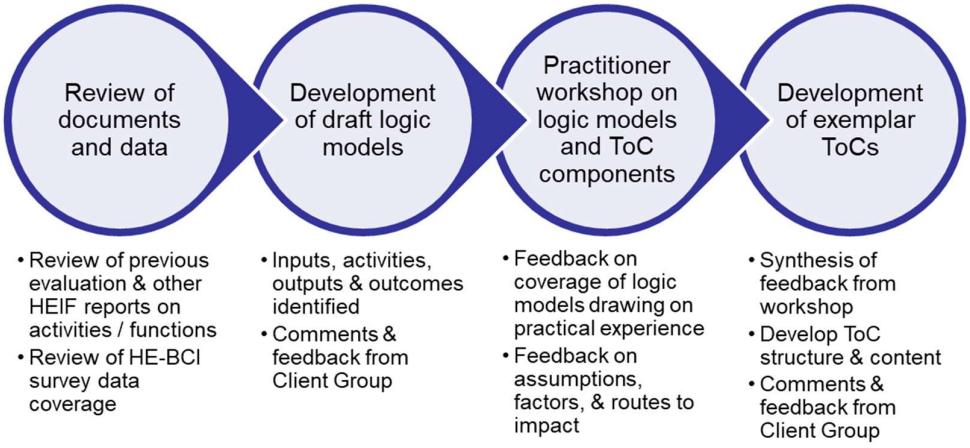
- to test in practice, the viability in the development of the ToCs for agreed KE functions
- identify the level of complexity of the ToCs and the potential pathways to impact that would need to be considered and explored in the implementation of an evaluation
- inform an assessment of the extent to which this has implications for the appropriateness of Contribution Analysis and/or Realist Evaluation for each function, and any other potential implications for evaluation approaches and recommendations.

3.19 The exemplar ToCs (and underpinning logic models) also provide the basis for an initial evaluation (updated as appropriate) in the next overall evaluation of HEIF.

Developing exemplar logic models and theories of change

3.20 The process for developing the logic models and ToCs are summarised in Figure 3-3.

Figure 3-3: Approach to the development of exemplar theories of change



Source: SQW

3.21 As noted in Section 1, the scope of this study did not involve seeking to develop a comprehensive and original list of HEIF activities, outputs and outcomes. The approach therefore drew on existing documents, data and feedback from the Client Group and KE practitioners at a workshop organised by PraxisAuril. The workshop was attended by KE practitioners from 18 HEIs across England funded by HEIF. The practitioners represented institutions covering a wide range of institution types, from large research intensive and multi-disciplinary HEIs, to small specialist HEIs.

3.22 The feedback from the practitioners focused both on the specific functions (with break-out group sessions completed for each of the two functions covered, which have informed the depictions set out below), and more generally regarding the potential for the use of novel theory-based evaluation for KE. Five points of general feedback were noted, which informed the development of the theories of change, and will need to influence evaluation approaches.

3.23 First, practitioners highlighted the importance of recognising the **cross-cutting and enabling nature of HEIF, which is commonly used to fund staff that work *across* as well as *within* specific KE functions**. Given the breadth of activity covered by HEIF it was recognised by practitioners that the function-level focus was important, and that the work was seeking to enable Research England to evidence more fully the understand the value created by different KE functions. However, the 'connecting and translational' role of KE practitioners supported by HEIF across functions will need to be reflected in the overall conceptual framework for the theory-based evaluation.

3.24 Second, and linked to this, the **relationships between function activities and routes to outputs/outcomes was highlighted**. For example, for the '*Skills and human capital development, including enterprise education*', outcomes from formal/informal support to academics to develop their entrepreneurial capacity, such as spin-off/start-up companies, may also rely on IP advice, access to appropriate physical premises, and access to finance that are delivered via other functions. Again, this does not mean that the function-level analysis is not valuable, but that the links across functions will need to be recognised in progressing any theory-based evaluation.

3.25 Third, **practitioners highlighted that the entry points to KE can vary across functions**, with businesses, for example, engaged in multiple different ways and at different points across

the logic models. The pathways to impact are therefore expected to vary, and there are likely to be feedback loops where outputs and outcomes from activities lead to engagement in other forms of activity e.g. engagement in CPD by businesses and the outcomes this generates may lead to engagement in other forms of strategic partnerships or training.

3.26 Fourth, practitioners **highlighted the importance of the local and regional context within which individual HEIs are located in informing both the nature, and pathways to impact, of HEIF supported KE activity**. These external factors, alongside institution-level decisions and strategies will need to be considered in any theory-based evaluation methodologies.

3.27 Fifth, **the strategic effects of HEIF were highlighted as important, which function-level analysis alone will not capture fully**. This included changing institutional strategies and priorities owing to KE activities and particularly the work of KE practitioners to manage and align the range of activities, which in turn influences the level of resource and priority placed on KE activity that complements and supports HEIF. As noted above, the purpose of this work was *not* to seek to map all of the potential effects of HEIF activity, however, these strategic effects were regarded by practitioners as important outcomes that a non-quantitative approach to evaluation may help to identify, providing evidence on 'how' HEIF generates impact in this way alongside the more specific functional-level effects.

3.28 In this context, the logic models and ToCs for the two functions are set out below. Two points are noted. First, the ToCs do not seek to identify the links between functions and related feedback loops, as these will span the full range of KE functions. These elements of the ToC are important, and they would need to be identified in the delivery of a theory-based evaluation (as discussed in more detail in Section 4).

3.29 Second, the TOCs identify (using red connectors and outlines) the pathways to outputs and outcomes that are captured in the HE-BCI survey, and which may be the focus of the evaluation (consistent with the priorities discussed in Section 2).

- The boxes with 'thick' red outlines identify the outputs/outcomes that are captured explicitly in HE-BCI (e.g. income, from various sources)
- The boxes with 'thin' red outlines identify the intermediate outputs/outcomes that lead to final outcomes, but that are not themselves captured explicitly
- The red connecting lines identify the causal links between activities, outputs and final outcomes captured in HE-BCI. In some cases (specifically when focused on income), this link is direct from activities to outcomes (where the delivery of this activity is recorded in HE-BCI as a direct source of income e.g. income from CPD activities).

3.30 TOCs that shows only these pathways to evidenced outcomes are presented in Annex B for clarity.

Facilitating the research exploitation process[19]

Inputs	Activities	Outputs	Outcomes

• HEIF funding
• HEI staff time (incl. research faculty, dedicated 'change agents', KE mentors, business development staff, admin/support staff)
• Students' time (incl. research students)
• Time from customers and partner organisations
• Time from alumni (e.g. investor panels)
• Other sources of funding
  • Research councils
  • Government
  • ERDF
  • HEIs' own investments
  • Private investment (e.g. corporate partners)
• Pre-existing knowledge and technical understanding
• Pre-existing partnerships, networks and contacts
• HEI assets
  • Facilities (incl. venues for events)
  • Equipment
• HEI policies and regulatory frameworks
	• **Access points for external organisations**
	  • Online portals
	  • Support and advice
	• **Business development**
	  • Collaborative research facilitation
	  • Contract research facilitation
	  • Programme management
	  • Support and advice to academics on research exploitation opportunities
	  • Developing capabilities of HEI staff (e.g. stakeholder analysis)
	  • Client management (e.g. key partnerships)
	  • Contracts/legal support
	  • IP advice
	• **Consultancy services**
	  • Delivery of consultancy
	  • Case making (external funding applications, business plans)
	  • Management of academic consultancy activities
	• **External relations (incl. business, other HEIs, public sector)**
	  • Identification of strategic partners
	  • Engagement with strategic partners
	  • External fundraising for research
	  • Networking events
	  • Network management
	  • Student placement years
	• **Press communications and marketing (incl. awareness-raising about HEI capabilities)**
	  • Online marketing and comms
	  • Offline marketing and comms
	• **Dialogue with policy-makers**
		• Businesses engaged, including SMEs
		• Partners signposted to relevant academics / other provision
		• Collaborative relationships supported
		• Research partnerships established
		• Collaborative research projects undertaken
		• Corporate partnerships established
		• Business cases for exploitation developed
		• Businesses supported, including SMEs
		• Other external partners supported (e.g. public service providers)
		• IP applications improved/developed
		• IP applications submitted
		• Patent applications
		• Licenses
		• Disclosures
		• Consultancy projects secured
		• Contract research projects secured
		• Student-focused projects undertaken (e.g. studio projects)
		• Work placements completed
		• Strategic partnerships established
		• Employees trained / supported
		• HEI staff trained
		• Inputs to HE curriculum / offer
		• Attendance at events
		  • Free events
		  • Chargeable events
			• **Institution**
			  • Income generation, including consultancy and IP income
			  • Higher retention of graduates as university staff
			  • Stronger and more expansive/diverse networks
			  • Enhanced relationship with industry and key employers, leading to other partnerships and investments
			  • Improvement in research excellence assessments (directly via impact and indirectly)
			  • Increased KE expertise
			• **Staff**
			  • Increased cross-faculty collaboration
			  • Increased awareness of industry/business needs, informing teaching and research
			  • Patents filed
			• **Students**
			  • More practical/applied teaching content
			  • Enhanced skills
			• **External partners**
			  • Technical solutions / challenges / opportunities addressed
			  • Improved performance (e.g. profit, productivity, turnover and lower costs)
			• **Society and the economy**
			  • Increased transfer of technology and knowledge from HEIs to business and wider markets
			  • Improved policy-making
			  • Wider economic and social benefits

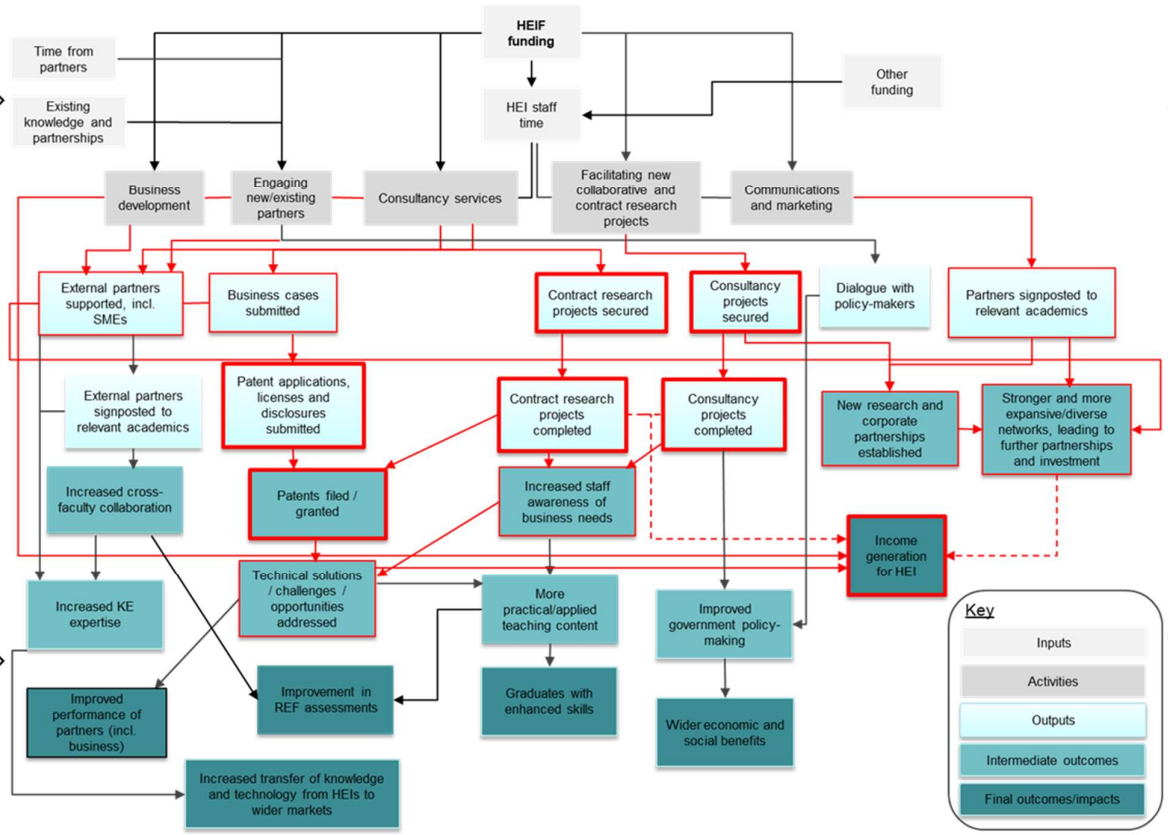
[19] KE practitioners at our workshop noted that student placements are relevant to this function as well as to skills and human capital development. However, this logic model is primarily focused on engagement with external partners specifically for the purpose of establishing partners, and so student placements are not included in this function-level logic model (but are included in skills and human capital development).

Assumptions …

- Regulatory frameworks influence the outputs and outcomes that HEIs seek to deliver (as opposed to simply influencing metrics reported by HEIs)
- HEIF provides flexibility, that enables HEIs to focus on activities and target groups that meet their strategic aims and local/regional priorities
- HEIF acts as a political signal, telling HEIs KE activity is something they should be engaging in and that is expected of them
- Supporting funding streams are sufficiently long-term to enable outcomes
- HEIF resources create capacity/time for staff to mobilise resources and facilitate relationships, that otherwise would not exist
- Demand is evident, and there are effective routes to engagement with businesses, public sector and the wider community for function activity
- Local private/public sector demand shapes HEIs networks and offer, but HEIs have a wider geographic focus than just their local area/region
- Businesses in particular can struggle to articulate their additionality (esp. SMEs) and therefore need big-writing/case making support
- Businesses are aware of these limitations and aware of HEI support that is available to them

… and risks / barriers

- Activities would happen in any case, using other resource (including other Govt funding schemes)
- Practitioners prioritise other KE functions
- Path dependency (e.g. limited previous activity, level of engagement with businesses) limits quality/scale of offer and demand
- Other sources of funding that support/ scale up activities are reduced or stop
- Businesses might be unaware they can reclaim R&D expenditure and/or partner with HEIs
- Academics find the process of identifying and engaging with non-academic partners too complicated
- Policy/strategic approach to delivering funding is inefficient or uncoordinated



Factors influencing the logic model …

Internal to institutions

- **Pre-existing knowledge and experience**, including existing networks and contacts, and technical/subject matter expertise
- **Pre-existing partnerships and collaborations** with other HEIs, private companies, government bodies etc.
- **Processes, policies and infrastructure within the HEI**, including both formal policies and strategies as well as institutional/departmental culture
- **Strategic support** from KE practitioners and senior leadership, including governance, management of funding pots, network building

External to institutions

- **Regulatory frameworks,** including the KE Concordat and frameworks such as REF, KEF and TEF
- **Other sources of funding**, such as state aid
- **HEIs' geographic base**, encompassing factors including local private enterprises, local authorities (and their strategies), and skills supply
- **The priorities of external partners**. Business partners might want activities to focus on meeting their business needs, academic partners might have aims relating to their institutional mission/goals.

Source: SQW

Skills and human capital development

Inputs

- HEIF funding
- Other sources of funding
- HEI staff time (incl. research faculty, trainers, faculty involved in delivery of career services, admin/support staff)
- Time from training and work placement providers
- Time from students, business and academics to undergo training

Activities

- CPD/short courses including:
  - career development opportunities
  - short courses/evening courses
  - bespoke training
  - workshops on business skills, personal transferrable skills, entrepreneurship etc.
- Lifelong learning including:
  - for academics and students
  - professional and personal education
  - employability
- Work placements/project experience
  - apprenticeships/internships/work experience
  - support for work-based placements and student consultancy projects
- Curriculum development
  - engagement with external orgs e.g. employers, trade bodies
  - engagement with academic community to translate external inputs in the curriculum
- Enterprise and Entrepreneurship training including on:
  - entrepreneurship education courses
  - executive education
  - technical/ negotiation skills
  - support for staff engagement / capacity in enterprise (one-to-one, one-to-many)
  - enterprise-related careers advice to students

Outputs

- Businesses supported, including SMEs
- Individuals trained
  - in businesses / public sector
  - students
  - HEI staff
  - in the wider local community (non-employer led
- Learner days delivered
- Qualifications secured
- Memberships of professional bodies secured
- Student placements/project experience delivered
  - KTP
  - non-KTP
- Business benefiting from placements/project experience
  - KTP
  - Non-KTP
- New curricula developed that includes inputs from external organisations
- Revised curricula that includes inputs from external organisations
- Individuals with increased knowledge, skills and competence in enterprise / entrepreneurship
- Students with increased knowledge and awareness of enterprise as a career opportunities
- Networks developed with local/regional businesses

Outcomes

To the institution:

- Income generation
- Improved teaching and research capabilities
- Enhanced relationship with industry and key employers, leading to other partnerships and investments
- Enhanced reputation for graduate employment opportunities, leading to increased applications
- Increased retention of graduate talent locally/regionally

To target groups:

- Skills development for key actors incl:
  - Students
  - HEI staff/academics (incl capacity to engage in KE)
  - in businesses / public sector
  - in the wider local community (non-employer led)
- Employment outcomes
  - Improved employability and career outcomes for students
  - recruitment outcomes for businesses (positions filled, new talent recruited)
- Improved business performance (e.g. profit, productivity, turnover)
- Improved public sector performance (e.g. efficiency, service quality)
- Increased access to education
- Increased entrepreneurial attitude and competence
  - Students; HEI staff; in the wider local community (non-employer led)
- New enterprises started-up (including social enterprise)
  - Graduates; staff
- Employment, turnover and investment in start-ups
  - Graduates; staff

To society/the economy:

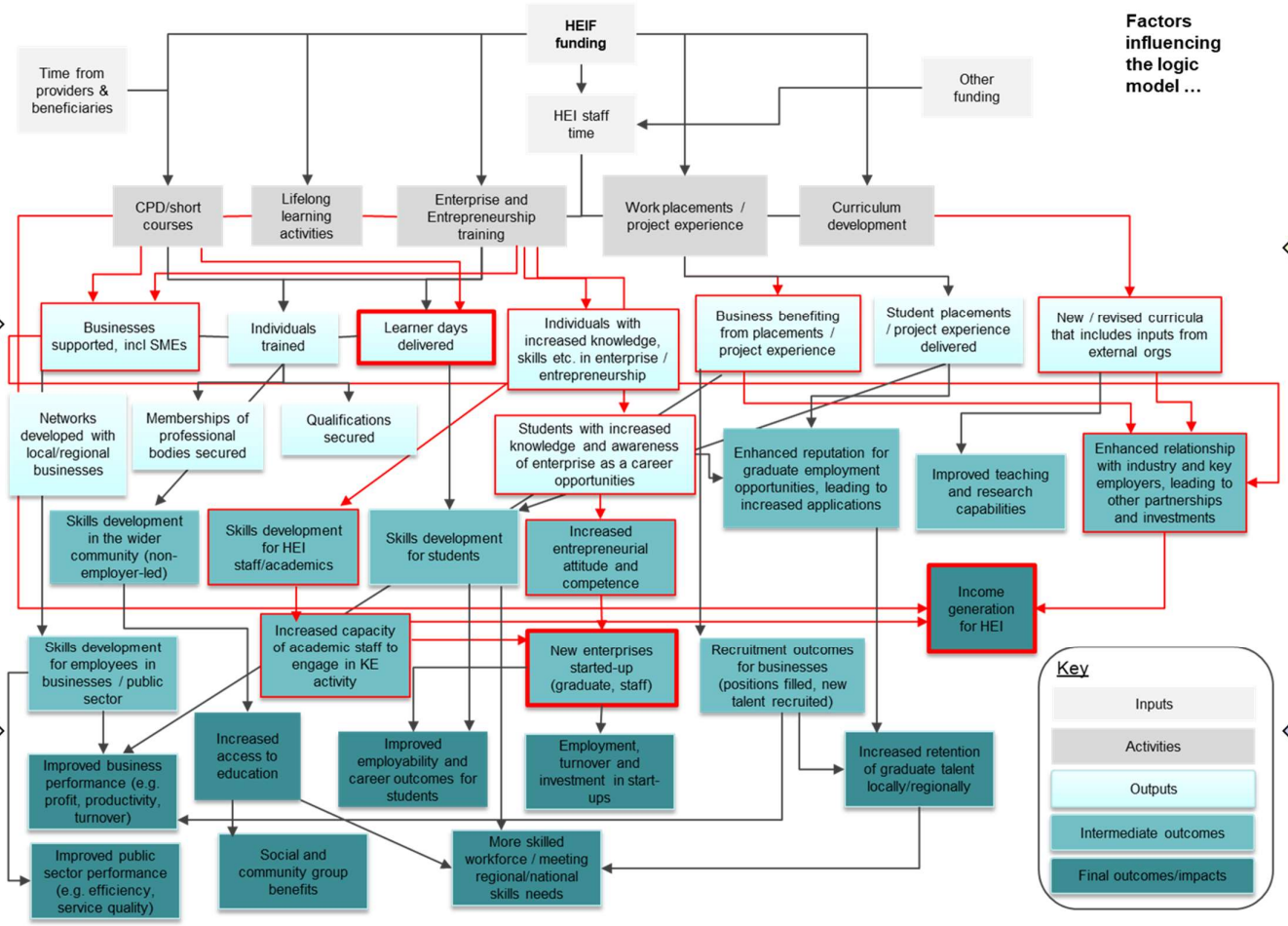
- More skilled workforce and meeting regional/national skills needs
- Social and community group benefits

Assumptions …

- HEIF enables delivery of skills and human capital development activity that would not otherwise be delivered at the same scale/quality
- HEIF provides flexibility, that enables institutions to focus on activities and target groups that meet their strategic aims and local/regional priorities
- New or improved skills developed via support and activity formal and informal) leads to changes in behaviours and attitudes of those engaged, including further involvement in KE by staff and entrepreneurship actions by staff/students
- Demand is evident, and there are effective routes to engagement with businesses, public sector and the wider community for KE skills and human capital development activity
- Academics are conducive to inputs from external organisations in curriculum development, and other capacity development in KE
- Feedback loops are evident, with activities leading to enhanced relationships (e.g. with businesses) that leads to other forms of KE engagement and longer-term strategic partnerships
- Income generation is delivered via a range of KE routes from skills and human capital development, both directly and indirectly

… and risks / barriers

- Activities would happen in any case, using other resource (including other Govt funding schemes)
- Practitioners prioritise other KE functions (incl HEIF supported)
- Path dependency (e.g. limited previous activity, level of engagement with businesses) limits quality/scale of offer and demand
- Offer does not meet the needs of target groups, and/or barriers and market failures prevent pathways to outcomes from activities/outputs and changes in behaviours do not lead to tangible effects e.g. related to time, access to finance etc.
- Training activities duplicates or displaces other provision (e.g. crowding out private sector providers)



Factors influencing the logic model …

… Internal to institutions

- **Leadership, strategy and priority placed on KE capacity development**, influencing resource allocation (incl related to research and teaching), staff incentives, recruitment, culture etc
- **Availability of appropriate facilities and staff**, particularly related to CPD and executive education including basic 'hygiene factors' (e.g. parking, accessibility, appropriate IT infrastructure etc.)
- **Delivery of/participation in non-HEIF related schemes for KE capacity development**, which may include local/regional schemes with target groups, and access/use of other funding streams
- **Discipline focus,** and alignment to changing external demand/need
- **Complementary KE offers** e.g. incubation space and support, tech-transfer and IP advice

…External to institutions

- **Policy and regulatory environment**, influencing levels of demand and need for capacity development particularly amongst businesses / public sector, funding landscape for complementary activities
- **Local and regional economic, enterprise and spatial context** e.g. related to access to finance opportunities; scale, nature and concentration of business base; strength of business and other networks and innovation ecosystems around institutions; accessibility of institution (particularly for non-employer led activity)
- **Macro-level economic conditions** e.g. influencing levels of entrepreneurship, access to employment opportunities, business investment in skills development

Source: SQW

Implications for the evaluation framework

3.31 Drawing on the logic models and ToCs set out above, and observations from the practitioner workshop, the following implications for the evaluation framework and potential approaches emerge.

3.32 First, the development of logic models and ToCs is clearly challenging in a HEIF context, especially given the 'fuzzy' boundaries between the activities contained within individual functions and the complexity of the system. However, at function level the relationships between inputs, activities, outputs and outcomes are possible to articulate and postulate, drawing on practitioner input. This suggests that in headline terms **theory-based approaches (or the application of their principles) that rely on ToCs to account for this complexity are potentially viable for the evaluation**.

3.33 Second, the ToCs are clearly complex, reflecting the realities of KE activity. They are made particularly complex by the potential range of feedback loops and links between different functions, with links likely to be evident across all functions. This suggests that **a Realist Evaluation approach (which requires the development of multiple CMOCs[20]) is not likely to be viable and proportionate at the function level**, given the very wide range of contexts and mechanisms that would need to be considered.

3.34 Third, there are some challenges associated with Contribution Analysis in a HEIF context, including variation in delivery models between HEIs (noted earlier in Section 3). However, in developing the ToCs we have identified each function's underpinning assumptions and risks, and influencing factors (both internal and external to HEIs). This could inform and enable a Contribution Analysis approach at the function level; therefore, **Contribution Analysis appears to be the most appropriate method identified and considered in the study**, meeting the stated requirements of the theory-based evaluation.

3.35 Fourth, given the broad coverage of the ToCs and the complexity noted above**, a Contribution Analysis approach should potentially focus on those outputs and subsequent outcomes that are already demonstrated in the existing evidence base[21],** and which are linked to specific activities, limiting the potential level of variation in implementation. For these outcomes, the role and relative contribution of HEIF could be tested throughout the identified pathway to impact, taking into account the internal and external factors identified and considering the assumptions and risks, providing evidence on 'how' HEIF has created value within that function. The relative contribution of HEIF may be found to vary by outcome within the function, which is not unexpected, and this would inform an overall 'contribution story' at the function-level.

3.36 Fifth, a Contribution Analysis approach could also consider equivalent analyses for the other outcomes identified, to provide a fuller picture of the value created by each KE function. However, without an existing evidence base on these outcomes, these analyses would need to provide an assessment of both 'how' and 'what' impact has been generated; this would require gathering primary data (for example on effects on businesses, policy makers, academics) in order to confirm that outcomes have been realised, which may not be a priority. Potentially,

[20] Including multiple different configurations for each outcome.
[21] e.g. in the HE-BCI survey

the ToCs could be used as the basis for a Realist Evaluation examining these individual outcomes, examining selected cases (e.g. a selection of HEIs initially, which could be scaled-up as appropriate) to manage the scope of any required primary research. This would first evidence a specific outcome (e.g. an increased capacity of academic staff to engage in KE activity), then work backwards from this to identify the mechanisms that enabled the outcomes to be realised, and then the contexts in which the mechanisms did or did not enable this outcome to be realised. This could supplement and inform the overall Contribution Analysis for the relevant function. This outcome-specific approach would also be possible for outcomes that are already evidenced and could be considered if Research England was interested in assessing specific pathways to outcomes in more detail (which may be of particular policy interest). The potential approach is discussed in the Section 4.

3.37 Sixth, an important consideration is how the 'connecting and translational' role of KE practitioners working across HEIF across functions will be accounted for in the analysis. In our view, drawing on the feedback from practitioners, the variation in how this role is delivered across institutions precludes the development of a detailed ToC for this activity, and therefore a formal Contribution Analysis or other theory-based evaluation approach specifically; this 'connecting and translational' role essentially feeds into most (if not all) of the potential ToCs across functions. HEIF-funded staff inputs within individual functions will be captured via the function-level Contribution Analysis. However, **to provide a full and rounded explanation for 'how' HEIF has generated outcomes, the analysis will need to consider how KE practitioners have influenced outcomes by making links between the functions**, and potentially generated other benefits that are 'greater than the sum of the parts' (including strategic effects), helping demonstrated *how* HEIF has generated value.

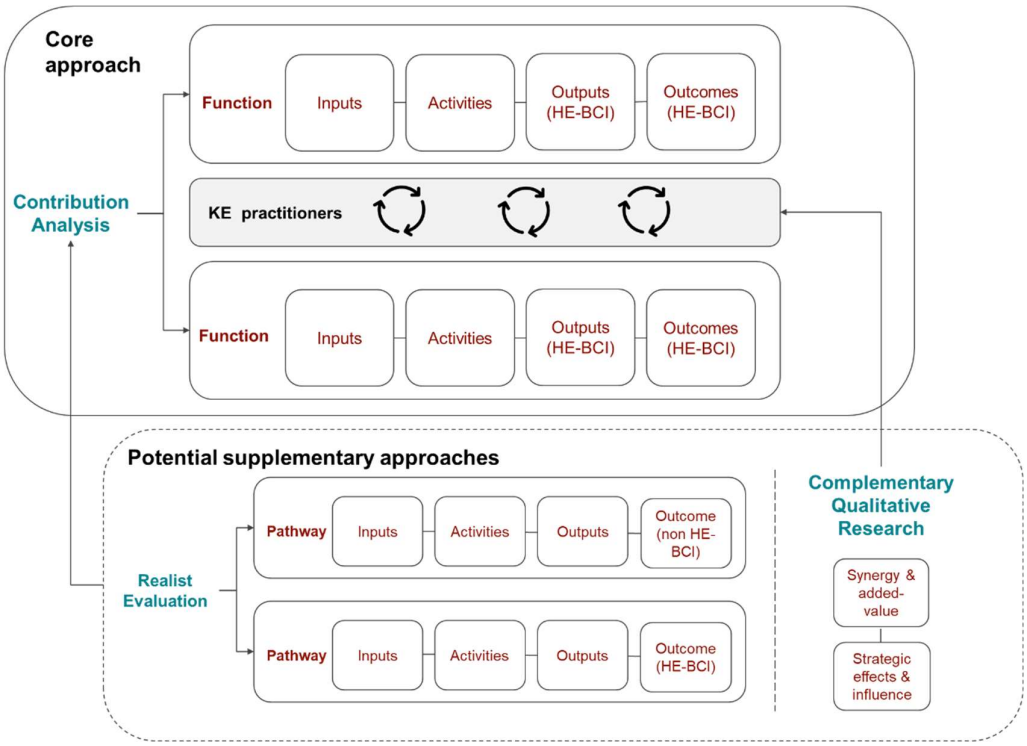
3.38 Our approach to this issue is set out in more detail, alongside the wider proposed conceptual framework in Section 4.

4. Proposed evaluation framework

Conceptual framework

4.1 The proposed conceptual framework for theory-based evaluation approaches to complement quantitative evaluation as part of the next evaluation of HEIF is set out in Figure 4-1.

Figure 4-1: Conceptual framework for novel evaluation approaches in qualitative evaluation



Source: SQW

4.2 The conceptual framework operates at two levels:

- A '**core approach**', involving Contribution Analysis at the level of KE functions, focusing on those outcomes that are captured in the HE-BCI survey, providing evidence on how HEIF has contributed to outcomes by arriving at a plausible explanation of its role relative to other factors and influences.
- Two '**potential supplementary approaches**' that would seek to generate additional evidence on how HEIF generates value and the full range of value within and across functions. The two approaches would involve:
    - fine-grained analysis at the level of particular outcomes of interest, including those which are not captured in HE-BCI survey data, to generate a fuller picture of how HEIF generates outcomes through activity within KE functions. The outcome-specific focus of this level of analysis would enable the use of Realist Evaluation techniques. This evidence would be used to inform the overall Contribution Analysis, utilising the evidence generated on '*what*

works, for whom, to what extent, and in what contexts' from the Realist Evaluation to augment and exemplify in detail the role of HEIF.

- complementary qualitative research on the 'connecting and translational' role of KE practitioners supported by HEIF, seeking to deepen the insight generated at a function level, focusing on the synergy and added value, and wider strategic effects and influence this role generates across institutions.

Considerations

4.3 Five points are highlighted in relation to the conceptual framework. First, considering the linkages between functions, the need for a practical and proportionate method of research, and the strengths/weaknesses of methodologies that are consistent across functions, the Contribution Analysis approach is proposed for both functions covered in detail in this research, and we suggest for the other five functions currently framing KE and HEIF activity. The exemplar ToCs demonstrate that whilst within functions there is variation in the extent to which some pathways to impact are relatively simple (in theory), and other are more complex, both functions share similar issues in terms of the wide range of internal and external influencing factors, linkages to other function areas, and non-linear relationships between activities, outputs and outcomes. Therefore, **a consistent conceptual and analytical approach across functions is proposed**. Put simply, the study suggests that different novel methodologies should not be used for different KE functions, rather, different novel methodologies (i.e. Contribution Analysis and Realist Evaluation) are more appropriate to be applied at different levels of activity: function level and (potentially) individual outcome level, as illustrated in the conceptual framework.

4.4 Second, Contribution Analysis at the function level could in principle be completed individually by function, meaning that not all functions would need to be covered in a single evaluation process. However, an important element in the approach will be to consider explicitly during the data collection and analysis, the role of activities within other KE functions in contributing to outputs and outcomes realised. At this scoping stage, these linkages have not been developed fully across all of the functions (with a focus on developing exemplar logic models and ToCs for two functions only). If the approach is progressed, it will be important to develop initial logic models for the other functions and identify the key linkages between functions in advance of undertaking function level analysis. Whilst the challenges of the 'fuzziness' of functions is recognised – as discussed at the practitioner workshop – this framework has been applied by Research England for HEIF expenditure analysis and has also been used in previous evaluation and research studies as an organising framework for HEIF support for KE. Activities will need to be allocated to specific functions, so the links and relationships between them can be established transparently to inform the Contribution Analysis.

4.5 Third, it is recognised that the outcome-specific analysis element of the potential supplementary approach is not currently considered a priority for Research England. This should therefore be regarded as a potential exploratory option that could be implemented in order to strengthen the evidence base but is not essential. The core approach could be implemented independently without this complementary approach.

4.6 Fourth, considering the 'connecting and translational' role of KE practitioners working across HEIF across functions, as illustrated in the framework, we anticipate this issue would be addressed with a combination of 'bottom-up' and 'top-down' evidence.

- Bottom-up: as part of the function level Contribution Analysis, and in the development of the final ToCs, the 'connecting and translational' role of KE practitioners working across HEIF across functions should be identified explicitly as an assumption that underpins the delivery of activities to be tested via consultation. In some cases, it may be found that this assumption does not hold (i.e. the 'connecting and translational' role is not an important factor in explaining how outputs and outcomes are delivered via HEIF), whereas in other it does. This should be considered in the analysis explicitly and inform the overall function-level contribution story in relation to the identified outcomes.

- Top-down: complementing this, as a potential supplementary approach, a specific programme of qualitative research to consider how the 'connecting and translational' role of KE practitioners operates across HEIF funded institutions, alongside how and why this influences activities and reported outcomes.

4.7 The findings from the two approaches could be brought together to provide an integrated assessment of the contribution of the 'connecting and translational' role of KE practitioners working across HEIF across functions, and whether this has impacted on how HEIF has generated value and impacts, both at the level of individual functions and in terms of strategic effects and influence. This could be reported as part of the overall evaluation of HEIF.

4.8 Fifth, all potential theory-based evaluation methodologies will necessarily involve the engagement of those involved in the delivery of the intervention, meaning that some primary research with institutions and KE practitioners is required. The proposed approach to Contribution Analysis seeks to minimise the level of burden through focusing particularly on those outcomes where there is existing evidence from HE-BCI on the delivery and scale of outputs/outcomes. As a method-neutral approach, which can draw on mixed-methods, Contribution Analysis will also draw on documents and other data. However, testing the ToCs and assessments of relative contribution will require primary qualitative research with institutions and the wider KE landscape. The complementary research, and outcome-specific research (if progressed) would also require engagement. The next Section sets out in more detail what the implementation of the conceptual framework could involve practically.

5. Implementation

5.1 This Section sets out in further detail the proposed methodologies in the conceptual framework and how these could be implemented practically to inform the overall evaluation of HEIF.

Core approach

Contribution Analysis

5.2 The proposed implementation of the Contribution Analysis approach is set out below. In practice, we recognise that the research is likely to be undertaken in parallel to and as part of a broader programme of cross-sectional evaluation research of HEIF. This may lead to some cross-overs and opportunities for shared research activities and data collection (and minimise research burden on participants). However, at this point, the timing and format of this wider research is not confirmed, meaning we have set out a specific research process for the Contribution Analysis.

5.3 We have also assumed that this work will be progressed across all KE functions in parallel, although this is not essential. The set-up phase should be completed for all functions to ensure consistency and coverage of the links between functions, however, the data collection and reporting could be completed for individual functions at different points and/or for a sub-set of functions. This said, there is also likely to be value in completing the Contribution Analysis across the full range of KE activity to allow for economies in delivery (e.g. with an overall synthesis and reporting as part of the full HEIF evaluation).

5.4 Three phases of work are proposed which map on to the six stages of Contribution Analysis, as set out in Table 5-1

Table 5-1: Implementation of Contribution Analysis

Research Phase	Stage in Contribution Analysis
A. Set-up and establish theory of change	1. Set out the attribution problem 2. Develop a ToC and risks to it
B. Data collection	3. Gather the existing evidence on the ToC
C. Analysis and reporting	4. Assemble and assess the contribution story, and challenges to it 5. Seek out additional evidence 6. Revise and strengthen the contribution story

Source: SQW

Phase A: Set-up and establish theory of change

5.5 The first step in the Contribution Analysis will be to set-up the conceptual approach and ToC for each function. This will involve initially defining the attribution problem to be tested. In this case of this work, this will be consistent across each function, reflecting the purpose of the theory-based evaluation within the context of the wider evaluation of HEIF.

5.6 The following two core research questions are proposed as the basis for the Contribution Analysis of HEIF functions:

- *What role did HEIF play in bringing about the KE outputs/outcomes in this function?*
- *How important was this role relative to other factors in explaining how and why the outputs/outcomes were realised?*

5.7 The next stage will involve developing the ToC for each function. This will follow a similar process to that undertaken as part of this study for the two exemplar ToCs, and involve a mix of review of data and documents to inform an initial logic model that sets out the activities, outputs and outcomes of relevance to the function, and then engagement with KE practitioners that work within the relevant function area to translate this into a ToC (see Figure 3-3).

5.8 Alongside establishing the expected pathways to outputs and outcomes, this will involve identifying the relevant assumptions and risks/barriers relevant to the ToC, and the internal and external contexts that will need to be considered. The two exemplar ToCs should also be reviewed and re-assessed in advance of evaluation research to reflect any potential changes in context at the point of the evaluation. The 'connecting and translational' role of KE practitioners working across HEIF across functions should be identified explicitly in the ToCs, including how this may vary across functions.

5.9 However, it should also be noted that reflecting the focus of the Contribution Analysis, the ToCs should focus explicitly on those outputs/outcomes that are covered via existing HE-BCI survey data (with the wider outcomes in the exemplar ToCs removed for the purposes of this element of the evaluation[22]).

5.10 When each of the function-level ToCs has been established, the key links between activities and outputs/outcomes *across* functions should be identified and mapped to inform the analysis. This process should not involve seeking to establish every possible relationship between functions (which is not likely to be viable, even where the focus is on outputs/outcomes covered in HE-BCI survey data), rather it should identify those relationships that are considered to be most important based on practitioner experience, and theory. This is likely to be an iterative process, that may require a series of engagements with practitioners with experience across functions. In the data collection stage, other linkages may be established via primary research, which can be used to re-assess the ToC at the analysis stage.

Phase B: Data collection

5.11 With the ToCs established, the Contribution Analysis will turn to data gathering to provide the evidence for testing the ToCs and answering the research questions. We are aware that Research England is exploring approaches to the HEIF strategy and plans templates to identify data needed for assurance of HEIF for the next HEIF round due to begin from 2020-21. This may potentially offer opportunities to systematically collect information of planned activities and investments that can be used to minimise the requirements for primary data collection.

[22] Those activities and outputs that are not covered in HE-BCI, but which are hypothesised to lead to outcomes that are, should be retained in the theories of change in order for these links to be tested.

However, dependent on the timing of the evaluation, this may not be viable. Further, if there is scope to seek evidence specifically related to functions in monitoring processes (e.g. via the Annual Monitoring Statement returns), then that would aid the analysis.

5.12 However, these issues noted, it is proposed that the Contribution Analysis data gathering would involve four principal strands of research.

- **Desk review of documents and data**: this would involve (i) a review of the existing evidence from previous reports and evaluations regarding the activity delivered within the function, the outputs and outcomes this has generated, and any available evidence on the role of HEIF in this. This should include evidence developed centrally by Research England and any institution-level evaluations and studies that may be available (identified potentially through a 'call for evidence' from institutions). (ii) the collation of relevant data from the HE-BCI for the outputs/outcomes relevant to the function for the period covered by the evaluation and (iii) a desk-based review of Annual Monitoring Statements from institutions for the period covered by the evaluation. This review of the Annual Monitoring Statements would involve the development of an appropriate coding framework to capture evidence for each relevant function (e.g. based on activity and function descriptions) and the qualitative feedback from institutions relevant to the different elements of the ToC i.e. the uses of HEIF, links and relationships to other KE funding mechanisms, barriers and challenges experienced, external and internal factors influencing outputs/outcomes (including the role of KE practitioners), and other relevant evidence impacting on the ToC. Evidence would be coded by key contexts (e.g. institution type, timing), to inform systematic qualitative analysis in the development of the contribution story.

- **Practitioner survey**: a survey of practitioners across HEIF funded institutions, segmented by function involvement/engagement. This would be a census survey (i.e. sampling all institutions in receipt of HEIF) and focused on those individuals (as identified by institutions) that can comment on the activity delivered within the relevant function. At this stage, we expect that multiple respondents per institution could be identified to ensure sufficient coverage across and within function areas. The survey would be undertaken via telephone to maximise engagement (with an online survey option available where preferred), and would gather data on:

  - the nature and scale of activities delivered by the institution within the function using HEIF resource and non-HEIF based on the activities identified in the logic model/theory of change, and how this compared to plans

  - why specifically HEIF resource was required to deliver the activity, what other options were considered, and why they were rejected

  - the nature of influence of HEIF on behaviours, attitudes, and approaches in delivery of function activity (e.g. strategic planning, management, nature of the delivery offer etc.)

  - the nature and scale of influence of other factors on behaviours, attitudes, and approaches in delivery of function activity

  - for each of the outputs/outcomes reported in HE-BCI by the institution:

- the level of HEIF resource additionality i.e. what is the most likely 'counterfactual situation' for this output/outcome without HEIF, including relating to timing, scale, and quality
- other factors that have contributed to the delivery of outputs/outcomes (internal and external)
- the relative importance of HEIF resource to these other factors i.e. whether it had no influence; contributed to outputs/outcomes but was not necessary; was an important contributory factor alongside others; was the critical contributory factor.

- **In-depth interviews with a KE director/senior manager of institutions in receipt of HEIF**: these interviews would seek to explore in depth with senior-level staff at institutions the relative contribution of HEIF alongside other factors influencing HE-BCI outputs and outcomes, and how this may vary (and inter-relate) across different functions. The aim would not be to engage with all institutions, rather to cover a representative sample of respondents by institution type (likely sampled by the HE clusters, which also aligns broadly with the level of HEIF resource), which is an important contextual factor in the exemplar ToCs, with at least 2-3 institutions from each group covered. These interviews would complement the survey evidence and play an important role in considering the linkages between functions, the role of KE practitioners in facilitating these linkages, the factors influencing decision making regarding the use of HEIF, and its relative importance when compared to other sources of funding in the delivery of KE activities. We anticipate the interviews would be held face-to-face. The evidence would be analysed, with key findings relevant to each function coded and cross-cutting messages identified.

- **In-depth interviews with KE director/senior managers of institutions <u>not</u> in receipt of HEIF**: Contribution Analysis does not require formally evidence from non-supported institutions. However, evidence on how the pathways to HE-BCI outcomes have been realised without HEIF would help with the contribution story and explaining how HEIF plays a part in bringing about the KE outputs/outcomes. It is important to recognise that given the funding allocation for HEIF, which is based in part on performance in HE-BCI outcomes, the institutions not in receipt of HEIF are different to HEIF supported institutions; the institutions that were not funded via HEIF in the latest funding round were generally smaller and/or teaching oriented institutions, which may have different KE arrangements and imperatives than institutions support by HEIF. This said, within the qualitative theory-based evaluation approach, this context can be considered and allowed for in the analysis and understanding of how outcomes are realised in the absence of HEIF, and where they are not and why not will help to inform the contribution story. As such, we propose that the evaluation would include 8-10 interviews with senior KE representatives from non-HEIF supported institutions as part of the data collection process. There may be practical challenges in engaging with this group that would need to be factored into the approach e.g. communicating clearly the purpose and potential benefits of the research to the overall KE funding landscape and understanding to encourage participation.

- **Stakeholder and partner interviews**: alongside the institution-level consultation, we propose a programme of stakeholder interviews with representatives from across the KE landscape that will provide additional qualitative perspectives on the observed role of HEIF within and across functions, how this relates to other factors and influences in realising HE-BCI outcomes, and its relative contribution. Given the complexity of the ToCs and context, this process should focus on a limited number of interviews with 'well-informed partners and stakeholders', rather than a broader engagement of 'interested parties'. For example, this could include interviews with representatives at Research England itself, representatives of PraxisAuril and the National Centre for Universities and Business (at Executive and potentially Board level), BEIS, Innovate UK, and a selection of LEPs/Combined Authorities with responsibility for knowledge exchange and innovation.

5.13 Contribution Analysis approaches often also employ case study research, and this would be possible, and potentially add value to the evaluation. However, for this evaluation, for case studies to be meaningful at a function level, this would potentially require a very significant volume of research: for example, five case studies per function (to ensure that the case studies are broadly representative of different institution types) would involve potentially 30-35 case studies. At this stage, case studies are therefore not proposed, however, if Research England considered prioritising particular functions, then case studies should be considered, which would seek to engage with institutions to consider in more detail their experience of the pathways to outcomes in the relevant function, and the relative contribution of HEIF and other factors.

Phase C: Analysis

5.14 With the data collected, the Contribution Analysis will move on to the synthesis and reporting stage. For each of the functions, this will involve drawing on the range of evidence from the documents and data, survey, and interviews, to assemble and assess the 'contribution story', framed around the research questions i.e. what is the evidence on the role that HEIF has played in bringing about the KE outputs/outcomes in each function, and its importance relative to other factors in explaining how and why the outputs/outcomes were realised.

5.15 In the context of the attribution problem and research questions, a plausible 'contribution story' would be evident if[23]:

- a reasoned ToC for the function has been established, and the role that HEIF played in the activities anticipated in the ToC was realised in practice
- the chain of expected outputs and outcomes can be shown to have occurred
- other influencing factors have been shown not to have made a difference or the decisive difference, or HEIF has been shown to have made a difference relative to or alongside other influencing factors.

5.16 We anticipate that the initial 'contribution stories' would then be tested via a series of workshops with practitioners and partners/stakeholders that have been involved in the

[23] Based on White and Phillips (2012) *Addressing Attribution of Cause and Effect in Small n Impact Evaluations*, International Initiative for Impact Evaluation Working Paper 3.

research. The feedback of the workshop would inform the need for any additional analysis/interpretation of the evidence, to revise and finalise the 'contribution story' for each function.

5.17 The function-level contribution stories could then inform an overall report drawing on the Contribution Analysis from across the functions.

Supplementary approach

Realist Evaluation of outcome pathways

5.18 This approach would be used to examine individual pathways to outcomes, assessing the role of different activities (both HEIF-funded and otherwise) in delivering a single outcome or small number of closely related outcomes within functions.

5.19 Figure 3-2 (page 20) sets out the key steps involved in delivering a Realist Evaluation, which are grouped under three individual sub-sections within this supplementary approach:

- **Develop evaluation framework**: setting out the process by which an intervention is expected to lead to its outcomes/impacts.
- **Defining and collecting data**: choosing the methods and analytical approaches and tools that will produce evidence capable of answering your evaluation research questions.
- **Analysis and findings**: analysis based on a 'context-mechanism-outcome' (CMO) approach, identifying the mechanisms that enabled outcomes to be realised and the contexts in which these mechanisms did or did not do so.

Develop evaluation framework

5.20 As with Contribution Analysis, the first steps undertaken for a Realist Evaluation are to establish the study's key research questions and the theory of change that will be tested.

5.21 The following three research questions are proposed as the basis for a Realist Evaluation of pathways to impact within HEIF functions:

- *What are the mechanisms by which HEIF-funded activities lead/contribute to outcomes and impacts?*
- *What wider factors influence the translation of these activities into outputs, and outputs into outcomes/impacts?*
- *How important was the role of HEIF-funded activities relative to wider other factors in explaining how and why the outputs/outcomes were realised?*

5.22 Although we expect the above questions would be the basis of a Realist Evaluation undertaken for the purposes of this supplementary research, these research questions would be discussed, revised (if necessary) and finalised between Research England and the evaluation team prior to the following steps of this approach commencing.

5.23 The next stage will involve developing a ToC for each pathway to impact of interest for the supplementary research. This will follow a similar process to that undertaken as part of this study for the two exemplar ToCs, developing a more granular map of individual pathways within those exemplars, and involve a mix of review of data and documents to inform an initial logic model that sets out the inputs, activities and outputs of relevant to the outcome(s) of interest. This will be followed by engagement with stakeholders that work on that pathway[24] to translate the logic model into a ToC.

5.24 Alongside establishing the expected pathways to outputs and outcomes, this will involve identifying the relevant assumptions and risks/barriers relevant to the ToC, and the internal and external contexts that will need to be considered.

5.25 Once ToCs have been established for the pathway to impact of the specific outcome, the key links to other activities and outputs/outcomes in the pathway (drawing on the function-level theories of change developed for the Contribution Analysis where relevant) should be identified. Where those links have not been established (that is, where the Realist Evaluation is focused on an outcome not covered in the Contribution Analysis of HE-BCI outcomes), this mapping process should draw on stakeholder experience to identify the relationships considered to be most important to the realisation of outcomes.

5.26 This is likely to be an iterative process, that may require engagement with a range of stakeholders (depending on the nature of each pathway to impact). Primary research undertaken during the data collection stage may identify other linkages; if so these can be used to re-assess the ToC during analysis.

Defining and collecting data

5.27 The ToC sets out a hypothesis of how outcomes are realised. Subsequent research will then need to examine how exactly the realisation of these outcomes is influenced by HEI-level variation in the resources and stakeholders involved, activities undertaken, and the contexts within which they are delivered.

5.28 The scale and exact nature of outcomes realised will likely vary when different combinations of inputs, activity and context are used. At this stage in a Realist Evaluation, these combinations will be mapped out in a series of mini-theories, or CMOCs. Developing these configurations at this stage will help to frame and shape subsequent data collection and provide a more granular framework for the final analysis (that ensures all relevant variables are explored and their relative important assessed).

5.29 Each CMOC must be able to be read 'as a sentence'. Some contextual elements of likely relevance have already been identified in our function-level ToCs, as having high-level definitions of mechanisms involved in the delivery of some pathways to impact; these will need to be refined and added to using a small number of scoping interviews with key evaluation stakeholders from Research England, and potentially consultation with a panel of KE practitioners convened for the purpose of helping shape the evaluation design.

[24] Likely including (although not necessarily limited to) KE practitioners, academic staff, and external partners (incl. from academic, business and other external partners).

5.30 Data gathering will then collect evidence for testing the ToC. In order to understand the relative influence and important of different contextual variables and mechanisms, it would be necessary to undertake research with a sample of HEIs. As with the Contribution Analysis, a representative sample of HEIs would be examined. Likely sample criteria may include HEI cluster, geography, and level of HEIF funding (each of which are important contextual factors), with at least 3-4 HEIs from each group covered.

5.31 Realist Evaluation is 'method neutral' (i.e. it does not mandate any particular methods for data collection) and, as noted earlier in this chapter, the timing of and level of resources available to the next HEIF evaluation are not yet confirmed. However, data collection for a realist evaluation of a single pathway to impact would – at each sample HEI - likely involve the following key elements:

- **Desk review of documents and data**, providing evidence of inputs, activities, outputs and outcomes from the ToC being delivered/realised, plus any evidence on variations between institution in: rationale for approach taken; local context influencing progress; and enablers/barriers to realise of outcomes.
- **Interview with a KE director/senior manager,** exploring in depth the relative contribution of HEIF to the realisation of outputs and outcomes, including specific consideration of the relative contribution of other inputs/activities not funded by HEIF, the influence of differences in local context, and the specific mechanisms by which the HEI has sought to deliver outcomes.
- **Interviews with stakeholders and partners**, including individuals performing a range of different roles involved in delivery of the ToC being examined. These could include (where relevant) interviews with academic staff in the HEI, students, staff at partners (incl. but not limited to business, other HEIs, and public-sector clients/partners), plus other external bodies not involved in delivery but with a strategic interest in the outcomes (e.g. local/combined authorities). As with the KE director interview, these interviews will provide qualitative perspectives on the important contextual factors driving (or hindering) realisation of outcomes, and mechanisms utilised by different partners to do so (including those funded by HEIF or otherwise).

Analysis and findings

5.32 Once data has been collected, it will then be analysed against each of the hypothesised CMOCs developed at the outset of the evaluation. To do so, the evaluators will need to undertake the following key steps:

- Identify patterns of outcomes (e.g. outcomes commonly achieved, and any variation in outcomes between different HEI sub-groups).
- Review and code qualitative data collected during interviews, ensuring different contextual variables and mechanisms are coded consistently with each other and with the definitions used in developing the original configurations.
- Test these coded qualitative data, examining the combinations of different contextual factors and mechanisms that are most closely associated with the realisation (or non-

realisation) of the identified outcomes, including variation in combinations that are successful (or not) in different types of HEI.

5.33 This approach should produce a set of CMOC statements, that could be stated in a format such as: "*In this context, this mechanism was successfully used to generate these outcomes. In that context, a different mechanism generated these different outcomes.*"

5.34 The results of the above analysis will then be used to test and refine the pathway's original ToC. The evaluators will have identified how and in what contexts the pathway has realised outcomes, and therefore be able to conclude which ToC inputs, activities and contextual factors were most influential in generating outcomes (or, in some cases, which elements may have hindered realisation of outcomes). This will then enable them to develop a narrative setting out exactly how specific HEIF-funded inputs and activities have led to specific outcomes – helping reinforce the contribution analysis undertaken at the function- and programme-levels.

5.35 This analysis may well produce a range of levels of findings, ranging from some very specific conclusions (in this specific context, this specific mechanism is the most effective) to some more generalised ones (across different contexts, this mechanism was relatively more influential in support outcome realisation than other mechanisms).

Complementary qualitative research on the 'connecting and translational' role of KE practitioners

5.36 As discussed in Section 3, given the variation in how the role of KE practitioners across institutions supported by HEIF, the development of a detailed ToC that could inform a formal theory-based evaluation approach is not considered viable. However, as part of the supplementary framework, complementary qualitative research that applies some of the principles of theory-based evaluation is proposed in order to provide further evidence of the way in which this 'connecting and translational' role supported by HEIF may generate value.

5.37 It is proposed that this would involve the following:

- The development of a 'high-level ToC' for the 'connecting and translational' role: this would not seek to map all of the different activities that this role involves, and subsequent outputs and outcomes (i.e. which would essentially cover all of the KE outcomes at a function level), rather it would set out: why this role is considered important; alternative explanations that may challenge this; the mechanisms by which this role is anticipated to influence KE functions and wider activity; and the anticipated strategic results and influences. This could be co-produced via a targeted workshop with a number of practitioners representing different types of institutions supported by HEIF

- Testing this 'high-level ToC' via engagement with senior-level actors and decision-makers engaged in KE activity across a representative sample of institutions, via bilateral interviews or focus groups/workshop events. This would seek to gather examples and evidence to validate and/or question the underpinning logic set out in the high-level ToC

- A series of case examples, which would test the high-level ToC in particular settings in more detail via a mixed-methods approach (e.g. document review, interviews, site visit).
    - Potentially this could involve the use of QCA, which is based on the use of cases in order to answer the question 'what works best, why and under what circumstances', which may add further insight into the role of the KE practitioners. QCA is potentially useful here because it allows a robust analysis of evidence from case studies with modest sample sizes (say 15-20), to arrive at a more confident assessment of which factors alone or in combination contribute to the target outcomes, for example: increased institutional investment in KE; changes in institutional strategy etc. QCA relies on the definition of a relatively modest number of factors/conditions to test (e.g. fewer than 10) for each outcome, and with more conditions/factors, the analysis will result in the description of individual cases rather than key patterns, meaning it may not be suitable in this context. However, this could be considered in more detail if the supplementary approach is taken forward by Research England.
- Integrating these findings into the overall evaluation of HEIF, providing additional evidence on how support for staff working across functions has helped.

Other implementation considerations

Application to functions

5.38 As noted in Section 4, given the findings of the assessment of potential evaluation methodologies, and the important links between the functions highlighted by practitioners, a consistent conceptual and analytical approach across functions is proposed for any theory-based evaluation methodologies that are progressed. We do not consider that different methodologies should be used for different functions; this would also limit the extent to which the findings for different functions could be brought together and synthesised. This said, it is important to recognise that we have not developed logic models and ToCs for the other functions, which would be required in order to facilitate the proposed approach.

5.39 In the initial scoping stage of the study, it was agreed that following the development of exemplar ToCs and conceptual approach based on the 'Facilitating the research exploitation process' and 'Skills and human capital development, including enterprise education' functions, this would be validated/tested in principle against a further function, 'Commercialisation (technology transfer)'.

5.40 As set out in Section 1, this function focused particularly on two forms of activity:

- specialist advice and capacity for the institution on technology transfer, including technology sourcing, market analysis, technological and commercial due diligence, patenting and intellectual property management, licensing, spin-out creation, and legal and contract support in relation to technology transfer

- investment funds, providing finance of various kinds from proof of concept funds to demonstrate whether an idea has commercial merit, to investments (e.g. via equity, loans etc.) in start-up companies.

5.41 In considering the extent to which the proposed Contribution Analysis approach is likely to be appropriate in this context, the following points are noted:

- First, consistent with the feedback from practitioners in relation to the two functions considered in detail, there are likely to be links between the Commercialisation (technology transfer) function and other functions, highlighting the importance of the process of developing the ToC and articulating these key links in order to understand the potential ways in which HEIF generates value. Links are considered evident in relation to both ToCs developed in detail, for example specialist advice on spin-out creation within the Commercialisation (technology transfer) function may contribute to the delivery of new enterprises started-up within the Skills and human capital development function that has been generated via enterprise and entrepreneurship training.

- Second, and related to this, there will be a need for clarity between the 'Commercialisation (technology transfer)' function and the 'Business Development' function in particular. Previous depictions of the function coverage[25], and the engagement with practitioners, indicate that there are often cross-overs between the activities, notably around advice to academics regarding the commercial exploitation of research. For example, practitioners engaged in this work indicated that IP advice and Contracts/legal support should be included in the logic model and ToC for the 'Business Development' function, as this is often delivered as part of the business development offer in practice. Ensuring clarity as far as is practical on what activity-types are covered in each ToC respectively, but also recognising the links between the routes to outcome to test in the Contribution Analysis, will be important.

- Third, similar issues in relation to the internal and external factors driving performance are anticipated to be evident for this function, which the proposed Contribution Analysis would seek to understand and account for in the analysis. For example, a particular external factor in relation to investment funds will be the spatial context of institutions with very different levels of other sources of external finance for commercialisation evident across the country; this factor may influence the role of HEIF and how it generates value, which would be considered in the Contribution Analysis drawing on the evidence from primary research e.g. the role/contribution may be found to be particularly important where there is limited other funding available. Similarly, the way in which institutions deliver commercialisation support and investment funds varies (for example, through out-sourcing of activity, and partnership approaches), which would need to be considered in the development of the theory of ToC.

[25] For example, Understanding the Knowledge Exchange Infrastructure in the English Higher Education Sector, A report to HEFCE by PACEC and the Centre for Business Research, University of Cambridge, 2011

Approach to synthesis and assessment

5.42 As discussed above, function-level research using Contribution Analysis could be completed individually and Research England may wish to prioritise the research in one or a number of functions of most interest. As a qualitative approach, generating a series of contribution stories answering the research questions could be used to build-up evidence sequentially to generate an overarching narrative for how HEIF generates value across functions. This would draw on the evidence from the function-level analysis and seek to identify areas of consistent findings and/or where the plausibility of the contribution has been evidenced.

5.43 This said, the proposed method does include primary research with directors/senior manager of institutions and partners/stakeholders, and desk-based qualitative analysis, that would span the functions. Delivering several function-level analyses over a similar period would help to minimise the burden on consultees and maximise efficiencies in research tasks/costs.

Risks

5.44 **The approaches set out in the proposed framework are novel in a KE context. Their implementation therefore does pose a risk to Research England, and this is recognised explicitly and transparently by the study team**.

5.45 Considering the core approach specifically, key risks include:

- the variation in implementation of HEIF across institutions, even within individual functions, limits the robustness of the findings and the ability to develop plausible 'contribution stories'
- it is not possible to develop coherent ToCs for every function, limiting the extent to which these can be tested and form the basis for the contribution analysis e.g. owing to the issue of definition across functions, and the assumptions and theory underpinning relationships between activities and outputs/outcomes across functions
- the required level of input from institutions – in the development of the ToCs and in primary research – is considered disproportionate, limiting engagement and consequently the quality and coverage of the evidence base.

5.46 Two (not necessarily mutually exclusive) options appear to be evident to seek to manage these risks:

- Research England could 'pilot' the Contribution Analysis approach with one function, to test its deliverability, robustness, and the extent to which it does provide the nature of evidence that is sought for the theory-based approaches. If successful, the approach could then be 'rolled-out' as part of the next full evaluation of HEIF, including drawing on any lessons learned. If the pilot was not successful – for theoretical or practical reasons – this could inform either a revision to the method's research approach or focus (e.g. focusing research at the level of individual outcome pathways, or for a sub-set of institutions with very similar delivery models to limit the level of variance in implementation and context, rather than taking a function-by-function approach), or

demonstrate that such approaches are not in practice viable for the programme and should not be progressed in future.

- Research England could apply the 'principles' of Contribution Analysis as part of the next overall evaluation of HEIF, informing the approach taken to primary research and analysis, but not seeking to deliver the formal 'six step' approach set out above that would align with the requirements of the methodology in full. This would involve, for example: (i) developing a high-level ToC for HEIF that includes broad activities, outputs, and outcome descriptions and the key assumptions, risks and factors influencing its realisation (ii) including in any primary research with institutions and partners/stakeholders specific questions regarding the other internal and external factors influencing the delivery of KE outputs/outcomes, and their importance relative to HEIF and (iii) analysing this primary evidence (and any available supporting evidence) systematically against the ToC. This would provide a less 'plausible' conclusion in relation to role and relative contribution of HEIF than the full formal approach, however would complement the quantitative analysis.

5.47 Similar risks are associated with the potential supplementary research:

- The Realist Evaluation in particular would represent an exploratory approach in a KE context. In this case, a pilot of the approach on a single outcome may also be appropriate.

- Risks associated with the complementary qualitative research on the connecting and translational' role of KE practitioners is less pronounced. This does not seek to apply a formal theory-based approach and is more consistent with standard qualitative research approaches. This said, the risk around the level of engagement required with institutions is also evident, and would need to be managed carefully, alongside the wider proposed research approaches.

6. Conclusions and recommendations

Conclusions

6.1 Below we set out the main conclusions of the review, including the theory-based methodologies identified as most appropriate to a future evaluation of HEIF, and the way(s) in which they could potentially be applied to such an evaluation.

6.2 The over-arching objective of this study was to explore the potential (novel) theory-based evaluation methodologies that could be applied as part of the next full-scale evaluation of HEIF, complementing quantitative evaluation approaches. The focus was on identifying methodologies that can be used to better explain and test the causal relationships between HEIF-supported KE activities and their outputs and outcomes, focusing on the specific 'functions' of KE defined and developed on behalf of Research England, and subsequently applied in research and programme management of HEIF.

6.3 The methodologies would be focused on explaining 'how' outputs and outcomes are realised (i.e. providing accountability for the use of HEIF funding). Given the complexity of HEIF, they would be methodologies capable of producing valid and reliable findings when used to evaluate a complex funding programme, supporting a wide variety of different activities in a range of settings. Given the existing data collection associated with HEIF and its flexible approach, these methodologies should also seek to minimise as far as practical the inputs from institutions supported by HEIF, although it as recognised that some primary research would likely be required in all cases.

6.4 A review of academic literature and wider evidence found that the application of theory-based methodologies to the evaluation of KE would be novel. We found no clear examples from elsewhere of such an evaluation being conducted, and therefore identified no methodologies that have been conclusively demonstrated to work when applied to a programme like HEIF.

6.5 We did, however, identify a range of methodologies with the (exploratory) potential to be applied to KE. Each methodology identified had its strengths and weaknesses but based on the criteria set out above two were deemed most appropriate potentially to meet the requirements for the evaluation of HEIF: Contribution Analysis and Realist Evaluation.

6.6 Both Contribution Analysis and Realist Evaluation rely on the development of a ToC at the outset of the evaluation process. This ToC then serves as a hypothesis for the subsequent evaluation research to test. We focused on two KE functions as 'exemplars' and, with input from KE practitioners, developed a ToC for each in order to test the potential applicability and deliverability of the theory-based methods.

6.7 The process of developing the ToCs highlighted that the pathways to outcomes for HEIF are complex, although focusing specifically on those outputs and outcomes that are captured in the existing HE-BCI survey means that theory-based methods may be proportionate. The process of developing the ToCs also indicated that there will be important linkages and inter-dependencies between different KE functions, including the role of KE practitioners. Identifying and understanding these will be important for the subsequent evaluation.

6.8 The desk-review and findings from the development of the exemplar ToCs informed a conceptual framework, that combines:

- a 'core approach' focused on applying Contribution Analysis at the level of the functions, and focused specifically on the outputs and outcomes that are captured in the existing HE-BCI survey;
- 'potential supplementary approaches' utilising Realist Evaluation, considering individual outcomes in more detail; and
- complementary research on the 'connecting and translational' role of KE practitioners (identified as particularly important during this study) to provide a fuller picture of the value of HEIF.

6.9 Practical methods to implement the approaches have been set out in the main report body. For the 'core approach' these methods would support the initial development of logic models and ToCs for each KE function, then the subsequent implementation of the 'six step' approach of Contribution Analysis. Evidence would be gathered to inform a 'contribution story', generating plausible evidence on the role that HEIF played in bringing about the KE outputs/outcomes in each function (as captured in the HE-BCI survey), and how important HEIF was (relative to other factors) in explaining how and why outputs/outcomes were realised.

6.10 It is highlighted that the approaches set out in the proposed framework are novel in a KE context. Their implementation therefore does pose a risk to Research England, and this is recognised explicitly and transparently by the study team. This should inform the next steps in taking forward the findings of the work.

Recommendations

6.11 In this context, below we set out a series of recommendations based on the findings and exemplar approaches presented in this report. However, it is recognised that there may be changes to the coverage of HE-BCI (with a review by HESA currently underway[26]), and we understand that Research England is currently reviewing the coverage of the KE functions. Both developments will impact on the proposed approach to the function-level analysis. It is also noted that the specific timing of the next evaluation of HEIF is not yet confirmed. This noted, the following broad recommendations are made:

- **Research England should revisit the exemplar logic models/ToCs once the implications of any changes to HE-BCI and KE functions are known**. The purpose would be to check that they accurately reflect any updated function definitions and ensure that data on outputs/outcomes will be available from HE-BCI to inform the function-level analysis. Logic models and ToCs for the remaining KE functions could also be developed at this point, but we recommend this should only be done once the HE-BCI coverage and KE functions are confirmed.
- The novelty of applying the identified methodologies to the evaluation of HEIF is a risk that should be recognised. **We therefore recommend that Research England**

[26] Initial outputs expected in 2020: https://www.hesa.ac.uk/innovation/records/reviews/he-bci-major-review

pilot the use of Contribution Analysis with one of the finalised KE functions initially. The timing of this is dependent on the wider evaluation activity, however this could be undertaken in advance of the full research phase.

- If this is successful, Contribution Analysis could then be rolled-out for the full evaluation (covering all functions).
- If this is not successful (e.g. because the contribution story does not provide plausible evidence, or because its practical implementation is regarded as overly burdensome by partners and Research England), the principles of Contribution Analysis could be applied to qualitative research for the full evaluation.

- **Research England should consider including questions in their templates and monitoring systems for the next phase to collect data that would support the qualitative evaluation**. This would help to minimise the coverage of data to be collected via primary research (e.g. surveys). This could include consistent information collection on activities delivered against the function logic models.
- The implementation of the 'potential supplementary approaches' should be considered by Research England when the timing of the full evaluation of HEIF is confirmed. Again, **we recommend this approach be piloted for the next evaluation to provide additional learning and insight**, demonstrating the value and role of HEIF in delivering wider knowledge exchange outcomes to help inform policy and decision-making.
- More broadly, **Research England and institutions in receipt of HEIF may wish to consider whether methodologies identified in this research that support development/learning could be applied in the continuous improvement of HEIF**. Particular methodologies of interest are Outcome Harvesting and Most Significant Change.

Annex A: Glossary of terms and concepts

A.1 Definitions of key terms and concepts used within the main body of the report are set out below.

- **Case study research**: research that seeks to test a research hypothesis and/or examine best practice in delivery of an intervention by producing detailed case studies of instances of the intervention being delivered or of the experiences of individual participant individuals/organisations.
- **Econometrics:** the branch of economics concerned with using mathematical techniques (particularly statistics) to examine economic systems.
- **Grey literature:** materials produced outside of the usual commercial and academic publishing and distribution channels. Examples of grey literature publication types include: reports; work papers; government documents; and evaluations.
- **Hypothesis (plural: hypotheses):** a proposed explanation to be tested during subsequent research.
- **Logic model:** a visual depiction of the relationship between resources put into a project and that project's ultimate results. Individual elements of a logic model are defined below:
  - **Inputs**: the resources used to deliver a project (incl. financial and human)
  - **Activities:** the activities undertaken in support of a project's objectives.
  - **Outputs:** the immediate products of the activities undertaken (e.g. the number of training sessions delivered or participants receiving support).
  - **Outcomes and impacts:** the medium and longer-term knock-on effects of the outputs (e.g. increased skills or knowledge due to participating in a training session).
- **Mixed-methods study**: a research study conducted using a mixture of different research methods, including those gathering both qualitative and quantitative data.
- **Qualitative data:** information about qualities, that cannot be recorded in numeric form (e.g. an individual's employment status). Also known as categorical data.
  - Although qualitative data is not numeric, it is possible to categorise qualitative data and by doing so convert it into numerical (quantitative) data (e.g. counting the number of project participants who are university students).
- **Quantitative data:** information that can be measured and written down in numeric form.
- **Theory of Change:** a visual depiction of the relationship between resources put into a project or programme and that project's/programme's ultimate results. Differs from

a logic model in that a Theory of Change maps the relationships between each individual element of project/programme delivery, showing the actual or intended pathway by which individual inputs lead to specific activities, outputs and outcomes, whereas a logic model does not.

- **Theory-based evaluation:** an approach to evaluation that examines both the outcomes realised by a programme and how the programme's activities led to those outcomes.

Annex B: Literature review findings

Search Strategy

B.1 The literature review used systematic searches which combined a general search strategy that was developed after testing a combination of keywords on the following bibliographic databases: Econlit (EBSCO) and Web of Science (Clarivate Analytics). Bespoke searches were also undertaken for promising methodologies. The databases searched have relatively poor indexing therefore requiring broader search terms than planned to increase sensitivity which increased the volume of returns that had to be screened. However, for manageability we limited the number of databases searched. However, even though we did not search Medline, the proportion of returns from the health field was high.

1. University OR Higher Education OR HEI
2. Research OR technology OR innovation
3. Knowledge Mobilisation OR Knowledge Exchange OR Diffusion OR commercialisation
4. Evaluation OR Monitoring OR impact assessment OR logic model
5. #1 and #2
6. #5 and #3
7. #6 and #4

B.2 Search results were stored in Endnote and duplicates removed.

Screening

B.3 Two reviewers screened abstracts and titles to produce a long list of texts for potential inclusion, which was reviewed by a third reviewer (an evaluation expert with understanding of Knowledge Exchange). Full texts of promising texts were then retrieved and considered for inclusion.

Included

- Qualitative methodology
- Application in knowledge exchange or analogue (similar) activity

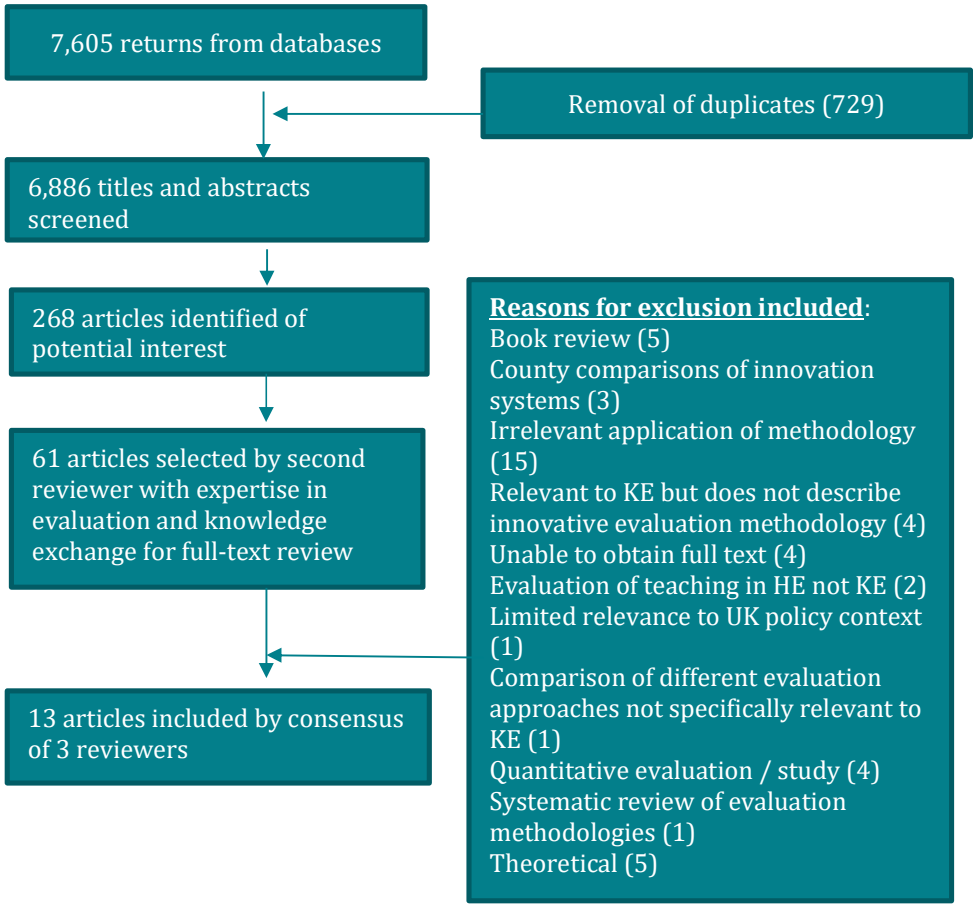
Excluded

- Quantitative methodology
- Theoretical outline of methodology

Search Results

B.4 Search results are summarised in Figure B-1. In all 13 of the 61 shortlisted articles were chosen for full review[27] and individually summarised under in the 'Findings' sub-section of this annex (page A-7 onwards).

Figure B-1: Overview of search results



[27] See Table **B-1** for a list of those excluded from the shortlist of 61 and reasons for their exclusion.

Table B-1: Excluded texts

Reference	Reason for Exclusion
Abma, T. A. (2019). "Collaborative, Participatory, and Empowerment Evaluation. Stakeholder Involvement Approaches." Evaluation and Program Planning 74: 18-19.	Book review
Alvarez-Coque, J. M. G., F. Mas-Verdu and N. Roig-Tierno (2017). "Technological innovation versus non-technological innovation: different conditions in different regional contexts?" Quality & Quantity 51(5): 1955-1967.	Country comparison of innovation systems
Aus, J. P. (2009). "Conjunctural causation in comparative case-oriented research." Quality & Quantity 43(2): 173-183.	Irrelevant example of application of methodology.
Befani, B. and J. Mayne (2014). "Process Tracing and Contribution Analysis: A Combined Approach to Generative Causal Inference for Impact Evaluation." Ids Bulletin-Institute of Development Studies 45(6): 17-36.	Unable to obtain full text
Bennett, A. and J. T. Checkel (2015). Process Tracing From Metaphor to Analytic Tool Preface.	Unable to obtain full text
Berbegal-Mirabent, J., D. E. Ribeiro-Soriano and J. L. S. Garcia (2015). "Can a magic recipe foster university spin-off creation?" Journal of Business Research 68(11): 2272-2278.	Relevant to KE topic but does not describe innovative evaluation methodology
Biggs, J. S., L. Farrell, G. Lawrence and J. K. Johnson (2014). "A practical example of Contribution Analysis to a public health intervention." Evaluation 20(2): 214-229.	Irrelevant example of application of methodology. (contribution analysis)
Blackman, T., J. Wistow and D. Byrne (2013). "Using Qualitative Comparative Analysis to understand complex policy problems." Evaluation 19(2): 126-140.	Irrelevant example of application of methodology (QCA).
Brandon, P. R. (2014). "Book Review: Participatory Evaluation Up Close: An Integration of Research-based Knowledge." American Journal of Evaluation 35(2): 291-294.	Book review
Cambre, B., P. C. Fiss and A. Marx (2013). CONCLUSION: THE PATH FORWARD. Configurational Theory and Methods in Organizational Research. P. C. Fiss, B. Cambre and A. Marx. 38: 311-319.	Unable to obtain full text
Caren, N. and A. Panofsky (2005). "TQCA - A technique for adding temporality to qualitative comparative analysis." Sociological Methods & Research 34(2): 147-172.	Theoretical elaboration of existing methodology (QCA)
Copestake, J. (2014). "Credible impact evaluation in complex contexts: Confirmatory and exploratory approaches." Evaluation 20(4): 412-427.	Irrelevant example of application of methodology (QUIP).
Cousins, J. B. and J. A. Chouinard (2012). Participatory Evaluation Up Close: An Integration of Research-Based Knowledge.	Book review
Coxe, A. M. and C. A. Reiter (2003). "Fuzzy hexagonal automata and snowflakes." Computers & Graphics-Uk 27(3): 447-454.	Not relevant
Crespo, N. F., R. Rodrigues, A. Samagaio and G. M. Silva (2019). "The adoption of management control systems by start-ups: Internal factors and context as determinants." Journal of Business Research 101: 875-884.	Irrelevant example of application of methodology (QCA).
Delahais, T. and J. Toulemonde (2012). "Applying contribution analysis: Lessons from five years of practice." Evaluation 18(3): 281-293.	Irrelevant example of application of methodology (contribution analysis)

Reference	Reason for Exclusion
Denk, T. (2010). "Comparative multilevel analysis: proposal for a methodology." International Journal of Social Research Methodology 13(1): 29-39.	Irrelevant example of application of methodology.
	Irrelevant example of application of methodology (QCA).
Denk, T. and S. Lehtinen (2014). "Contextual analyses with QCA-methods." Quality & Quantity 48(6): 3475-3487.	Country comparison of innovation systems
Essegbey, G. O., O. Sakyi-Dawson, D. Kossou, B. Ouologuem, F. Dembele, R. Adu-Acheampong and J. Jiggins (2017). "External influences on agro-enterprise innovation platforms in Benin, Ghana and Mali - Options for effective responses." Cahiers Agricultures 26(4).	Irrelevant example of application of methodology (Bayesian updating)
Fallesen, P. and R. Breen (2016). "Temporary Life Changes and the Timing of Divorce." Demography 53(5): 1377-1398.	Unable to obtain full text
Ferrer-Martin, R. M., M. R. Sepulveda, F. J. Reyes-Zurita, M. Medina-O'Donnell, A. Perez-Jimenez, C. E. Trenzado and E. E. Rufino-Palomares (2014). NOVEL TOOLS FOR CONTINUOUS EVALUATION IN UNIVERSITARY STUDIES. Edulearn14: 6th International Conference on Education and New Learning Technologies. L. G. Chova, A. L. Martinez and I. C. Torres. Valenica, Iated-Int Assoc Technology Education a& Development: 1207-1210.	Book review
Hargreaves, M. B. and D. Podems (2012). "Advancing Systems Thinking in Evaluation: A Review of Four Publications." American Journal of Evaluation 33(3): 462-470.	Relevant to KE topic but does not describe innovative evaluation methodology
Kitagawa, F., M. S. Barrioluengo and E. Uyarra (2016). "Third mission as institutional strategies: Between isomorphic forces and heterogeneous pathways." Science and Public Policy 43(6): 736-750.	Irrelevant example of application of methodology (Developmental evaluation)
Lam, C. Y. and L. M. Shulha (2015). "Insights on Using Developmental Evaluation for Innovating: A Case Study on the Cocreation of an Innovative Program." American Journal of Evaluation 36(3): 358-374	Evaluation of innovation in teaching in HE not KE
Lawrenz, F. and D. Huffman (2003). "How can multi-site evaluations be participatory?" American Journal of Evaluation 24(4): 471-482.	Limited relevance to UK
Laycock, A. F., J. Bailie, N. A. Percival, V. Matthews, F. C. Cunningham, G. Harvey, K. Copieys, L. Pater and R. Bailie (2019). "Wide-Scale Continuous Quality Improvement: A Study of Stakeholders' Use of Quality of Care Reports at Various System Levels, and Factors Mediating Use." Frontiers in Public Health 6.	Quantitative evaluation/study
Lee, M. T., R. L. Raschke and R. St Louis (2016). "Exploiting organizational culture: Configurations for value through knowledge worker's motivation." Journal of Business Research 69(11): 5442-5447.	Theoretical
Leeuw, F. L. (2012). "Linking theory-based evaluation and contribution analysis: Three problems and a few solutions." Evaluation 18(3): 348-363.	Irrelevant example of application of methodology (contribution analysis)
Lemire, S. T., S. B. Nielsen and L. Dybdal (2012). "Making contribution analysis work: A practical framework for handling influencing factors and alternative explanations." Evaluation 18(3): 294-309	Evaluation of innovation in teaching in HE not KE
Leonard, S. N., R. N. Fitzgerald and G. Riordan (2016). "Using developmental evaluation as a design thinking tool for curriculum	Relevant to KE topic but does not describe

Reference	Reason for Exclusion
innovation in professional higher education." Higher Education Research & Development 35(2): 309-321.	innovative evaluation methodology
Lewis, D. G. R., E. M. Gerber, S. E. Carlson and M. W. Easterday (2019). "Opportunities for educational innovations in authentic project-based learning: understanding instructor perceived challenges to design for adoption." Etr&D-Educational Technology Research and Development 67(4): 953-982	Comparison of different evaluation approaches, not specifically relevant to KE
Luskin, R. J. C. and T. Ho (2013). "Comparing the intended consequences of three theories of evaluation." Evaluation and Program Planning 38: 61-66.	Quantitative evaluation
Maag, S., T. J. Alexander, R. Kase and S. Hoffmann (2018). "Indicators for measuring the contributions of individual knowledge brokers." Environmental Science & Policy 89: 1-9.	Theoretical
Mahoney, J. (2012). "The Logic of Process Tracing Tests in the Social Sciences." Sociological Methods & Research 41(4): 570-597	Irrelevant example of application of methodology (realist).
Marchal, B., M. Dedzo and G. Kegels (2010). "A realist evaluation of the management of a well-performing regional hospital in Ghana." Bmc Health Services Research 10	Theoretical
Mark, M. M. and G. T. Henry (2013). "Logic models and content analyses for the explication of evaluation theories: The case of emergent realist evaluation." Evaluation and Program Planning 38: 74-76.	Theoretical
Miller, R. L. (2010). "Developing Standards for Empirical Examinations of Evaluation Theory." American Journal of Evaluation 31(3): 390-399	Systematic review of evaluation methodologies
Minary, L., J. Trompette, J. Kivits, L. Cambon, C. Tarquinio and F. Alla (2019). "Which design to evaluate complex interventions? Toward a methodological framework through a systematic review." Bmc Medical Research Methodology 19: 9.	Irrelevant example of application of methodology (QCA).
Opstrup, N. (2017). "When and why do university managers use publication incentive payments?" Journal of Higher Education Policy and Management 39(5): 524-539	Quantitative evaluation (training in knowledge translation)
Park, J. S., J. E. Moore, R. Sayal, B. J. Holmes, G. Scarrow, I. D. Graham, L. Jeffs, C. Timmings, S. Rashid, A. M. Johnson and S. E. Straus (2018). "Evaluation of the "Foundations in Knowledge Translation" training initiative: preparing end users to practice KT." Implementation Science 13: 13.	Book review
Patton, M. Q. (2015). "Empowerment Evaluation: Knowledge and Tools for Self-Assessment, Evaluation Capacity Building, and Accountability." Evaluation and Program Planning 52: 15-18.	Irrelevant example of application of methodology (QCA).
Pattyn, V., A. Molenveld and B. Befani (2019). "Qualitative Comparative Analysis as an Evaluation Tool: Lessons From an Application in Development Cooperation." American Journal of Evaluation 40(1): 55-74.	Country comparison of innovation systems
Proksch, D., M. M. Haberstroh and A. Pinkwart (2017). "Increasing the national innovative capacity: Identifying the pathways to success using a comparative method." Technological Forecasting and Social Change 116: 256-270	Irrelevant example of application of methodology (realist).
Rushmer, R. K., D. J. Hunter and A. Steven (2014). "Using interactive workshops to prompt knowledge exchange: a realist evaluation of a knowledge to action initiative." Public Health 128(6): 552-560.	Does not describe innovative methodology

Reference	Reason for Exclusion
Saunders, A. M. and S. Sin (2015). "Middle manager's experience of policy implementation and mediation in the context Scottish quality enhancement framework." Assessment and Evaluation in Higher Education 40: 135-150.	Quantitative evaluation
Upton, S., P. Vallance and J. Goddard (2014). "From outcomes to process: evidence for a new approach to research impact assessment." Research Evaluation 23(4): 352-365.	Relevant to KE topic but does not describe innovative evaluation methodology

Findings

Text 1: Abboud (2016)

Outcome Harvesting	Abboud, R. and C. Claussen (2016). "The use of Outcome Harvesting in learning-oriented and collaborative inquiry approaches to evaluation: An example from Calgary, Alberta." Evaluation and Program Planning 59: 47-54.
Published Abstract The Community Development Learning Initiative (CDLI) in Calgary, Alberta, Canada aims to be a network that brings together neighbourhood residents, community development practitioners and other supporters to learn and act on neighbourhood-based, citizen-led community development projects. In 2013, the CDLI initiated The Evaluation for Learning and Dialogue Project to provide the opportunity for organizations and supporters to work together to establish a shared vision and goals through discussions about evaluation learning and outcomes. It was intended that the project would be a useful learning tool for participating organizations by enabling them to engage in an evaluative methodological process, and record relevant information and to compare and learn from each other's projects. Outcome Harvesting was chosen as the evaluation methodology for the project. This article reviews critical learning from the project on the use of Outcome Harvesting methodology in the evaluation learning and outcomes of local community development projects, and it provides lessons for other jurisdictions interested in implementing this methodology	**Features** Case (study) based: ◉ Configurational: Counterfactual: Experimental: Generative causation: Mapping: Modelling: Participatory: ◉ Predictive: Quasi-experimental: Statistical association: Synthesis: Theory based:

Overview of approach
 	
• Outcome Harvesting is an evaluation approach in which evaluators, grant makers, and/or programme managers and staff identify, formulate, verify, analyse and interpret 'outcomes' in programming contexts where relations of cause and effect are not fully understood"
• It identifies, monitors and learns from the changes by asking: What changed, for whom, when, where and why does the change matter?
• It is a highly participatory approach that identifies, formulates, verifies and makes sense of outcomes and activities when relationships of cause and effect are unknown.
• Unlike many evaluation approaches that begin with stated outcomes or objectives, this approach looks for evidence of outcomes (through reports, personal interviews and other
	Overview of application
	
	• Outcome harvesting (OH) was developed for use in international development and this paper describes an alternative application in evaluating community development in Canada.
	• Four different approaches were considered by Community Development Learning Initiative (CDLI) in Calgary, Alberta, as part of its Evaluation for Learning and Dialogue Project. The four approaches were:
	  • Outcome mapping
	  • Outcome harvesting
	  • Most significant change

documented information) and explanations for those outcomes, in what has already happened through a process the creators call 'sleuthing.' • Outcomes are defined specifically as "observable changes in the behaviour, relationships, activities and actions of 'boundary partners' i.e. those individuals, groups or organizations with whom a project or program interacts and works with directly to affect change • The method consists of six iterative steps that help uncover the changes in behaviours and actions, and it actively engages a variety of important stakeholders in the process – the change agent (the individual or organization that influences an outcome), the social actor (the individual, group, community, etc. that changes as a result of the social actor's activities), the harvest user (the individual(s) who will use the information gleaned from the OH for a variety of reasons and the harvester (the person responsible for the OH process)	• Reflective practice process. • The emphasis was on learning and not accountability and evaluation should enhance relevance, ownership or deep involvement on the side of practitioners in the evaluation process. • OH was chosen for two reasons. First, the ability to deal with complex and non-linear initiatives where the outcomes may be unknown in advance, to elucidate changes and on the utilization of the findings meant it would be a responsive tool to the diverse landscape of CD initiatives in Calgary. A tool useful to a variety of organizations to uncover the sometimes-intangible outcomes of their CD initiatives that would be difficult to articulate in traditional evaluation frameworks. Second, create a shared 'space' where a variety of people doing CD work could talk about it in a way that promoted shared learning and dialogue. • The promise of OH is to retrospectively uncover the context, behaviours and other key influencers in program or initiative implementation was enticing because of the desire of the CDLI to promote deep and insightful dialogues so that people could learn from each other's work without focusing specifically on outcomes. Therefore, the learning that is possible from OH, which focuses on the whom, when, where and why it matters as much as the what, would be the basis of these conversations and would help to move them beyond merely comparing and contrasting activities**.**
Assessment of novelty Still novel outside the field of international development.	**Methodological issues addressed by methodology** Multiple alternative pathways to outcomes. Where outcomes are unknown in advance and difficult to articulate. Traditional' evaluation approaches – positivist methodologies that favour linear and logical forms of project implementation and evaluation were considered not to offer theoretical or practical platforms for community development (CD) evaluations. The same might be said of commercialisation of technologies developed in universities that have multiple pathways (especially health technologies). However, the similarity stops there as the focus on was supporting learning and information sharing rather than accountability in community development.
Strengths **Authors:** Given that the purpose of the project was to 'test' if OH could be used across a variety of CD contexts, the evaluators were purposeful in facilitating conversation with the pilot site members as to the process of using OH in their projects. To varying degrees, the four sites were impressed with the level of outcome data, as well as how a more nuanced understanding of their work emerged throughout the process. Some discussed the heaviness of the workload it required (surprisingly, by the largest team; the smaller	**Weaknesses** **Authors:** There are a range of outcome descriptions that each contains varying aspects of a required outcome description. In some cases, there is a lack of specific information, such as dates when the change in behaviour, relationships, activities or actions occurred (i.e., Out-comes #1–3). In other instances, outcome descriptions failed to include how they, as the change agents, contributed to the change (i.e., Outcomes #1 through to #5). Only one of the six outcomes generated could be

initiatives did not necessarily speak about how cumbersome it was) and how the process, particularly how to develop and elucidate the outcome statements, was challenging and confusing. Overall, the sites were impressed with the knowledge that emerged about their initiatives. Outcome Harvesting supports deep learning The theoretical and epistemological foundations of this type of participatory, iterative and emergent methodology provided a mechanism by which participants could support the short- and long-term claims they had made about their work. The types of outcome statements that emerged moved beyond single, linear causal paths, which although they may be appropriate in some settings, is questionably appropriate in interventions such as CD because of the risk of overstating the causal contributions of the intervention. Another way that OH supported deep learning in the pilot projects was through the use of "useable/useful questions" that were created by the Project Team to help guide the inquiry. The value of OH in the monitoring of project implementation was evident during the pilots. The agency that conducted the evaluation in mid-cycle reported that the OH process highlighted opportunities for continuous improvement of the program design, which elicited some course correction in mid-cycle. The value of the evaluation process was made tenable for this agency; the real- time data provided a feedback loop in time for program improvement as they continued in the cycle. Moreover, they discussed that since this was a newer project, implemented via a theory-based change model, there were some uncertainties during the planning and implementation phases. Completing the OH process helped to understand and manage these uncertainties better. OH does not reduce these sorts of uncertainties, but allows for an approach that is formative in nature and produces knowledge about the program that is explicitly contingent on the time and the place of the current cycle and could inform future implementation and cycles.	considered fulsome, given that it includes information on whom the change happened for (with quantitative data to indicate the #'s of participants), when the change happened, and how the change agent contributed to that change. It was also slightly more sophisticated, in that it developed an explanation of the significance of the outcome. This occurred as a result of engagement with and feedback from members of the pilot sites team who felt a 'long' version of the outcome statement would be better able to answer the useable question. First, Outcome Harvesting reverses the logic of traditional monitoring and evaluation (Wilson-Grau & Britt, 2012). Identifying outcomes, whether planned or not, and explaining how the pilot sites contributed, was a significant shift in thinking for all pilot site participants. This was further complicated by the way in which pilot project documented data and observed CD project processes. It was a sobering realization for several pilots that not only was there little data to review, but their notes and emails were focused much more on trouble-shooting issues arising during project implementation than on outcomes for participants. Perhaps not surprising, the pilot site with the most complete and fulsome outcome description. During the substantiation phase, a beneficiary stated that the crafted outcome statements were "too simple" and that a lot more substantial personal and professional change had occurred during, but more significantly, after the project cycle. This feedback forced the pilot site participant to re-think and re-craft the outcome statement so that it captured the extent of the change. Fortunately, OH provides a framework for creating significant outcome statements that are specific, measurable, achievable, relevant and timely.
Resource intensiveness of approach Even where possible data collection was built around existing activity it did create significant requirements on behalf of participating pilots.	**Robustness, Reliability and Replicability** The advantage of outcome harvesting is that can be used in situations where it is difficult to develop logic model or implement monitoring frameworks because of heterogeneous activities. The downside is there is trade off in terms of robustness, reliability and replicability.

Text 2: Archibald (2018).

Realist evaluation	**Archibald, M. M., et al. (2018).** Transdisciplinary research for impact: protocol for a realist evaluation of the relationship between transdisciplinary research collaboration and knowledge translation. BMJ Open 8(4): 7
Published Abstract **Introduction** Transdisciplinary teams are increasingly regarded as integral to conducting effective research. Similarly, knowledge translation is often seen as a solution to improving the relevance and benefits of health research. Yet, whether, how, for whom and under which circumstances transdisciplinary research influences knowledge translation is under theorised, which limits its potential impact. The proposed research aims to identify the contexts and mechanisms by which transdisciplinary research contributes to developing shared understandings and behaviours of knowledge translation between team members. **Methods and analysis** Using a longitudinal case-study design approach to realist evaluation, we outline a study protocol examining whether, how, if and for whom transdisciplinary collaboration can impact knowledge translation understandings and behaviours within a 5-year transdisciplinary Centre of Research Excellence. Data are being collected between February 2017 and December 2020 over four rounds of theory development, refinement and testing using interviews, observation, document review and visual elicitation as data sources. **Ethics and dissemination** The Health Research Ethics Committee of the University of Adelaide approved this study. Findings will be communicated with team members at scheduled intervals throughout the study verbally and by means of creative reflective approaches (eg, arts elicitation, journaling). This research will be used to help support optimal team functioning by identifying strategies to support knowledge sharing and communication within and beyond the team to facilitate attainment of research objectives. Academic dissemination will occur through publication and presentations.	**Features** Case (study) based ◉ Configurational[28] ◉ Counterfactual Experimental Generative causation[29] ◉ Mapping Modelling Participatory Predictive Quasi-experimental Statistical association Synthesis Theory based ◉

[28] Configurational view of causation involves a configuration / constellation of conditions, or more precisely with a combination of single causes producing the same effect. Different combinations may lead to the same outcome; and similar combinations may lead to different outcomes, because individual conditions can affect the outcome in opposite ways, depending on what other factors they are combined with.

[29] Generative causation depends on identifying the 'mechanisms' that explain effects – the inference basis for 'theory based' and 'realist' approaches to impact evaluation. Generative causation is strong on explanation but weak on estimating quantities or extent of impact.

Overview of approach	Overview of application

• Realist evaluation is a type of theory-driven evaluation method used to understand if, how, for whom and under what circumstances an intervention 'works' to produce an intended outcome
• Unlike other forms of theory-driven evaluation, realist approaches have a particular focus on understanding how causation works and why programme outcomes work or do not work in different contexts.
• In realist evaluations, researchers seek to uncover how various contexts (C) work with underlying mechanisms (M) to produce particular outcomes (O), which are theorised through possible CMO interactions or configurations.
• Such CMO configurations are explanatory pathways, underpinned by implicit theories that can be made explicit through the realist evaluation process.
• The philosophical premise of scientific realism distinguishes realist evaluation from other types of theory- driven evaluations. Here, reality is both knowable yet relative to the researcher, and actors possess innate capacity for change.
• Causal mechanisms are embedded in 'social relationships and contexts as much as individuals' which makes realist evaluation a highly appropriate approach to developing an explanation about the impact of transdisciplinary collaboration on knowledge translation within a team setting.
• The process of realist evaluation is iterative, cycling between (1) theory development (ie, generating a working theory/hypothesis), (2) theory verification (ie, hypothesis/theory testing throughout data collection), and (3) theory refinement (ie, refining the hypothesis/ theory based on emerging data).
• A middle-range theory is generated, which lies between the working hypothesis and a fully operational, explanatory theory.
• Data collection is pragmatic and method neutral—selection of data sources and methods is guided by what is needed to test the working hypothesis.
	• Transdisciplinary research and knowledge translation are increasingly more common in health research, but the relationship between the two concept is poorly defined. There is the need for a stronger theoretical explanations of if, how, why, for whom and under what circumstances transdisciplinary collaboration influences knowledge translation and this can be generated through realist evaluation (RE).
	• RE is seen by the authors as a way of maximising the benefits of collaboration in health research who see it as being contingent on:
	  ○ Upon shared understandings (eg, the nature of the research problem, roles of team members, team objectives and translational goals).
	  ○ How diverse team members understand and enact the concept and process of knowledge translation (ie, the interactive and iterative process of knowledge creation, sharing and use for better health outcomes, and involving multiple system stakeholders) can produce barriers to knowledge creation and knowledge sharing activities conducted within the team.
	  ○ Barriers having downstream effects on the uptake and impact of the knowledge produced.
	  ○ Tendency to regard academics as homogenous groups with shared goal and understandings of knowledge translation. Consequently, little is known about knowledge translation within academic communities, and very little research has explored processes of knowledge translation within transdisciplinary research teams
	  ○ Transdisciplinary research advancing knowledge translation in relation to complex, multifactorial health problems that often exceed the capacity of any single discipline.
	• The paper sets out a protocol for a RE that involves the development of an explanatory theory of how, for whom and under what circumstances transdisciplinary collaboration can impact knowledge translation is necessary to support such processes, and to identify which outcomes are affected by transdisciplinary team approaches in certain contexts.
	• Findings will inform implementation of responsive, context-driven strategies to maximise the impact of collaborative efforts, transdisciplinary approaches and research use.

Assessment of novelty RE is now well established in a number of evaluation fields. Now being used in the area of knowledge translation	**Methodological issues addressed by methodology** Experimental designs provide information on average effect of an intervention by comparing two or more groups (control / intervention; control/intervention/ alternative interventions) and in social policy interventions can suffer from lack of

RE originated in the area of criminal justice with Ray Pawson and Nick Tilly setting out the approach in their book in 1997. There has been increasing acceptance of the approach across a number of different policy areas including the redesign of health services, therefore, its use in the assessment of interdisciplinary research and knowledge translation in the health research is a logical development.	fidelity (not consistently executed). They do not provide information on the how, why and when intervention might work, which is the focus of RE.
Strengths **Authors:** This evaluation will be one of the first internationally to examine if, how, for whom and why transdisciplinary research collaboration impacts knowledge translation understandings and behaviours. This research will provide insight into understandings of knowledge translation within a transdisciplinary team, thereby identifying (developmentally) misaligned understandings of knowledge translation processes and activities, and concurrent strategies for supporting shared understandings.	**Weaknesses** **Authors:** Although realist evaluation can provide insights into which conditions impact which outcomes (and how), no single study cannot produce universally transferable findings. Despite its strengths, participant observation can limit both the depth of data provided to the researcher and the extent of confidentiality afforded to participating members given the researchers position.
Resource intensiveness of approach Moderate to high	**Robustness, Reliability and Replicability** Good for capturing learning and for understanding in what context specific interventions do or not work. The paper is a protocol for the implementation of a RE method, rather than the results. The success of its practical application is therefore no known at this point.

Text 3: Banks (2017).

Realist evaluation	**Banks, J., et al. (2017).** The researchers' role in knowledge translation: a realist evaluation of the development and implementation of diagnostic pathways for cancer in two United Kingdom localities. Health Research Policy and Systems 15: 11
Published Abstract **Background:** In examining an initiative to develop and implement new cancer diagnostic pathways in two English localities, this paper evaluates 'what works' and examines the role of researchers in facilitating knowledge translation amongst teams of local clinicians and policy-makers. **Methods:** Using realist evaluation with a mixed methods case study approach, we conducted documentary analysis of meeting minutes and pathway iterations to map pathway development. We interviewed 14 participants to identify the contexts, mechanisms and outcomes (CMOs) that led to successful pathway development and implementation. Interviews were analysed thematically and four CMO configurations were developed. **Results:** One site produced three fully implemented pathways, while the other produced two that were partly implemented. In explaining the differences, we found that a respected, independent, well-connected leader modelling partnership working and who facilitates a local, stable group that agree about the legitimacy of the data and project (context) can empower local teams to become sufficiently autonomous (mechanism) to develop and implement research-based pathways (outcome). Although both teams designed relevant, research-based cancer pathways, in the site where the pathways were successfully implemented the research team merely assisted, while, in the other, the research team drove the initiative. **Conclusion:** Based on our study findings, local stakeholders can apply local and research knowledge to develop and implement research-based pathways. However, success will depend on how academics empower local teams to create autonomy. Crucially, after re-packaging and translating research for local circumstances, identifying fertile environments with the right elements for implementation and developing collaborative relationships with local leaders, academics must step back.	**Features** Case (study) based: ◉ Configurational: ◉ Counterfactual: Experimental: Generative causation: ◉ Mapping: Modelling: Participatory: Predictive: Quasi-experimental: Statistical association: Synthesis: Theory based: ◉
Overview of approach See Archibald (2018) above.	**Overview of application** • Using a case study approach an RE of initiative to develop and implement new cancer diagnostic pathways in two English localities was undertaken in two main phases.

	• ***Phase1:*** Documentary analysis of the meeting minutes and decisions taken by the two reference groups was carried out alongside mapping of the formation and shape of the pathways to develop an understanding of the process and different outcomes from pathway development and implementation. This process addressed the following questions: (1) Did the sites develop new pathways for each of the specified cancers? (2) Were the pathways developed within the study time frame? (3) To what extent did the sites draw on the research of the Discovery Programme and/or other recent cancer diagnostic research? (4) Were the developed cancer diagnostic pathways implemented? • The documentary analysis generated accounts of the pathway development in each locality, which fed into a comparative matrix table including data on both pathway content and timing. The table also served to inform the development of draft context–mechanism–outcome (CMO) configurations, which, along with the programme theory, shaped the strategy for data collection in the second phase of the study. • ***Phase 2:*** qualitative interviews with purposefully sampled participants from each reference group (from clinicians (primary and secondary care), service managers and nonclinical stakeholders) at the two sites, provided insights into the context and mechanisms at play, for example, the rationale behind the decisions that were taken in relation to pathway development. The topic guide was developed from an assessment of the pathway development documentation described above and the programme theory that was outlined at the start of the evaluation. Formal consent to participate was taken prior to the interview. Interviews were audio recorded, transcribed verbatim and fully anonymised. Interview transcripts were analysed thematically. • Coded data were organised into CMO configurations (e.g. codes for 'working relationships' and 'leadership' were categorised as contextual factors) which underwent a. process of testing and refining the programme theory and developing CMOs, at regular research meetings as data analysis continued, until the final CMOs and revised programme theory were agreed.
Assessment of novelty See Archibald (2018) above.	**Methodological issues addressed by methodology** See Archibald (2018) above.
Strengths **Authors assessment:** A major strength of this research was that two sites were studied, such that potential CMOs could be contrasted and compared to create more robust theories about what worked. However, these CMOs need further testing elsewhere to provide greater confidence in the generalisability of findings.	**Weaknesses** **Authors assessment:** A key challenge of this study was the continual changes in healthcare policy and arrangements nationally and locally that required the pragmatic adaptation of the research team.

An unusual aspect was that the main interviewer was a policy-maker. With his experience of managing change in healthcare economies, he brought extra depth to the data collection, analysis and interpretation process. **Our assessment:** The advantage of the approach is that is scalable to include other sites should they become available.	**Our assessment:** the development of diagnostic pathways is inherently difficult to do and its evaluation of diagnostic tests has become a specialist area in health technology assessment.
Resource intensiveness of approach Relatively moderate in this study, e.g. number of interviews undertaken (14 out of 25 individuals approached took part) but does account of the time needed to develop theatrical underpinning and different CMO configurations.	**Robustness, Reliability and Replicability** The two-phase approach provides confidence in the robustness and reliability of using RE in the evaluation of knowledge translation. Approach is well described and could be replicated but requires a high level of interpretation by researchers of data.

Text 4: Baumgartner (2017).

Qualitative Comparative Analysis	Baumgartner, M. and A. Thiem (2017). Model Ambiguities in Configurational Comparative Research. Sociological Methods & Research 46(4): 954-987
Published Abstract For many years, sociologists, political scientists, and management scholars have readily relied on Qualitative Comparative Analysis (QCA) for the purpose of configurational causal modelling. However, this article reveals that a severe problem in the application of QCA has gone unnoticed so far: model ambiguities. These arise when multiple causal models fare equally well in accounting for configurational data. Mainly due to the uncritical import of an algorithm that is unsuitable for causal modeling, researchers have typically been unaware of the whole model space. As a result, there exists an indeterminable risk for practically all QCA studies published in the last quarter century to have presented findings that their data did not warrant. Using hypothetical data, we first identify the algorithmic source of ambiguities and discuss to what extent they affect different methodological aspects of QCA. By reanalyzing a published QCA study from rural sociology, we then show that model ambiguities are not a mere theoretical possibility but a reality in applied research, which can assume such extreme proportions that no causal conclusions whatsoever are possible. Finally, the prevalence of model ambiguities is examined by performing a comprehensive analysis of 192 truth tables across 28 QCA studies published in applied sociology. In conclusion, we urge that future QCA practice ensures full transparency with respect to model ambiguities, both by informing readers of QCA-based research about their extent and by employing algorithms capable of revealing them.	**Features** Case (study) based ◉ Configurational ◉ Counterfactual Experimental Generative causation Mapping Modelling ◉ Participatory Predictive Quasi-experimental Statistical association Synthesis ◉ Theory based ◉
Overview of approach QCA is used to model causation in configurational data using Boolean logic using truth tables. However, any set of data a number of different models and this paper present an approach to assessing ambiguities between models.	**Overview of application** Use of census data on local populations of grassroots associations in 22 rural municipalities of Hordaland County in Norway gathered between 1980 and 2000, to identify the essential combinations of causal conditions for their growth and volatility.
Assessment of novelty QCA was developed in 1987, but only recently been more widely adopted in different fields of evaluation. The novelty of the approach described in this paper is that provides an approach for assessing the strength of models developed using QCA.	**Methodological issues addressed by methodology** Being a structured approach to comparing cases allows attribution of effects to intervention. This paper sets out a methodology for assessing the merits of models developed by QCA.

Strengths **Authors assessment:** This paper sets out an approach to assessing ambiguities created when using QCA to model attribution. **Our assessment:** The approach is potentially useful if adopted in software used for QCA analysis.	**Weaknesses** **Authors assessment:** This paper describes potential ambiguities in QCA models resulting from software used short cutting to preferred solution rather than providing a range of alternative models for consideration **Our assessment:** the approach set out is reliant on being adopted by developers of software to run QCA.
Resource intensiveness of approach Requires specific expertise in the QCA. The level of resource is dependent on	**Robustness, Reliability and Replicability** This paper is specifically concerned with assessing the robustness of QCA models.
General Comment On reflection, this paper is problematizing the use of QCA. It is suggesting there are issues with current software that favour the identification of a preferred model to early. Less methodological innovation, more the identification of the need.	

Text 5: Beach (2018).

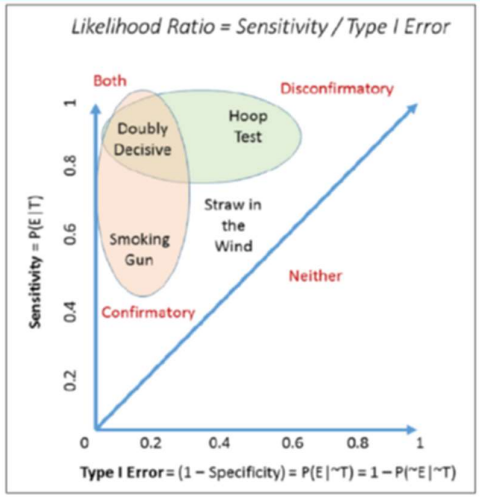
QCA followed by process tracing	Beach, D. (2018). "Achieving Methodological Alignment When Combining QCA and Process tracing in Practice." Sociological Methods & Research 47(1): 64-99.
Published Abstract This article explores the practical challenges one faces when combining qualitative comparative analysis (QCA) and process tracing (PT) in a manner that is consistent with their underlying assumptions about the nature of causal relationships. While PT builds on a mechanism-based understanding of causation, QCA as a comparative method makes claims about counterfactual causal relationships. Given the need to ensure alignment between the ontological understandings of causation that underlie a method and methodological practice, the different ontological foundations result in methodological guidelines that contradict each other, forcing the analyst to choose whether to be more in alignment with one or the other method. This article explores the implications of contrasting guidelines in a practical case study, where a QCA for sufficiency is followed by two PT case studies of positive cases.	**Features** Case (study) based ◉ Configurational ◉ Counterfactual Experimental Generative causation ◉ Mapping Modelling ◉ Participatory Predictive Quasi-experimental Statistical association Synthesis ◉ Theory based ◉
Overview of approach Process tracing (PT) can be used after QCA to develop within-case evidence of causal relationships. It allows inferences about whether the causal links of interest identified by QCA are necessary or sufficient by creating a counterfactual variation. This achieved by transforming a single case study into a form of comparative most-similar-systems design, where we hold everything else equal but the hypothesized necessary condition, and then speculate using logical arguments about whether the outcome would have been any different had the condition not been present. This article employs an fsQCA for sufficiency in a more theory-building fashion, followed by a PT case study of the most robust conjunction from the fsQCA. PT is used in a theory-testing manner, with the emphasis being a structured empirical test of whether there is evidence suggesting that a hypothesized causal mechanism exists between the found conjunction and the outcome.	**Overview of application** The paper uses as its example an exploratory evaluation of congruence between voter views and governmental positions in EU constitutional politics.

Guidelines for proper case selection that are aligned with the type of causal claims being made in fuzzy-set QCA (fsQCA) are more restrictive than those for PT case studies aimed at tracing mechanisms. Existing guidelines for case selection for PT after an fsQCA for sufficiency suggest we should only select typical cases where the fuzzy score for membership in the conjunction is lower than membership in the outcome, and that we should only select cases that are only members of one conjunction.	
Assessment of novelty Interest in QCA and PT as a methodology has grown in past five years.	**Methodological issues addressed by methodology** This article explores the practical challenges one faces when combining qualitative comparative analysis (QCA) and process tracing (PT) in a manner that is consistent with their underlying assumptions about the nature of causal relationships. While PT builds on a mechanism-based understanding of causation, QCA as a comparative method makes claims about counterfactual causal relationships.
Strengths **Authors assessment:** N/A **Our assessment:** Potentially very useful in areas of high expenditure and of strategic importance, otherwise too resource intensive.	**Weaknesses** **Authors assessment:** N/A **Our assessment:** The approach involves the need for users to develop expertise in its use that will be upfront cost and limit the availability of potential contractors.
Resource intensiveness of approach The example used relied on secondary data but potentially very resource intensive.	**Robustness, Reliability and Replicability** This approach is still to be developed and refined but the combination of QCA and PT should increase robustness of findings.

Text 6: Befani (2017).

Contribution analysis (Process tracing with Bayesian Updating)	Befani, B. and G. Stedman-Bryce (2017). "Process Tracing and Bayesian Updating for impact evaluation." Evaluation 23(1): 42-60.
Published Abstract Commissioners of impact evaluation often place great emphasis on assessing the contribution made by a particular intervention in achieving one or more outcomes, commonly referred to as a 'contribution claim'. Current theory-based approaches fail to provide evaluators with guidance on how to collect data and assess how strongly or weakly such data support contribution claims. This article presents a rigorous quali-quantitative approach to establish the validity of contribution claims in impact evaluation, with explicit criteria to guide evaluators in data collection and in measuring confidence in their findings. Coined 'Contribution Tracing', the approach is inspired by the principles of Process Tracing and Bayesian Updating, and attempts to make these accessible, relevant and applicable by evaluators. The Contribution Tracing approach, aided by a symbolic 'contribution trial', adds value to impact evaluation theory-based approaches by: reducing confirmation bias; improving the conceptual clarity and precision of theories of change; providing more transparency and predictability to data-collection efforts; and ultimately increasing the internal validity and credibility of evaluation findings, namely of qualitative statements. The approach is demonstrated in the impact evaluation of the Universal Health Care campaign, an advocacy campaign aimed at influencing health policy in Ghana.	**Features** Case (study) based Configurational ◉ Counterfactual Experimental Generative causation ◉ Mapping Modelling Participatory Predictive Quasi-experimental Statistical association Synthesis Theory based ◉
Overview of approach Process Tracing is a method, tool and a technique for data collection and analysis, reflecting its focus on theory development as much as on the search and assessment of evidence for a causal explanation. It draws causal inferences from 'historical cases', broadly intended as explanations of past events. It is based on a mechanistic understanding of causality in social realities, and starts from the reconstruction of a causal process intervening between an independent variable and an outcome, which could for example be a Theory of Change, a complex mechanism or a CMO configuration. The method distinguishes between:	**Overview of application** Evaluation of the Universal Health Care Campaign in Ghana by civil society led movement to promote health related policy priorities prior to Presidential election. Oxfam commissioned the evaluation, which tested a number of contribution claims. While this application may not appear directly relevant to HEIF, the workshop with KE practitioners identified supporting policy innovation was an important aspect of HEIF activity for universities, particularly those with social work and policy departments.

a) Possible 'reality', or an ontological entity which might or might not exist or have materialized; which is usually unobservable i.e. Theory of Change; b) The evaluator's hypothesis on the existence of that reality; and c) The observable and therefore testable implications of the existence of such reality. Process Tracing aims to minimize the inferential error we risk making when producing statements about an ontological causal reality. The backward perspective takes advantage of the fact that, at the time of the investigation, the mechanism has presumably had enough time to leave traces which provide a strong indication of its existence. Process Tracing recognizes that not all of these traces are equally informative, and as a consequence focuses on assessing the quality, strength, power, or probative value that select pieces of evidence hold in support of (or against) the causal mechanism. One of its advantages is that it allows a clear distinction between 'absence of evidence', which has little inferential power and does not add much value to what the researcher already knows, and 'evidence of absence', which on the contrary can strongly challenge a hypothesis, if it contradicts observable implications stemming from such a hypothesis. In Process Tracing, four well-known metaphors are often used to describe the different ways evidence affects our confidence about a certain mechanism or Theory of Change: • Hoop test(disconfirmatory):, • Smoking Gun test(confirmatory):, • Straw-in-the-Wind test (neither confirmatory nor disconfirmatory)and • Doubly-Decisive test (both confirmatory and disconfirmatory)	



The concepts of Process Tracing can be modelled using Bayesian Updating that allows the inferential power or probative value of a piece of evidence E for a theory T can be measured. This includes 'sensitivity' (true positives rate) - the probability that the evidence confirms that the theory holds when this is in fact the case) and false positives rate or 'Type I error' (the probability that the evidence confirms that the theory holds when this is actually not the case).

The larger the difference between the true positives rate and the false positives rate, the higher the probative value of evidence E for theory T.

Intuitively, this means that if an observed piece of evidence has a higher chance of being observed if theory T holds true (sensitivity), than if theory T does not hold true (Type I error), this constitutes a confirmation of the theory. If the opposite is true, and the evidence has a higher chance of being observed if the theory does not hold, compared with if the theory holds, observation of that evidence weakens the theory.

Finally, if the evidence has a similar chance of being observed whether the theory holds or not (sensitivity is roughly the same as Type I error), observing it will not significantly alter our confidence in the theory.

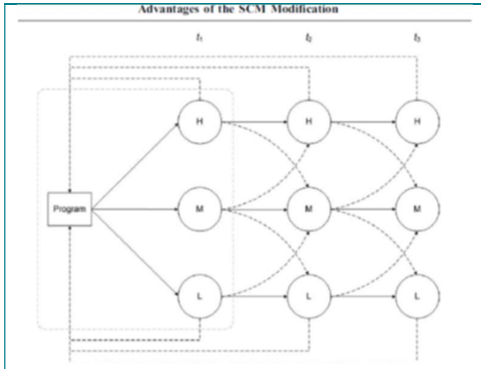
In Bayesian Confidence Updating, different pieces of evidence have different values of sensitivity and specificity, hence different likelihood ratios, and thus different abilities to alter the evaluator's initial confidence in the contribution claim. The evaluator is required to be transparent about their assumptions and confidence on the existence of the claim, and to 'declare' its observable implications ('if the claim holds true or does not – what should I expect to observe? With what probability?'). Just like in a judicial trial where evidence is produced in favour or against a defendant and the jury is left to assess the probative value of that evidence, if the prosecution cannot produce any significant evidence of guilt or if the defence finds proof that the suspect is innocent, then the suspect is considered innocent by the jury.	
Assessment of novelty High, recent development.	**Methodological issues addressed by methodology** Contribution Tracing approach aims to: reduce confirmation bias; improve the conceptual clarity and precision of theories of change; provide more transparency and predictability to data-collection efforts; and ultimately increasing the internal validity and credibility of evaluation findings. It uses Bayesian updating to model probability of evidence required for 4 process tracing tests.
Strengths **Authors assessment:** Contribution Tracing provides replicable and transparent testing of mechanisms and qualitative statements. Has the potential to contribute to qualitative evaluation methods being taken more seriously for two main reasons. The first is that the questions guiding data collection directly address the core reason we conduct data collection, which is to increase (or decrease) our confidence in a hypothesis. Every piece of evidence considered is assessed in terms of its ability to alter this confidence, and different Process Tracing tests can be used for this purpose, allowing the evaluator to fruitfully and transparently navigate a high number of hypotheses and a high number of pieces of evidence at the same time. This will allow evaluators to gradually connect claims with pieces of evidence and obtain a shortlist of a few, plausible claims which are strongly supported. The second reason is that, similarly to what happens in traditional quantitative and statistical methods, confidence can be measured, both with probability and qualitative rubrics. The measurements and qualitative assessments of confidence can be fully shared with a 'jury' of experts, stakeholders, or fellow evaluators, who can either validate or refine the assessments. The process is replicable and will produce confidence intervals describing	**Weaknesses** **Authors assessment:** None given **Our assessment:** Dependent on good documentary evidence. May be subject to recall bias if stakeholders need to be consulted to fill evidence gaps.

the power of given pieces of evidence to change our confidence in the contribution claim, increasing the robustness and ultimately the 'objectivity' of qualitative evaluation findings **Our assessment:** Ability to use existing available documentary sources to trace backwards could reduce the need for new primary data collection. The use of panels to assess the probative value of evidence provides transparency and means of dealing with different types and sources of evidence.	
Resource intensiveness of approach Likely to be resource intensive, though we are aware evaluators are developing a 'lite' version.	**Robustness, Reliability and Replicability** Likely to be very high if the evaluator adheres to main steps and principles.

Text 7: Coryn (2009).

Success Case Method with Longitudinal follow-up	Coryn, C. L. S., et al. (2009). "Adding a Time-Series Design Element to the Success Case Method to Improve Methodological Rigor An Application for Nonprofit Program Evaluation." American Journal of Evaluation 30(1): 80-92.
Published Abstract Brinkerhoff's Success Case Method (SCM) was developed with the specific purpose of assessing the impact of organizational interventions (e.g., training and coaching) on business goals by analyzing extreme groups using case study techniques and storytelling. As an efficient and cost-effective method of evaluative inquiry, SCM is attractive in other contexts as well, although few examples of such uses are to be found in the published literature. However, modifications of the SCM concept and design are sometimes necessary for implementing the approach in nonprofit environments where business goals are not necessarily an explicit objective. This method note demonstrates how SCM was modified and extended to a social service context, in which the program evaluated was aimed at reducing chronic homelessness and unemployment. Modifications included defining success in a non profit setting and adding a time-series element to the design features of traditional SCM to increase methodological rigor.	**Features** Case (study) based: ◉ Configurational: Counterfactual: Experimental: Generative causation: Mapping: Modelling: ◉ Participatory: ◉ Predictive: Quasi-experimental: Statistical association: Synthesis: Theory based:
Overview of approach Success Case Method (SCM) evaluates the impact of interventions on business goals. In general, this involves assessing how well an organizational intervention is working (primarily in for-profit contexts) by focusing on extreme groups and identifying the contextual factors that differentiate successful from unsuccessful adopters of new initiatives. An alternative approach to examine causal associations when more scientifically rigorous, sophisticated, and elegant designs are unethical, unpractical, too costly, or simply unfeasible. SCM is conducted using a five-step procedure: 1. Focus and plan the SCM 2. Create an impact model 3. Survey all program recipients to identify success and non-success cases 4. Interview a random sample of success and non-success cases and document their stories	**Overview of application** Evaluation of a local-level program aimed at reducing and preventing homelessness and chronic unemployment. The program was developed in response to a well-documented need in a mid-sized Midwestern city, with the long-term objective of reducing recipients' need for public assistance. The program was established as a joint effort between three local service providers who initially offered independent programs to the same target group but under rules and regulations established by three different funding sources with different outcome goals. The collaborative program provided pooled services to recipients for 1 year in an effort to stabilize housing and employment needs. One specific feature of the program was that services were tailored to individual needs via intensive case management, housing subsidies, and other unique service provisions (e.g., budget management workshops, time management workshops, interviewing skills training). SCM was implemented component of a larger, comprehensive evaluation.

5. Communicate findings, conclusions, and recommendations. The approach is modified with the addition of time-series design element was also intended to reduce some of the threats to internal validity inherent in most single-group designs by identifying and eliminating as many plausible, competing explanations for observed effects as possible. The authors assert their modified SCM increases methodological rigor through the addition of longitudinal design elements, causal inferences can be better supported when stronger cause probing designs (e.g., randomized experimental designs, regression discontinuity designs, interrupted time series designs) are not feasible.	Rationale for using SCM to assess the program's impact on program recipients was to plausibly eliminate rival hypotheses about core factors leading to sustainable success of service recipients and to do so in an efficient and cost-effective manner
Assessment of novelty SCM has been around a while but has significant promise if build into a process of organisational learning.	**Methodological issues addressed by methodology** The addition of longitudinal element to SCM aims reduce threats to internal validity.
Strengths **Authors assessment:** Figure below illustrates some of the advantages offered by adding a longitudinal, time-series design element to traditional SCM, which include (a) the ability to identify growth (dashed, upward turning arrows) and decay (dashed, downward turning arrows and the reasons for them (e.g., delayed manifestation of an effect following a treatment), (b) the ability to identify long-term program effects and for who and why (or why not) those effects are, or are not, sustained, and (c) the ability to provide useful feedback to the program at various points during the evaluation (represented by the feedback loops from t1, t2,and t3 to the program in Figure).	**Weaknesses** **Authors assessment:** Even though SCM is an efficient and cost-effective method of assessing the impact of organizational interventions, the modification for applying the method in the context of human and social service (i.e., non profit) program evaluation does reduce both of these advantages to some extent by adding to the time and resources required to conduct the time-series SCM. However, these limitations are offset by the gains in understanding of the onset (i.e., whether immediate or delayed) and permanence (i.e., whether continuous or discontinuous) of effects (as well as greater insight into who the program works for, who it does not, and why, by analysis of subsets or units across time) that would otherwise not be detected with traditional SCM **Our assessment:** The use of the methodology may be limited to interventions that have greater risk or estimates of benefits to justify its use. There are also usually risks associated with longitudinal studies such as attrition and increasing volume of data collection reducing the quality of data collected.

 **Our assessment:** This approach is probably more applicable to individual universities interested in developing a better understanding of which of their HEIF (or HEIF type activities funded by other sources) have the longest and most sustainable impact. Logistically it would be difficult to execute at a fund level, though it could potentially provide valid information for accountability purposes where it focuses effort on identifying the factors that differentiate successful from unsuccessful applications of HEIF for specific purposes.	
Resource intensiveness of approach Potentially very resource intensive with the addition of new waves.	**Robustness, Reliability and Replicability** Potential to increase robustness of findings. May be difficult to replicate studies.

Text 8: Dart (2003).

Most significant change Also been called "the Evolutionary Approach to Organisational Learning", the "Story Approach" and "monitoring without indicators".	Dart, J. and R. Davies (2003). "A dialogical, story-based evaluation tool: The most significant change technique." American Journal of Evaluation **24**(2): 137-155
Published Abstract The Most Significant Change (MSC) technique is a dialogical, story-based technique. Its primary purpose is to facilitate program improvement by focusing the direction of work towards explicitly valued directions and away from less valued directions. MSC can also make a contribution to summative evaluation through both its process and its outputs. The technique involves a form of continuous values inquiry whereby designated groups of stakeholders search for significant program outcomes and then deliberate on the value of these outcomes in a systematic and transparent manner. To date, MSC has largely been used for the evaluation of international development programs, after having been initially developed for the evaluation of a social development program in Bangladesh (Davies, 1996*). This article provides an introduction to MSC and discusses its potential to add to the basket of choices for evaluating programs in developed economies. We provide an Australian case study and outline some of the strengths and weaknesses of the technique. We conclude that MSC can make an important contribution to evaluation practice. Its unusual methodology and outcomes make it ideal for use in combination with other techniques and approaches.	**Features** Case (study) based: ◉ Configurational: Counterfactual: Experimental: Generative causation: Mapping: Modelling: Participatory: ◉ Predictive: Quasi-experimental: Statistical association: Synthesis: Theory based:
Overview of approach MSC involves the regular collection and participatory interpretation of "stories" about change rather than predetermined quantitative indicators. It is through the deliberation and dialogue that surrounds the selection of stories that represent significant change that impacts are surfaced that valued by stakeholders MSC has seven key steps (Davies, 1996): (1) the selection of domains of change to be monitored, (2) the reporting period, (3) the participants, (4) phrasing the question, (5) the structure of participation,	**Overview of application** Target 10 is a collaborative dairy extension program that works with farmers to improve farm productivity and profitability in a sustainable manner. The program focuses on issues of high priority to the industry, such as grazing management, business management, dairy cow nutrition, soils and fertilizers, and natural resource management. Information is extended to farmers through courses, discussion groups, newsletters, comparative analysis, field days, focus farms, demonstrations and other activities. In 1998, the program employed around 50 staff and operated across four regions of the Australian State of Victoria

(6) feedback, and (7) verification.	
Assessment of novelty Developed for evaluation for international development. This text describes its use in extension programme in Australia to improve productivity and profitability of dairy farmers. MSC has points in common with the critical incident technique (CIT). Both MSC and CIT ask stakeholders to recall memorable experiences of what they consider to be critical (significant) events. Both techniques can also involve ongoing reporting of events or incidents deemed to be significant by stakeholders. The key difference between them is that CIT focuses on variations from prescribed practice and generates negative information, whereas MSC searches for significant outcomes through an inductive process, and tends to generate mainly positive information.	**Methodological issues addressed by methodology** MSC can be conceived as a form of dynamic values inquiry whereby designated groups of stakeholders continuously search for significant program outcomes and then deliberate on the value of these outcomes. This process contributes to both program improvement and judgment. Of late there has been discussion of the importance of values and values inquiry in evaluation. Advocates suggest that criteria used to judge programs do not always reflect stakeholder values and that uncovering these values can help to ensure that programs meet needs. Dynamic values inquiry is a central and critical part of MSC. When key stakeholders select stories of significant change, they partake in an ongoing process of deliberation about the value of individual outcomes. As MSC occurs over time, the dialogue is responsive to the changing nature of the program and its context. Stories about the impact of interventions can infiltrate the collective memory of an organization, helping program staff to gain and retain a more deeply shared understanding of what is being achieved. This creates a common base to enter into dialogue about what is desirable in terms of expected and unexpected outcomes.
Strengths **Authors assessment:** MSC can result in a multitude of positive effects on programs and be conceptualized in a number of ways. But to us, its key strength lies in the ability to facilitate a dynamic dialogue between designated stakeholders. This dialogue concerns the question "what do we really want to achieve and how will we produce more of it?" Another strength of MSC is that people seem to enjoy the process; this seems to be largely due to the storytelling process. It is also refreshingly different and thought provoking. Because of these strengths, MSC is a particularly useful addition to evaluation portfolios for participatory programs that have diverse, complex outcomes, with multiple funders and stakeholder groups. **Our assessment:** Potentially useful methodology in complex interventions where outcomes are unclear at the onset and the programme design is not fixed and needs to be adaptive to change.	**Weaknesses** **Authors assessment:** It would be misleading to suggest that MSC was implemented smoothly and easily across the program. At various stages in the 12-month trial, problems arose and were addressed where possible. As the process was an iterative one, it was possible to modify each round on the basis of feedback from the previous round. The main problems were associated with the time taken to run the process and the need to develop a system to ensure confidentiality. Some people also disliked the competitive aspect of the process, feeling disillusioned when their stories did not get selected. **Our assessment:** Requires a high level of buy-in from stakeholders and may be difficult to maintain in the presence of time and resource constraints. While generating information on impacts (including unintended consequences not usually included in a logic models) may lend itself more to continuous quality improvement activity than impact evaluation to support accountability.

Resource intensiveness of approach	Robustness, Reliability and Replicability
Can be resource intensive even when built into normal programme management processes.	The iterative approach and multiple steps increase robustness and reliability of evaluation findings.
* Davies, R. J. (1996). An evolutionary approach to facilitating organisational learning: An experiment by the Christian Commission for Development in Bangladesh. Swansea. UK: Centre for Development Studies [online]: http://www.swan.ac.uk/cds/rd/ccdb.htm. This paper has also been published, with some variations, in D. Mosse, J. Farrington, and A. Rew (1998) Development as process: concepts and methods for working with complexity. London: Routledge/ODI (pp 68–83); and in Impact assessment and project Appraisal, 16. No. 3, September 1998, 243–250.	

Text 9: Gates (2018).

Critical Systems Heuristics	Gates, E. F. (2018). "Toward Valuing With Critical Systems Heuristics." American Journal of Evaluation 39(2): 201-220.
Published Abstract Evaluation is defined by its central task of valuing-the process and product of judging the merit, worth, or significance of a policy or program. However, there are no clear-cut ways to consider values and render value judgments in evaluation practice. There remains contention in the evaluation field about whether and how to make value judgments. No approach to valuing eliminates the uncertainty, plurality, and potential for conflict that comes with considering values. This article explores what critical systems heuristics (CSH), an area of applied systems thinking, might contribute to four long-standing issues regarding valuing: envisioning the social value of evaluation, framing the evaluand and evaluation, selecting and justifying criteria, and determining the roles of the evaluator(s) and stakeholders in valuing. CSH contributes concepts and tools that, in theory, support more reflective, responsible valuing although further practical application is needed.	**Features** Case (study) based Configurational Counterfactual Experimental Generative causation Mapping Modelling Participatory Predictive Quasi-experimental Statistical association Synthesis Theory based
Overview of approach This paper sets out a potential approach that evaluators might use in determining the value of interventions based on developing a set of heuristics (rules of thumb) grounded in "critical" systems thinking. CST is a version of systems thinking that pays attention to power relations and, specifically, how decisions are (and should be) made about which interrelationships and perspectives are considered relevant and irrelevant in an inquiry or intervention. It is not an evaluation approach in itself but complementary structured approach for framing evaluations that can be used with any type of evaluation. It involves: • **Considering the Social Value of Evaluation**: this involves the evaluator being aware of their and their client position regarding whether evaluations should advance select social and political values and, if so, which and whose values. • **Framing the Evaluand and Evaluation.** This includes setting boundaries within which an evaluand will be examined and an evaluation can be designed and carried out. Evaluators generally lack practical frameworks for making these choices. Typically, these choices are primarily descriptive (regarding what is) and not normative (regarding what could or should	**Overview of application** No specific example given in the text. This papers novelty relates to setting out an evaluator might go about undertaking an evaluation of complex interventions using one or more the novel methods identified.

be). Boundary critique and critical system heuristic questionnaire (CSH-Q) may offer one set of tools to aid in explicit and systematic choices between boundaries and framings.
• **Surfacing existing assumptions** and ways of understanding an evaluand and evaluation.
• **Consider consequences (practical, political and ethical) and alternatives.** Practical consequences influence the logistics of implementing a policy or program, conducting an evaluation, or the wider situation of interest. Political consequences influence who or what has and does not have power, control, voice, and agency in an evaluand or evaluation. Ethical consequences are related with good/right and bad/wrong and stem from the fact that all framing is influenced by ethical beliefs, worldviews, or normative assumptions about how things should be in society
• **Select, justify, and revise as needed the framing**(s) used while remaining open to contestation and revision. Inevitably, consideration of alternative framings and consequences will be limited, as alternatives are possible.
• **Selecting and Justifying Criteria**: Evaluative claims about the merit, worth, and significance of social policies and programs are made based on value-laden criteria and factual claims. There are numerous sources of potential criteria including stated objectives, effectiveness, relevance, equity focus, efficiency, social impact, sustainability, cultural relevance and responsiveness, established requirements, expert opinion, and needs assessment.
• CSH suggests a way of determining and representing value by constructing and contrasting descriptive and normative maps. This involves comparing a descriptive map of what is with a normative (or ideal) map of what should be to generate an evaluative assessment, judgment, critique, or conversation.
• **Determining Roles of the Evaluator(s) and Stakeholders in Valuing:** Alkin, Vo, and Christie (2012)* distinguish between three roles for the evaluator in making judgments of the value of a policy or program. These are:
  ○ stakeholders rather than the evaluator have the primary responsibility for such judgments.
  ○ stakeholders together with the evaluator are jointly responsible for rendering value judgments. This involves evaluators providing the data and establishing a framework for valuing
  ○ the evaluator alone is responsible for rendering value judgments. This typically involves valuing based on evaluator values, evaluation expertise, program expertise, or scientific appraisal.
• CSH invites evaluators to incorporate a witness role for groups, interests, and worldviews potentially affected by the situation of interest an evaluand (and evaluation) addresses to participate directly or be represented in the valuing process.	

Assessment of novelty High as the paper describes an approach to assessing value of outcomes in evaluation studies.	**Methodological issues addressed by methodology** Framing evaluands and evaluations more explicitly and systematically could potentially have several benefits, although these are preliminary and require empirical examination. First, this process could help identify and anticipate potential practical, political, and ethical consequences of an evaluand on different groups and interests.
Strengths **Authors assessment:** N/A **Our assessment:** Provides useful basis for increasing transparency and acceptance of judgements made by evaluators.	**Weaknesses** **Authors assessment:** N/A **Our assessment:** Need for a more practically written set of guidelines for practitioners.
Resource intensiveness of approach Would add to the resource requirement as adds another dimension to the process.	**Robustness, Reliability and Replicability** Will help increase the robustness of other approaches.
* Alkin, M. C., Vo, A. T., & Christie, C. A. (2012). The evaluator's role in valuing: Who and with whom. New Directions for Evaluation, 133, 29–41.	

Text 10: Kittel (2013).

Process tracing	Kittel, B. and D. Kuehn (2013). "Introduction: reassessing the methodology of process tracing." European Political Science 12(1): 1-9.
Published Abstract Although having been practised in the Social Sciences for decades, it was only in recent years that process tracing has gained prominence in methodological debates in political science. In spite of its popularity, however, there has been little success in formalising its methodology, defining its standards, and identifying its range of applicability. This symposium aims at furthering our understanding of the methodology by discussing four essential aspects: the underlying notion of causality, the role of theory, the problem of measurement in qualitative research, and the methodology's relationship with other forms of qualitative inquiry. It brings together methodological and substantive articles by young European scholars and summarises a round-table discussion with Peter A. Hall held at a workshop at the University of Oldenburg, Germany, in November 2010.	**Features** Case (study) based Configurational Counterfactual Experimental Generative causation Mapping Modelling Participatory Predictive Quasi-experimental Statistical association Synthesis Theory based
Overview of approach Process tracing (PT) emerged as a response to the debate on the usefulness of QCA for stablishing causal inferences. However, there are questions relating to the principles and practical application of the process tracing in testing causal theories. The authors have noted the causal inference methodological debates across the quantitative–qualitative divide in the 1990s and early 2000s has moved on from the question of internal validity to external validity. The debate around internal validity centred on whether causal inference are possible through small-n techniques such as process tracing. The ability to test causal statements through tracing the processes that lead to a certain outcome has been accepted. Some consider process tracing is second only to experimental research in uncovering causal relationships and testing causal arguments, because contrary to statistical analyses it allows the identification of causal mechanisms and proximate causal relationships, and thus does not fall prey to the correlation-causation fallacy.	**Overview of application** This paper provides an overview of the issues and does not provide a specific example of the use of PT.

PT relies on case studies and these present issues of generalisation and external validity, which essentially correspond to the same problems in large-n research. However, the main contribution of case studies is their ability 'to incrementally refine middle-range contingent generalisations, either by broadening or narrowing their scope or introducing new types and subtypes through the inclusion of additional variables' (George and Bennett, 2004: 124). Process tracing can be a powerful tool for a rigorous 'origin science' as it allows the analysis of temporal sequences of events that lead to the explanandum. As such, it is a promising technique for studying phenomena such as institutional change, precisely because rare, dynamic, and highly contingent events do not lend themselves readily to quantification and statistical analysis. Three topics are of importance for the application of process tracing. • Debate on the notion of causality and its implications for the method. • The role of theory. While all process-tracing theorists and practitioners agree that the ultimate goal is developing and testing theories, there seems to be a split concerning the necessity to specify the expected process a priori. • Measurement in qualitative research. While the literature usefully conceptualises and operationalising the main research variables of interest (the independent and the dependent variables) it has not addressed conceptualising and systematically measuring what Collier et al call 'causal process observations' (CPOs). CPOs are the set of data employed in-case analyses and are the main data source on which process-tracing inference rests. The absence of conceptualisation and standardised measurement procedures may lead to measurement error and undermines the ability to replicate findings. Finally, there seems to be some confusion concerning the status of process tracing within the Social Science toolbox and its relation to other forms of qualitative research, particularly historical analysis. By ignoring other disciplines such the historians' established set of methods and techniques reduces the ability to improve the quality of process tracing	
Assessment of novelty Its origins can be traced back to 1975 as research methodology that has only recently gained interest as an evaluation methodology.	**Methodological issues addressed by methodology** This paper sets out the methodological debates around the use of process tracing. Process tracing was developed to test causal inferences in small n studies. The debate has moved from internal validity to external validity and sets out four issues in its practical use (notion of causality, role of theory, measurement in qualitative research, and its relationship with other forms of qualitative enquiry). [Befani's paper above on Contribution tracing gives a better overview of process tracing methodology]

Strengths	Weaknesses
Authors assessment: N/A **Our assessment:** Process tracing has the advantage that can deployed retrospectively. It therefore similar to case reviews of rare health outcomes in medical research and as such exhibits similar advantages.	**Authors assessment:** N/A **Our assessment:** Dependent on the quality and availability of secondary data sources which may be incomplete and/or 'messy' making it difficult to undertake process tracing. This may be resolved by careful selection of cases with the most complete information that representative of different outcomes.
Resource intensiveness of approach	**Robustness, Reliability and Replicability**
Potentially high, unless limited to a small carefully selected number of cases.	If well documented and systematically undertaken – high.

Text 11: Salter (2014).

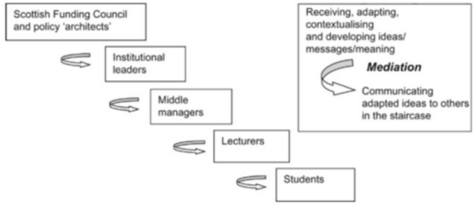
Realist evaluation **(Systematic review of use in Knowledge translation)**	Salter, K. L. and A. Kothari (2014). "Using realist evaluation to open the black box of knowledge translation: a state-of-the-art review." Implementation Science 9: 14.
Published Abstract **Background:** In knowledge translation, complex interventions may be implemented in the attempt to improve uptake of research-based knowledge in practice. Traditional evaluation efforts that focus on aggregate effectiveness represent an oversimplification of both the environment and the interventions themselves. However, theory-based approaches to evaluation, such as realist evaluation (RE), may be better-suited to examination of complex knowledge translation interventions with a view to understanding what works, for whom, and under what conditions. It is the aim of the present state-of-the-art review to examine current literature with regard to the use of RE in the assessment of knowledge translation interventions implemented within healthcare environments. **Methods:** Multiple online databases were searched from 1997 through June 2013. Primary studies examining the application or implementation of knowledge translation interventions within healthcare settings and using RE were selected for inclusion. Varying applications of RE across studies were examined in terms of a) reporting of core elements of RE, and b) potential feasibility of this evaluation method. **Results:** A total of 14 studies (6 study protocols), published between 2007 and 2013, were identified for inclusion. Projects were initiated in a variety of healthcare settings and represented a range of interventions. While a majority of authors mentioned context (C), mechanism (M) and outcome (O), a minority reported the development of C-M-O configurations or testable hypotheses based on these configurations. Four completed studies reported results that included refinement of proposed C-M-O configurations and offered explanations within the RE framework. In the few studies offering insight regarding challenges associated with the use of RE, difficulties were expressed regarding the definition of both mechanisms and contextual factors. Overall, RE was perceived as time-consuming and resource intensive. **Conclusions:** The use of RE in knowledge translation is relatively new; however, theory-building approaches to the examination of complex interventions in this area may be increasing as researchers attempt to identify what works, for whom and under what circumstances. Completion of the RE cycle may be challenging, particularly in the development of C-M-O configurations; however, as researchers approach challenges and explore innovations in its application, rich and detailed accounts may improve feasibility. Keywords: Realist evaluation, Knowledge translation, State-of-the-art review	**Features** Case (study) based: ☐ Configurational: ☐ Counterfactual: ☐ Experimental: ☐ Generative causation: ☐ Mapping: ☐ Modelling: ☐ Participatory: ☐ Predictive: ☐ Quasi-experimental: ☐ Statistical association: ☐ Synthesis: ◉ Theory based: ☐

Overview of approach Exploratory narrative systematic review to understand use of realist evaluation in knowledge translation. See Archibald (2018) above for an overview of the RE approach.	**Overview of application** • Note this is a review of the use of RE in knowledge translation. • Inclusion criteria: Primary studies examining a) the application or implementation of knowledge translation interventions or strategies b) using realist evaluation c) within healthcare settings were identified for inclusion. • To address the identified review questions, information was abstracted from each study identified for inclusion to address the core elements of the RE framework as follows: a) identification of linked C-M-O configurations to inform testable hypotheses, b) the use of multiple and/or mixed methods to interrogate the proposed C-M-O configurations, and c) explanatory focus; that is, did the study attempt to explain outcomes in terms of underlying mechanisms and contextual influences to present findings that helped to explain how the intervention might or might not have worked, for whom and under what circumstances. . • In order to address feasibility of RE, all information reported with regard to challenges associated with its application and the ways in which authors have attempted to address the challenges they encountered was also abstracted from each source document.
Assessment of novelty This is review of RE in knowledge translation which was relatively novel when undertaken as indicated by the small number of studies that met the study inclusion criteria.	**Methodological issues addressed by methodology** Feasibility of realist evaluation in knowledge translation. Although appealing for its theory-building and explanatory focus, adoption may have been limited by challenges associated with the application of realist evaluation. Any form of evaluative research that is intended to unearth underlying program mechanisms is likely to be labour and resource-intensive and, while interesting and informative, a full realist evaluation may not always be possible or appropriate.
Strengths **Authors assessment:** N/A **Our assessment:** Well conducted narrative synthesis with clear aims and objectives, search strategy, inclusion criteria and reporting.	**Weaknesses** **Authors assessment:** It has been suggested that theory-based evaluation intended to investigate and explain mechanisms is both labour- and resource-intensive. Pawson and Tilley noted that RE is an intellectually challenging process. There are no simple steps or strict methodological rules to follow and no standardized approach to take. Challenges or potential challenges associated with undertaking RE were noted in several of the publications included in the present review. It was anticipated that its application would result in a substantial number of reported difficulties or limitations. However, relatively few articles (n = 5) included a

	discussion of study challenges or limitations specific to the application of realistic evaluation. Overall, RE might be considered time consuming and resource intensive. Several authors noted a substantial investment of time required for discussion (particularly during development of initial C-M-O configuration and refinements to proposed configurations), while others reported making adaptations to the RE process in order to work within available project time and resources **Our assessment:** No significant concerns regarding the conduct of the systematic review.
Resource intensiveness of approach The review indicates RE is resource intensive.	**Robustness, Reliability and Replicability** Can produce robust findings but not designed to produce replicable and generalizable findings.

Text 12: Saunders (2015).

Evaluative research (Including Implementation staircase)	Murray Saunders, Cristina Sin and Steven Dempster (2015). Evaluative research in higher education policy analysis. In Theory and method in higher education research. (2015).
Published Abstract This chapter will focus on the use of evaluative research in higher education policy analysis. The approach will be illustrated by reference to higher education policy in Scottish higher education, with particular reference to the longitudinal evaluative research of support of teaching and learning (T&L) (the Quality Enhancement Framework or QEF). The chapter will discuss the features of the research process which are shaped by evaluation theory. We adopt a theoretical position on policy research which foregrounds the situated experience of policy as a core research focus. Policy is depicted as being underscored by an implicit theory of change which is used to structure and orientate the research focus. The design of the research is characterised by the involvement of potential users of the research output, with implications on the way in which findings are articulated, presented and ultimately used, along with aspects of the evaluative research design. The case study of the QEF will be contextualised, and the intersection between the design features and theoretical approaches, and the use and usability of research outputs, will be established.	**Features** Case (study) based ◉ Configurational Counterfactual Experimental Generative causation Mapping Modelling Participatory ◉ Predictive Quasi-experimental Statistical association Synthesis ◉ Theory based ◉
Overview of approach Evaluative research approach integrates three traditions of evaluation – utilisation-focused evaluation, theory-based evaluation and developmental evaluation. *Utilisation-focused evaluation* emphasises the task of providing key stakeholder groups with evidence they can use to judge the value and worth of strategies. It takes seriously the needs of commissioners of evaluative research, and has a strong sense of the need that such research has usability. *Theory-based evaluation* focuses on the connections between strategies and intentions. It has a firm basis in evidence, but is open to unintended and unanticipated processes and outcomes. It helps to articulate the tacit theories of change embedded in policy strategies, and the adaptations and modifications, which occur as a policy is created in practice. *Developmental evaluation* is an approach which is designed to support policy implementation in complex and dynamic environments, with the primary purpose of exploring possibilities and experimenting with innovations without the goal of	**Overview of application** This chapter describes the application of the implementation staircase in the evaluation of Quality Enhancement Framework (QEF) to support teaching and learning in higher education in Scotland. Evaluations sought to identify the situated experience of policy, that is how elements of the QEF were received, interpreted and enacted by professionals 'on the ground'. This involved investigating individuals' narratives concerning these policy elements, and considering how their narratives and reported enactments of policy are shaped by numerous influences. These influences are complex and multi-faceted and include: • conceptualisations of professionalism; meanings attached to membership of particular academic disciplines;

arriving at a fixed intervention. Uses the concept of 'social practice' as a lens for the depiction and understanding of what goes on in social domains like higher education. The idea of practice is a key aspect of socio-cultural theory, and all social life can be interpreted as consisting of a series or clusters of practices in different fields of activity, within families, friendship groups, at work and so on. Uses the metaphor of the implementation staircase to understand practice. The implementation staircase helps to illustrate the positions of the different actors: these actors' perspectives may differ, underlining the importance of constructing the experience of policy from the positions and points of view of the main actors in a policy environment. The implementation staircase metaphor is also evocative of the two-fold function of each group, which acts as both a receiver and an agent of policy messages. As a result, the message undergoes adaptation and is interpreted differently according to the distinctive and situated experience of each stake-holding group. The stakeholder positions and narratives identifiable on the implementation staircase for the QEF policy are illustrated in figure.	• feelings about the overarching culture and priorities of their organisations; • sense of 'goodness of fit' between a policy/strategy and their own professional concerns; • and the sense of power or powerlessness associated with their particular location on the implementation staircase In order to access these narratives and document change over a period of 10 years, between 2004 and 2014, evaluations have drawn on a combination of questionnaires and visits to all institutions in the Scottish higher education system (included semi-structured interviews with managers and Students' Union officers and focus groups (teaching staff and students). 
Assessment of novelty Moderate novelty. Innovation is from the combination of different evaluation traditions.	**Methodological issues addressed by methodology** Unclear. However, the approach does attempt to increase relevance and take account of power dynamics.
Strengths **Authors assessment:** • Understands social practice as a core construct in understanding policy in action • Is a nimble and adaptive process of research suited to complex policy environments • Highlights the socially and organisationally situated nature of responses to policy implementation • Uses estimations of value and worth by social actors as a resource for the derivation of practice clusters • Is method neutral but tends towards constructivist approaches • Distinguishes between policies-in-texts and policies-in-action • Provides sensitive and illuminative resources for decision-making	**Weaknesses** **Authors assessment:** N/A **Our assessment:** One potential weakness is by trying to meet a broader range of evaluation objectives could result in evaluations not delivering well on any specific objectives.

• Sees policy learning as an iterative and evolving process of development. **Our assessment:** Integrates a range of evaluation approaches that individually generate different set of benefits.	
Resource intensiveness of approach Moderate	**Robustness, Reliability and Replicability** Moderate

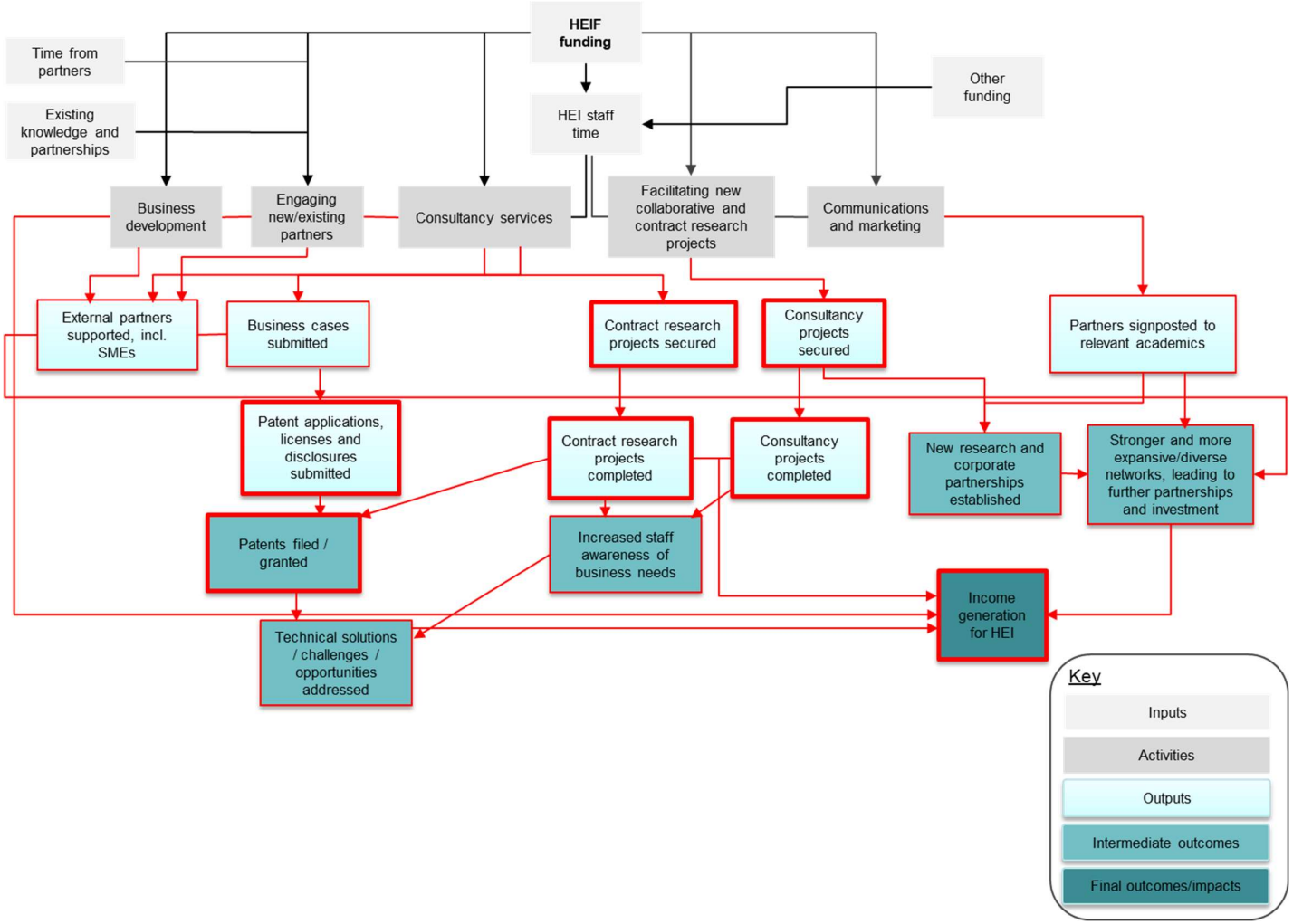
Text 13: Tho (2015)

Fuzzy set QCA	Tho, N. D. and N. T. M. Trang (2015). "Can knowledge be transferred from business schools to business organizations through in-service training students? SEM and fsQCA findings." Journal of Business Research 68(6): 1332-1340
Published Abstract Employing the ability–motivation–opportunity model, this study proposes that knowledge acquired from business schools by students, students' intrinsic motivation, and innovative culture of business organizations are factors that affect the transfer of knowledge from business schools to business organizations through in service training students. Using a sample of 843 in-service training business students in Vietnam, the results from SEM (structural equation modeling) support the hypotheses, except for the impact of innovative culture on knowledge transfer. However, the results from fsQCA (fuzzy-set qualitative comparative analysis) with the same data set reveal that none of the above-mentioned factors are sufficient conditions for knowledge transfer. Instead, combinations of these three factors are. Overall, it is believed that the study findings shed light on a new channel of knowledge transfer, that is, in-service training students, not investigated by prior research.	**Features** Case (study) based: ☐ Configurational: ◉ Counterfactual: ☐ Experimental: ☐ Generative causation: ☐ Mapping: ☐ Modelling: ☐ Participatory: ☐ Predictive: ☐ Quasi-experimental: ☐ Statistical association: ◉ Synthesis: ☐ Theory based: ◉
Overview of approach Structural equation modelling is a multivariate statistical analysis technique that combines factor analysis and multiple regression analysis to analyse structural relationships. This technique is used to analyse the structural relationship between measured variables and latent constructs. It involves estimating multiple and interrelated dependence in a single analysis using two types of variables endogenous and exogenous. Endogenous variables are equivalent to dependent variables and are equal to the independent variable. Fuzzy-set qualitative comparative analysis (fsQCA) is a variation on QCA but instead of dichotomous scale (0/1) involves scoring of variable between 0 and 1.	**Overview of application** Examines the main factors affecting knowledge transfer from business schools to business organizations through in-service training students. Employing the ability–motivation–opportunity model (AMO) knowledge transfer, the study identifies three determinants of knowledge transfer. Test hypotheses using a survey data set collected from in-service training students in Ho Chi Minh City by means of structural equation modelling (SEM) and fuzzy-set qualitative comparative analysis (fsQCA)
Assessment of novelty Relatively novel	**Methodological issues addressed by methodology** This paper uses both structural equation modelling (statistical approach) and Qualitative Comparative Analysis to model causation in service training in university-industry knowledge transfer in Vietnam. Fuzzy set QCA is more

	quantitative in nature than normal QCA that variables are scored between 0 and 1 rather than dichotomous (0/1).
Strengths **Authors assessment:** This study fills a gap in knowledge transfer: the transfer of knowledge from business schools to business organizations through in-service training students and further strengthens the theoretical aspects of the AMO model. **Our assessment:** Interesting comparison of a quantitative and qualitative approach to understanding relationships between variables.	**Weaknesses** **Authors assessment:** A closer examination of the configurations reveals that none of the conditions (intrinsic motivation, acquired knowledge and innovative) is sufficient conditions for the occurrence of knowledge transfer. Nevertheless, combinations of these conditions are. This is an INUS condition (insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result **Our assessment:** the text could have discussed more the application and merits of using fsQCA.
Resource intensiveness of approach Analytical methodology that can be used on secondary data. However, the collection and preparation of good data required can be resource intensive.	**Robustness, Reliability and Replicability** Moderate.

Annex C: Theories of Change to HE-BCI metrics

Facilitating the research exploitation process



Skills and human capital development (including enterprise education)