OXFORD

And then a miracle occurs—a review of theory of change models for societal impact of research

Ole Henning Sørensen*[1] , Stine Dandanell Garn[1], and Steffen Bohni Nielsen[1]

[1]The National Research Centre for the Working Environment, Lersø Parkallé 105, København Ø, 2100, Denmark
*Corresponding author. The National Research Centre for the Working Environment, Lersø Parkallé 105, 2100 København Ø, Denmark. E-mail: ohs@nfa.dk

Abstract

Through an umbrella review, this article identified and surveyed 24 societal impact of research (SIR) models. Most of these models were developed within health domains and in Anglo-Saxon countries. The authors mapped the SIR models against constituent components of a robust theory of change. The study found that logic models were predominantly used to conceive SIR models. Yet, only nine models had explicit causal links, and only two made explicit assumptions about why research contributes to societal change. The old proverb among evaluators—when using theories of change to describe change—"and then a miracle occurs ... ," rings uncomfortably true to the current state of SIR theorizing. Further theorizing and conceptual clarity are needed to advance the science of research impact.

Keywords: societal impact; research impact; research to practice; research to policy; theory of change; assumptions; evaluation.

1. Introduction

In recent years, research funders have increasingly demanded that researchers and research institutions address the impact of their research (Sørensen et al., 2022). In doing so, researchers, funders, and policymakers typically distinguish between the academic impact and societal impact of research (SIR) (D'Este et al., 2018).

Conceptualization of methodologies and instruments to measure academic impacts have evolved considerably (Aksnes et al., 2019) and have been quantified and codified in rankings of journals, research institutions, departments, and researchers to an extent where prominent institutions have called for a more balanced assessment of researchers and research institutions (Coara 2022).

There has been a similar call to conceptualize, measure, and demonstrate the societal impact of research (Kuruvilla et al., 2006; Pedersen et al., 2020). However, such conceptualization, let alone measurement of the SIR, is admittedly more complex and elusive.

Drawing on the extant literature, SIR models can be defined as conceptual frameworks seeking to describe and explain how research processes, outputs, interactions, and institutions contribute to changes in society, encompassing cultural, social, economic, environmental, and political domains beyond the academic arena (c.f. Greenhalgh et al., 2016; Reale et al., 2018; Pedersen et al., 2020). In the academic literature, several such *models or frameworks* (from here on models) to capture the SIR in various research domains have been developed (Greenhalgh et al., 2016). Essentially, such SIR models theorize how and why research contributes to societal change.

Invariably, the definitions of societal impacts differ along with their focus and underlying assumptions. Yet, SIR models offer valuable conceptual understandings of how research is supposed to bring about societal impact. Such models are often developed into *analytical approaches* (Pedersen et al., 2020) of

how SIR should be evaluated and measured, and in rare instances, *instruments* (Kramer et al., 2013; Sørensen et al., 2022) to measure and assess the SIR. Conceptualizations of SIR models are depicted in illustrative, narrative, or tabular formats.

Several reviews have provided insights as to the epistemological and methodological nature of these models, for example, dividing approaches up into positivist, constructive, realist, critical, and performative approaches (Greenhalgh et al., 2016) or describing the methods used for data collection, for example, interviews, surveys, expert reviews, statistics, altmetrics, and document analyses (Pedersen et al., 2020) in analytical approaches.

Despite these theoretical and methodological advancements, considerable ambiguity remains. As remarked by Greenhalgh and her colleagues:

> 'Different approaches to assessing research impact make different assumptions about the nature of research knowledge, the purpose of research, the definition of research quality, the role of values in research and its implementation, the mechanisms by which impact is achieved, and the implications for how impact is measured' (Greenhalgh et al. 2016:p.2).

As the above quote indicates, different SIR models inherently make *assumptions* about *research (knowledge)*, *how* researchers and research products (directly or indirectly) interact with a broader context of social actors, and *why* these interactions lead to societal changes. As such, these assumptions have far-reaching consequences. They are the very essence of the causal propositions of *how* and *why* research is believed to bring about societal impact (Weiss 1997; Nkwake 2013; Mertens 2016). In other words, SIR models are innately *theories* of how research brings about societal impact (Pawson and Tilley 1997). It is through explicating assumptions that one can develop plausible and cogent theorizing of

the interactions and impact pathways through which research is purported to bring about change can be developed.

This observation has implications for research program planners and evaluators. In short, assumptions clarify how and why research is expected to influence society. In doing so, it also brings to the fore the role of contextual factors. Explicit assumptions also surface critical weak links and risks in the SIR model (Chen 1990). Failing to develop a SIR model undermines the evaluability of the program and the testability of its causal claims (Janssens and de Wolf 2009). Finally, explicit assumptions support stakeholder dialogues and adaptive learning (Rogers 2008).

The practical implications are many. First, the unit of analysis may differ. It may be a research product (i.e. one or more scientific articles), a research programme (i.e. a concerted body of knowledge on a particular topic), or a research institution (the totality of activities and outputs produced from one organizational entity). Second, the depiction of research(ers)' interaction with society at large rests on different theories about social reality. Third, such assumptions may lead to a more or less explicated preference for different research designs and methods.

Therefore, it is salient and important to identify how such models are constructed and how theoretical assumptions about how research brings about societal impact are explicated in SIR models (Nkwake 2013). One may argue that such assumptions are the very foundation of such models.

In this article, we explore the theoretical assumptions ingrained and explicated in the current SIR models. We do so by posing the following research questions:

1) Which models for societal impact of research exist in the research literature?
2) How are these models constructed and depicted?
3) How are theoretical assumptions about societal impact explicated in these models?

As demonstrated above, and observed in other reviews (Greenhalgh et al., 2016; Pedersen et al., 2020) models of SIR draw from evaluation theory and methods. Before moving on with analysis of SIR, let us digress to this article's conceptual and methodological foundation in theory-based evaluation. Some further considerations are needed, when eliciting the underlying theoretical assumptions of SIR models.

1.1 Theory of change and societal impact of research models

As mentioned, when making assumptions about how research brings about societal impact, let alone set about to evaluate whether this was achieved—or not—one is invariably inscribed in a broader theoretical and methodological literature on theory-driven evaluation (aka theory-based evaluation). Even though the history of theory-driven evaluation stretches over more than half a century, conceptual ambiguity remains in this field (Coryn et al., 2011; Lemire et al., 2021). Therefore, we explicate some conceptual clarification.

At the heart of theory-driven evaluation lies the elucidation of how and why a planned change (i.e. a practice, project, program, or policy) is supposed to work. Carol Weiss famously distinguished *implementation theory* and *program (matic) theory* (1997:47). Implementation theory focuses on *how* an intervention is carried out (according to plan, quality, dosage, attainment of results, etc.). Program theory focuses on the mechanisms that intervene between the delivery of the intervention's activities and the outcome of interest. In other words, it focuses on *why* the intervention works and the underlying theoretical assumptions of these propositions. According to Coryn and his colleagues, such theoretical assumptions may be informed by existing social science theory (empirical and theoretical), implicit theory (stakeholders' theory), or emergent theory (based on observation of the intervention) (2010:203-05). Nkwake (2013) has distinguished between three main categories of assumptions that evaluators and stakeholders make; paradigmatic (worldview and conceptual), prescriptive (what actions should be taken), and causal (links between actions and results). In this article, we focus on the latter.

Within theory-driven evaluation, many different approaches have evolved, such as realist evaluation (Pawson and Tilley 1997) and contribution analysis (Mayne 2019), as well as a wide range of tools for elucidating theory from interventions such as logic models (Wyatt Knowlton and Phillips 2008) and outcomes frameworks (Taplin and Clark 2012). For an overview, see Lemire et al. (2023).

This article uses the label theory of change (ToC) as the overarching concept central to SIR models. We use *implementation theory* to describe the set of inputs, activities, outputs, and immediate, intermediate, and long-term outcomes in the models (i.e. the sequence of activities, outputs, and immediate outcomes in particular causal links). Essentially, implementation theory describes *how* a planned change is supposed to happen. We use the label *program theory* when we describe the assumptions about the causal links that explain *why* a planned change is supposed to work (within a broader configuration of contextual factors).

Drawing from Dhillon and Vaca (2018), one should distinguish between different components when constructing a theory of change: inputs (human, physical, technological, and financial resources), activities (actions taken), outputs (immediate product of the action), immediate outcomes (e.g. immediate outcome of training or service delivery on the recipient), intermediate outcomes (e.g. more distal behavioral changes of recipients), long-term outcomes (e.g. changes in the broader target population's condition), specific causal links (detailing how particular sets of activities and outputs bring about an outcome in a given context), assumptions (statements of if—then—because supporting each causal link) and mechanisms (recognizable causal patterns triggered under certain conditions) (Elster 1998:45).

Dhillon and Vaca posit that a *strong ToC*, and therefore a strong SIR model, must contain all components, thereby covering both the implementation theory and program theory aspects of a ToC. Drawing from dominant theory-driven approaches such as realist evaluation and contribution analysis, one may add that salient contextual factors assumed to mediate or moderate outcomes; potential unintended outcomes should also be identified (Pawson and Tilley 1997; Mayne 2017). Dhillon and Vaca observe that *standard ToCs* rarely include all components needed to form a strong ToC but only include inputs, activities, outputs, and immediate, intermediate, and long-term outcomes (Kellogg 2006). Figure 1 depicts these components.

Other proponents of theory-driven evaluation have emphasized that a strong (or robust) ToC must be structurally sound and plausible (herein measurable) (Mayne 2017). In a
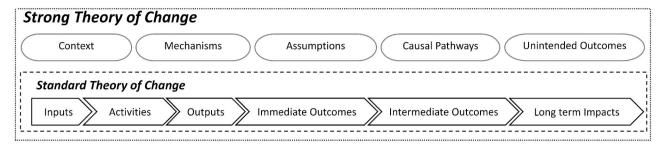
Figure 1. Components in a theory of change. This figure illustrates which components a theory of change should include and specifies five components that characterize strong theories of change.

similar vein, Kubisch and her colleagues argue that a ToC must be credible, feasible, and testable (Kubisch et al., 2010).

Generic models of how research brings about societal impact may be difficult to assess as plausible, feasible, and testable as this hinges on concrete applications. However, it is possible to determine whether the structural components for a strong ToC are in place. As such, we apply these structural components to conduct a componential analysis of the existing SIR models and reach a deeper understanding on how they were modelled.

1.2 Structure of the article

The article is structured as follows: First, we present the methodology used to identify the SIR models, the data extraction process, and the analytical strategy and techniques applied. Second, we present the findings. Third, we discuss the implications of our analysis for advancing theory and empirical research on the SIR. Finally, we summarize the findings and conclusions.

2. Methodology

In this section, we describe the methodology used in terms of sources and search strategy, coding and data analysis, and finally, the strengths and limitations of the chosen methodology.

2.1 Sources and search strategy

In recent years, many reviews have been published on the SIR, especially within the health domain. Therefore, we expect all major, original SIR models and frameworks to be mentioned in one or more of these reviews. This study is thus based on an umbrella review methodology. An umbrella review is the systematic collection and assessment of multiple (systematic) reviews on a particular research topic (Belbasis et al., 2022). The study aimed to identify, collate, analyse, and report the findings on these SIR models and frameworks.

We have not distinguished firmly between SIR *models* and SIR *frameworks*, as they seem to be used interchangeably in the literature. We conceive of SIR models and frameworks as specific and detailed ways to describe *causal links* and explain how research is *assumed* to bring about desired societal changes. The SIR model should be *empirically testable* in concrete applications. As such, we have distinguished SIR models from analytical approaches that prescribe how the societal impact of research can, and should be, evaluated against specific criteria, research designs, types of data, methods, etc.

We conceive of SIR frameworks as systematic approaches that help organize and guide research into how societal impact is created and documented, ensuring consistency and comprehensiveness. As our primary interest is in how to guide research institutions to improve societal impact, we have focused the literature search on SIR reviews that, at minimum, include the word 'framework.'

We searched for reviews on SIR models in the curated bibliographic database Web of Science (WoS) Core Collection and in the academic search engine Google Scholar, because it is broader and includes grey literature. We did not search other curated databases beyond WoS, as we considered it likely that all relevant scientific reviews for this purpose would be covered either in WoS or in Google Scholar. We used the following search terms: ('societal impact' (All Fields) or 'research impact' (All Fields)) and research (All Fields) and framework* (All Fields) and Review Article (Document Types). We limited the search to reviews from 2010 to 2024. We relied on 'framework' as an often-used synonym for 'model' and did not include 'model' as a search term because it is too broad and ambiguous.

We identified 326 reviews in WoS, of which 305 were excluded based on title, and 10 were excluded based on full-text assessment. The criteria for inclusion were that the article type is a review and that the review addresses models and frameworks concerning the societal impact of research. This process resulted in 11 eligible reviews. The search in Google Scholar yielded two more reviews that had not already been identified. The most recent and relevant SIR review identified was dated 2021.

Two reviewers screened titles and abstracts using Endnote 21 software. They sought consensus, and conflicts were resolved through discussions. Subsequently, one reviewer screened eligible full-text reviews, while the second reviewer screened excluded publications to validate exclusion decisions by reviewer one. Altogether, we identified 13 reviews that pertain to SIR models.

Six reviews focused on societal impact broadly (Grant et al., 2010; Penfield et al., 2014; Greenhalgh et al., 2016; Strahan et al., 2020; Razmgir et al., 2021; Smit and Hessels 2021), five on research impact within the health domain (Banzi et al., 2011; Milat et al., 2015; Raftery et al., 2016; Cruz Rivera et al., 2017; Deeming et al., 2017), and two on research impact within the social sciences and humanities (Reale et al., 2018; Pedersen et al., 2020).

Seven reviews mention assumptions made in the evaluated frameworks. A central argument is that frameworks rest on implicit or explicit assumptions concerning the nature of research, knowledge, and societal impact ranging from positivist to interpretative approaches (Raftery et al., 2016; Reale et al., 2018; Pedersen et al., 2020), and conceptions of impact pathways as linear or complex (Penfield et al., 2014; Smit and Hessels 2021). Additionally, three reviews highlight that

monetisation frameworks introduce assumptions related to time lags and the quantification of economic impact (Banzi et al., 2011; Greenhalgh et al., 2016; Raftery et al., 2016). Only three reviews address assumptions about impact pathways and solely in the context of frameworks incorporating theories of change or contribution mapping (Raftery et al., 2016; Pedersen et al., 2020; Smit and Hessels 2021). None of the reviews provides guidance on how to identify and operationalize such assumptions in practice.

We identified 80 references to original manuscripts about SIR models in these reviews. We retrieved the original manuscripts for further inspection. Two references were outdated and could not be retrieved. As new SIR models might have been published within the last five years, we conducted a hand search from 2019 to 2024 in leading journals on science research for more recent publications on SIR models using the search terms societal impact, theory of change, logic model, and programme theory. These journals included *Evidence and Policy, Health Research, Policy and Systems, Implementation Science, Journal of Health Services, Research and Policy, Research Evaluation,* and *Science Communication.* Our search identified four potentially relevant SIR models.

All retrieved manuscripts were included in NVivo 12 Pro for further analysis using file classification. Each manuscript was reviewed and screened. We excluded 60 of 84 articles based on seven exclusion criteria:

1) Not about societal impact (*n*=23);
2) SIR model or framework without causal links (*n*=6);
3) Practical cases or measurements, not a SIR model (*n*=11);
4) Performance evaluation and metrics, not a SIR model (*n*=9);
5) General report, not a SIR model (*n*=7);
6) Duplicate or variant of an existing model (*n*=1);
7) Out of scope (*n*=3).
8) In total, 24 distinct SIR models were identified. The process is depicted in the PRISMA diagram in Figure 2.

2.2 Data analysis

All manuscripts that were selected for further analysis in NVivo 12 Pro were categorized according to model characteristics using case classification (Table 1).

Some model properties could be derived from article characteristics in the file classification imported into NVivo from Endnote: article type (peer-reviewed article, report, PhD thesis, etc.), author affiliation (institution type), country of study, and type of research. The remaining information was extracted through document analysis.

2.3 Coding strategy and coding themes

The analysis combined both deductive and inductive coding. Deductive coding was guided by categories derived from existing reviews and theoretical articles, such as model components (Dhillon and Vaca 2018) and evaluation tools for devising a theory of change (Lemire et al., 2023). Inductive, open coding allowed emerging subcodes to be aggregated into thematic category codes. While most of these subcodes were framework-specific, manuscripts were recoded when subcodes were deemed to have broader applicability.

Manuscripts were coded in accordance with a coding manual. Two authors divided the manuscripts and conducted the initial coding independently, reading each manuscript in full.

Each author's coding was then validated by the other. To ensure consistency, the full author group met several times to resolve definitional issues and refine coding practices. Based on these discussions, the coding manual was revised and relevant manuscripts were revisited.

To increase transparency, the exact framework name used in each article was assigned as a subcode under the relevant theoretical category code. When a new category emerged, it was added to the coding manual. Furthermore, NVivo's search capabilities were used to mine manuscripts for keywords such as assumptions and related synonyms (such as propositions, hypotheses, suppositions, presumptions, and expectations) to reduce the risk of overlooking important points about assumptions.

2.4 Analytical strategy

For the data analysis, we applied different elements of the coding manual to address the various research questions. Table 2 presents an overview of how the coding elements were used.

2.5 Strengths and limitations

The chosen methodology's strengths are as follows: First, the review takes advantage of existing (systematic) reviews and thus saves significant resources on conducting exhaustive database searches. It presents an overview of the existing SIR models addressed in prior reviews, and the structured analyses cut across the reviews' different domains and purposes.

However, there are several disadvantages associated with umbrella reviews. For one, the validity hinges on the quality of the existing reviews. We have sought to address this apparent weakness by retrieving the original SIR model manuscripts. Second, umbrella reviews exclude data that are not reported in prior reviews. Recognizing these disadvantages, we conducted a hand search in leading science evaluation journals to identify SIR models published after the analyses conducted in the most recent review.

Another limitation is that this approach does not capture work from research groups that have not explicitly labelled their work as societal impact or as a framework, or that have not published in scientific journals. We are aware that some research groups have systematically applied theory of change and related approaches. Examples can be found in fields such as international humanitarian aid and agriculture—for instance, work connected to the CGIAR (Consultative Group on International Agricultural Research) programs (Mayne and Johnson 2015; Belcher et al., 2024). To the extent that we have encountered such research, it has informed the analyses presented in this article.

We believe that the current review is considerably more comprehensive than previous reviews because it covers many different domains. As such, it is the most comprehensive review of existing SIR models to date.

3. Findings

This section outlines the key findings from the review of the 24 SIR models. It is organized according to the three research questions.

Which SIR models for societal impact of research exist in the research literature?

The first step in our analysis was to examine the characteristics of the SIR models in the research literature.
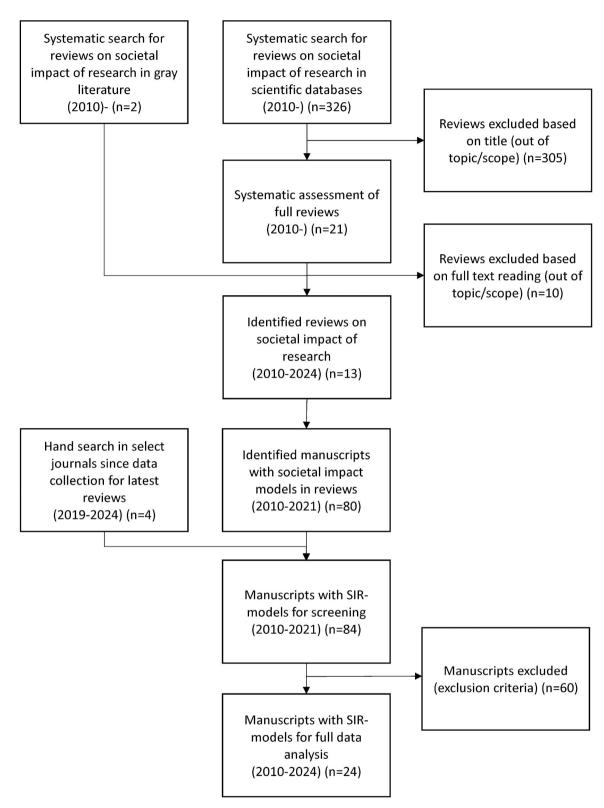
Figure 2. PRISMA diagram of search. The figure illustrates the flow of information through the different stages of the review process.

Table 3 illustrates that 16 of the 24 SIR models originate from the public health and clinical health domains, where most models' long-term impact is improved health and health care. The purpose of the models varies, but across domains, the common denominator is to assess SIR impact. The most frequent unit of analysis is at the project level (11). Five models address multiple levels, indicating that some models use a broad, integrative approach that examines societal impact across various levels. Institution level is found in five cases, reflecting their focus on the societal impact of research organizations. The community level is the least frequent, with three instances. The geographic origin of contributions shows

that an overwhelming number of the SIR models originate in Anglo-Saxon countries: six from the UK and five from each of Australia, Canada, and the USA. Continental Europe had the fewest contributions (3). Most SIR models were developed in a university setting, sometimes in collaboration with research centres or the government.

Most long-term impacts described by the selected 24 SIR models are specific to the domains in which they were conceived (Appendix 1 for details). The nine models from public health specify final impacts, such as improved health, better public health, enhanced health interventions, excellence in health care, and improved economy (Searles et al., 2016; Buxton and Hanney 1996; Kok and Schuit 2012; Cohen et al., 2015; Scott et al., 2014; Graham et al., 2012; Guinea et al., 2015; Cacari-Stone et al., 2014; Buckwalter et al., 2017). The five models from health science emphasize long-term impacts, such as improved health, more effective health services and products, and social and economic benefits (Leone et al., 2017; Bernstein et al., 2006; Logan and Graham 1998; CAHS 2009; Redman et al., 2015). Four articles within health care and occupational safety and health focus on health, improved primary health care, health-related societal changes,

reduced work-related injuries and illnesses, and research uptake in policy and practice (Morton 2015; Harper et al., 2020; Van Eerd et al., 2021; Williams et al., 2009). The six articles within agricultural research, environmental health, political science, and social science also focus on long-term impacts relevant to their respective domains (Young et al., 2014; Meagher et al., 2008; Engel-Cox et al., 2008; Joly et al., 2015; Phipps et al., 2016; Duryea and Parfitt 2007).

3.1 Summary of existing models

The SIR models cover a broad range of purposes and long-term impacts. Most models originate within the health domain and come from Anglo-Saxon countries. They address different impact levels, most frequently the project level, but almost as many address other levels. It is least common to address the community level.

The primary sources of the selected scientific articles about SIR models were academic institutions. Other sources may be more prevalent in books, grey literature, and practical guidelines, but they were not considered because we focus on academically vetted knowledge.

4. How are these models constructed and depicted?

We further investigated how the SIR models are constructed and depicted. Table 4 shows that the most common format is a figure alone, appearing in 16 instances, suggesting a preference for visual representation. Figures are most often used in logic models. Tables and the combination of tables with figures are used less frequently, appearing only six times. The use of narratives is only found in two SIR models.

Table 4 also illustrates which evaluation tools are utilized within the SIR models and the frequency of their application. The most commonly used tool is the logic model (typically in the style of the Kellogg Foundation Logic Model) (Kellogg 2006), appearing 12 times, which significantly outnumbers the other tools listed. The nested and hybrid model is the second most frequent, with five occurrences. Two are syntheses of the Kellogg Foundation Logic Model and the WHO and Payback Model, respectively. Other tools are used less frequently.

As documented above, the fact that the models are depicted as figures that apply elements of the Kellogg Foundation

Table 1. Case classifications.

Case classification	Example
SIR Model name	FAIT, CAHS, OMRU …
Model purpose	Inform policy, improve health …
Unit of analysis	Intervention, Programme, or Institution level
Long-term impact	Better health, Improved safety …
Domain	Agriculture, Public Health …
Model type	Original, Synthesis, Adaptation …
Model depiction	Tabular, Figure, Narrative, Combined …
Evaluation tool used	Logic Model; Log Frame; Concept Map; Outcome Framework, Theory of Change[a] …
ToC components	Inputs; Activities; Outputs; Outcomes; Assumptions; Mechanisms; Context …

We used Lemire and colleagues' categorization of models (2023). In their typology, they use theory of change as specific term for a particular model. They use *program theory* as the umbrella term for different models, whereas we use *theory of change*. We further added outcomes framework and impact plan not covered by their publication.

  a   In this context, ToC is considered an evaluation tool and not an overarching concept, as we do in the rest of the article.

Table 2. Analytical strategy for each research question.

Research question	Thematic category codes	Analytical approach
1) **Which models for societal impact of research exist in the research literature?**	• SIR Model name • Model purpose • Long-term impact • Domain • Author affiliation • Country of study	• Descriptive analysis of the identified SIR models
1) **How are these models constructed and depicted?**	• Model type • Unit of analysis • Model depiction • Theory of change components included • Evaluation tool used	• Descriptive analysis of how the SIR models are constructed and which components are included • Cross-tabulation of type versus evaluation tool
3. **How are theoretical assumptions about societal impact explicated in these models?**	• Model depiction • Evaluation tool used • Theory of change components included	• Descriptive analyses • Cross-case analysis of depiction, evaluation tool, and assumptions

Logic Model as an evaluation tool aligns well with the components of the models. Logic models are typically arranged in boxes with components from input, activities, outputs, and outcomes. These are the components that the majority of the SIR models contain. Inputs and outputs are only present in about half of the SIR models using logic models.

Finally, we have analysed how many of the SIR models include the elements of the standard ToC. None of the 24 models had all six standard ToC elements, while 20 had four or five elements. Table 5 illustrates that especially inputs and outputs were not described. Very few SIR models included elements of a strong ToC: 10 described context, nine explicitly described causal links, four described mechanisms, two explicitly described how to work with assumptions of impact pathways in the model, and none described unintended outcomes. Table 6 lists the five SIR models that included most components characterizing a strong ToC (3). It is notable that they include either mechanisms or assumptions.

4.1 Summary of how models are constructed and depicted

Our findings reveal that most SIR models lack the components Dhillon and Vaca (2018) address as essential for a strong ToC. They only comprise the implementation theory, which describes the set of inputs, activities, outputs, and immediate, intermediate, and long-term outcomes in the models. Thus, program theory components, such as assumptions, causal links, alternative outcomes, mechanisms, and context, are only included in a few SIR models.

Table 3. Summary of domain, level, and source of SIR models.

Category	Value	Number
Domain	Public health	11
	Clinical health	5
	Generic	2
	Other (occupational safety and health, social and political science, agriculture)	6
Level	Project level	11
	Institution level	5
	Community level	3
	Multiple levels	5
Country	UK	6
	Australia	5
	Canada	5
	USA	5
	Continental Europe (Spain, France, and the Netherlands)	3

5. How are theoretical assumptions about societal impact explicated in these models?

Lastly, we examined how theoretical assumptions about societal impact are explicated in the SIR models to provide inspiration for model developers and revisers. Our analysis showed that only two models, the Research Contribution Framework (RCF) (Morton 2015) and the Spirit Action Framework (SAF) (Redman et al., 2015), include theoretical assumptions in their SIR models about *why* research contributes to societal change. While several models stress the importance of stating assumptions about research impact, they neither incorporate these assumptions in their SIR model nor provide guidance on addressing them (e.g. Young et al., 2014). As assumptions are a key element of a robust SIR model, we elaborate on research question 3 by analyzing these two frameworks to illustrate how assumptions can be incorporated and explicated in practice.

Below, we present how these two SIR models present assumptions. We provide examples, elaborate on the types and origins of these assumptions (whether they stem from social science, stakeholders, or emergent theories), and describe how the authors suggest one can work with them in practice. The two SIR models are both situated within the health domain. They are based on empirical and theoretical foundations, using research interventions, policies, or programs as units of analysis.

5.1 The research contribution framework

As the name implies, the framework aims to investigate and demonstrate how research *contributes* to policy-making and practice. Drawing on literature about evaluation, research utilization, and empirical studies and using Contribution Mapping as a theory-based approach, it traces and evaluates the pathways through which research influences policy and practice. The model emphasizes the processes and ways research is adopted and used to identify its contributions.

Using the logic model as a tool, illustrated in figure and table formats, the RCF provides a detailed guide on identifying, explicating, and testing theoretical assumptions at each step in the causal chain, from research inputs to societal impact. Throughout this process, the authors argue that one should consider the influence of context and potential risks if the assumptions do not hold. A similar approach has been used in other fields (Belcher et al., 2024).

5.2 The spirit action framework

While the CRF solely aims to measure how research contributes to impact, the SAF model is, as the name suggests, an action framework. It also aims to *enhance* research use in

Table 4. Relation between evaluation tool and SIR model format.

	Logic model	Outcome framework	Nested and Hybrid Model	Stock and Flow diagram	Programme theory	Theory of change	Total
Figure	1,8,10,13,14, 17,18,21,23,24		4,5,22	9,12,20			16
Table	2,11						2
Figure and table		3	7		6,15		4
Figure and narrative[a]			16				1
Narrative[a]						19	1
Total	12	1	5	3	2	1	24

a　Narrative is interpreted as extensive narration of model content.

Table 5. Components in the SIR models compared to the type of evaluation tool used.

	Model components	Logic model	Outcome framework	Nested and hybrid model	Stock and flow diagram	Programme theory	Theory of change	*Total*
Standard ToC	01. Inputs	6	1	3	2	2	0	*14*
	02. Activities	11	1	3	3	1	1	*20*
	03. Outputs	4	0	2	0	0	1	*7*
	04. Immediate outcomes	8	1	2	1	2	0	*14*
	05. Intermediate outcomes	12	1	3	2	2	1	*21*
	06. Long term impacts	12	1	4	3	2	1	*23*
Strong ToC	07. Unintended outcomes	0	0	0	0	0	0	*0*
	08. Causal links (explicit)	6	0	2	0	1	0	*9*
	09. Assumptions	1	0	0	0	1	0	*2*
	10. Mechanism	2	0	2	0	0	0	*4*
	11. Context	6	0	1	1	1	1	*10*
Other	12. Vision-purpose	4	0	1	1	0	0	*6*
	13. Other components	9	1	1	0	0	0	*11*
	Total (unique)	*12*	*1*	*5*	*3*	*2*	*1*	*24*

Table 6. SIR models with a strong ToC.

ID	Model name	Strong ToC components
10	Melbourne Children's Knowledge Translation and Research Impact Framework	Causal links, mechanisms, context
14	The SPIRIT Action Framework	Causal links, assumptions, context
15	The Research Contribution Framework	Causal links, assumptions, context
21	Conceptual Model of Comprehensive Research Metrics for Improved Human Health and Environment	Causal links, mechanisms, context
22	ASIRPA	Causal links, mechanisms, context

policy-making by identifying and testing strategies that increase research utilization.

The SAF model is based on findings from a literature review and interviews combined with knowledge translation and implementation science theories. It was developed by a team of key stakeholders, including policymakers, researchers, and knowledge exchange specialists. According to the authors, the model can '*create testable hypotheses about the drivers of evidence use, causal pathways, and intervention strategies that are likely to be effective*' (Redman et al., 2015:152). The authors use the notion of hypotheses instead of assumptions, but we have interpreted it as essentially the same, as the hypotheses describe why the causal pathways work as they do.

The model employs a logic model as a tool, illustrated in a figure. The logic model should be used as a dynamic tool, meaning that if the assumptions are not confirmed, the logic model should be adjusted. The SAF model provides examples of practical application, such as illustrating hypothetical scenarios and how to identify intervention strategies and causal links. Additionally, it includes an overview of key concepts, measurement levels, and tools.

5.3 Examples of assumptions

Table 7 provides examples of how the two models incorporate assumptions, taking Dhillon and Vaca's suggestions for components needed for a strong Theory of Change (ToC) at the outset. As shown, the models incorporate assumptions for each component in the standard ToC. In the RCF, impact pathways can only be deduced from the assumptions leading from one step in the ToC to the next. In the SAF, the assumptions are associated with hypotheses about how four different

catalysts can lead to more research-informed policy and practice and, subsequently, better health services and programs.

The RCF's assumptions concern the relevance of the research to policy or practice, the capacity of stakeholders to engage with and use research, and the adaptability of the research to specific contexts. The model depicts in a table how assumptions are important to progress toward contributing to increased capacity, changed behaviour, and improvements in intended outcomes (societal impact). The model also suggests indicators that document such a progression.

The assumptions in the SAF concern what is needed for a catalyst to initiate the use of research. The target group's capacity to engage with research mediates its response to this catalyst. If an organization has sufficient capacity, research engagement actions will likely occur, facilitating research use.

Both models emphasize the importance of context. For example, they assume that research is more likely to be used if it is timely and relevant to the user's needs and *if* it fits with their current thinking. In several cases, particularly in the RCF, facilitating contextual factors are integrated into the descriptions of assumptions related to each component. For example, that research is timely and relevant.

Furthermore, both models stress the importance of considering not only the intended (positive) assumptions about the causal links from research input to contribute to societal change. The RCF argues that addressing the potential risks at each step in the causal chain and the unintended outcomes that research may produce is equally important. As such, the RCF model incorporates potential risks into its logic model alongside the assumptions of various causal links.

However, while assumptions for each component are described, an interesting finding is that the assumptions mainly concern the implementation theory (how research brings

Table 7. Types of assumptions explicated in the CRF and SAF.

Components	Examples of assumptions about societal impact explicated in the models
Inputs	Intended audiences perceive research knowledge as valuable and useful, e.g. organizations and staff in political organizations.
Activities	Activities delivered and implemented as intended. Activities aligned with the interests and needs of involved stakeholders.
Outputs	The researchers know and can reach the intended audiences. Medias do not distort research content.
Immediate outcomes	The intended audiences access the research communication and receive it as intended. The timing is right and relevant to the current needs of the intended receivers. Intended audiences appraise the produced research knowledge and interact with researchers.
Intermediate outcomes	Research findings are perceived as useful and relevant. Research is adapted and used for instrumental, tactical, conceptual, and/or imposed purposes, leading to behavioral change.
Long-term impacts	Research 'fits' with current thinking and is integrated with other knowledge. Research has contributed to informed health policies, policy documents, and better health systems and outcomes.
Context	Research is only used if it is relevant and timely to the intended audience's needs, interests, political context, public opinion, economic climate, etc. If the intended audiences do not have the capacity, skills, tools, and systems to support the research engagement actions and the use of research, they will most likely not be used.

about societal impact). As such, the models do not explain their assumptions about the programme theory, i.e. why the research brings about societal impact. In other words, what links the components together—the causal links—is missing. For example, an assumption in the SAF model is that the research is perceived as valuable and useful, but what it takes to make it so is not explained.

6. Where do the assumptions originate from?

The assumptions (and risks) in the RCF model are identified based on findings from a mix of existing literature on evaluation, research utilization, and an empirical study of research impact. In the SAF, assumptions are developed based on a literature review and input from key stakeholders.

Neither of the models clearly states what exactly underlies which assumptions and why they are included in the models. This is primarily described in the articles' methodology, so they are unrelated to each causal link or causal effect proposition.

7. How do the models describe how to work with assumptions?

The RCF proposes to create a model that illustrates the expected pathway(s) to societal impact. Subsequently, one should identify the common assumptions and risks associated with the activities and progression between stages involved in creating impact. Subsequently, one should identify indicators for progress. The indicators may lead to modifications of assumptions, and the combination of assumptions and indicators makes it easier to determine if a contribution has been made. Whether such assumptions and indicators should be theoretically or empirically driven is not explicitly expressed. While the project unfolds, the framework suggests annotating and revising the impact pathway model with new insights. Finally, a table listing pathways and their evidence can be produced.

Whether the assumptions or risks can be confirmed should be empirically assessed through suitable indicators. For example, indicators to validate such theoretical assumptions or risks can be whether the research project and activities are carried out as intended and measures of the audience's reaction and participation. Indicators of intermediary outcomes could be the level of stakeholders' awareness and response, participation, and engagement. It could also be research users' perspectives on usefulness, relevance and timing, research engagement, and use in policy and practice. The long-term outcomes may be built on population-level data, including how many were affected by the research.

7.1 Summary of how theoretical assumptions are explicated

Our analyses show that only two SIR models explicitly state assumptions about how research brings about societal impact. However, two interesting findings emerge when investigating the types of assumptions being made. First, the assumptions only describe *how* research is assumed to bring about change (implementation theory). Assumptions about the causal links in the models that explain *why* the research is supposed to create change are absent. Second, it is unclear from *where* each assumption originates, i.e. whether they are empirical, theoretical, or based on stakeholder input. There is only a rudimentary discussion of how such assumptions can be developed and described.

8. Discussion

The review of 24 SIR models highlights both the breadth of conceptual approaches and notable gaps in theoretical robustness. While models span diverse domains, most originate in health-related fields and are constructed primarily with logic models focusing on implementation theory components—inputs, activities, and outcomes—rather than the program theory explaining why change occurs. Only two models explicitly integrate assumptions, and even these address the "how" more than the "why" of societal impact. Mechanisms, unintended outcomes, and contextual configurations are often absent or underdeveloped, limiting the plausibility and testability of causal propositions as seen from a theory-based evaluation perspective. This uneven theoretical grounding underscores the need for greater conceptual clarity, explicit articulation of assumptions, and attention to contextual and cultural variability. These findings set the stage for a deeper examination of the conceptual, methodological, and theoretical issues in SIR modelling, as explored in the discussion, where we address (1) cultural dependencies, (2) definitional precision, and (3) the imperative for rigorous causal theorizing. Finally, (4) we propose a research agenda to advance theory development and empirical research on SIR.

8.1 Cultural and domain dependencies

A large share of the SIR models described in the current study are conceived in the health domain and Anglo-Saxon countries. This may not be a problem in and of itself, but it points to a need for cautious and judicious analysis of these models' generic properties. Whilst generic, exhaustive SIR models may scaffold specific application, they are obviously limited in their specificity. Surely, models that cover other domains and cultural and systemic contexts are called for. The accumulation and combination of generic (nomethetic) and case-specific (idiographic) SIR models may help advancing the both the science and use of SIR models.

8.2 A need for conceptual clarity

As shown in the methodology and findings sections, concepts are widely obfuscated. This umbrella review reveals a need for further conceptual clarity. First, one must distinguish between evaluation tools for thinking, such as logical models or outcomes frameworks, and substantive empirical models of how research affects societal impact. The latter may be derived using evaluation tools but should be testable and subject to empirical analysis in concrete applications.

Further definitional clarity is also needed in defining what constitutes a SIR model, what analytical approaches can be used to test it, and what indicators and instruments can be used to measure its components. This lack of conceptual clarity also implies that what is discussed as SIR models in the extant literature becomes unnecessarily confusing.

Based on our analysis, we propose the following definitions:

> A *societal impact of research model* is a substantive theoretical model that describes how, through which pathways, and why research contributes to societal change. The model may be generic and set forward a set of hypotheses about causality that can be adapted to analyse concrete cases and become empirically testable.

To describe the SIR model, researchers may use evaluation tools such as a theory of change (including logic models, outcomes frameworks, and others) to describe its components, their interrelationships, and the assumptions underlying it.

To empirically test the SIR model, researchers may use *analytical approaches*, which are structured approaches to *gathering, analysing, and evaluating* data. The analytical approach guides the selection of research questions, the development of the assessment methodology and data analysis plan, and the selection of indicators.

Impact indicators provide a set of measurements or values for constructs in the societal impact models, such as identified activities, outputs, outcomes, or impacts. Researchers may operationalize such indicators using different data collection methods and instruments.

Methods and Instruments are tools that researchers can use to *operationalize* one or more elements from the analytical framework and to *obtain, measure*, and *analyse* data.

8.3 A need for rigor when devising SIR models

An established SIR model is essentially a set of causal assumptions that propose how and why research contributes to change. The assumptions about such causal links are essential for the model's usability (Weiss 1997; Rogers 2008; Coryn et al., 2011). Our umbrella review identified the widespread use of logic models (or hybrids) to describe the model's implementation theory. However, we only identified two SIR models with explicit assumptions about how research brings about the impact (the program theory), and even in these articles, the 'why' was only vaguely explained.

Many models have implicit assumptions associated with impact pathways (the 'if … then') or in other parts of the models. Several articles mention that assumptions, hypotheses, or the like are important (e.g. Kellogg 2006; Young et al., 2014; Harper et al., 2020). However, assumptions need to be explicit for the theory to be plausible and empirically testable (Weiss, 1997). The old proverb among evaluators when using theories of change to describe change, often citing Sidney Harris' cartoon, ' … and then a miracle occurs … ', rings uncomfortably true to the current state of SIR theorizing (Archibald et al., 2016). The authors behind the ASIRPA framework make a similar point about the lack of descriptions of mechanisms, which we can only agree on as we only found four SIR models that describe mechanisms (Joly et al., 2015)

There is an urgent need for more attention to underlying (theoretical) assumptions in why research is supposed to bring about societal impact and through which impact pathways and contexts such changes occur. These assumptions (and associated mechanisms) should be explicitly stated, for example, by using 'if, then, because' statements as recommended by Dhillon and Vaca (2018) and by making explicit whether such assumptions derive from social science theory (empirical and theoretical), implicit theory (i.e. stakeholders' theory) or emergent theory (i.e. from data collected in concrete cases). Also, making assumptions explicit forces considerations about unintended consequences, risks, and, invariably, ethics. Indeed, more rigor in thinking is needed. There are places to find inspiration, such as in learning from concrete cases (e.g. Rishma et al., 2018), educational material (e.g. Funnell 2011), practical guides (e.g. Guijt 2013), and recommendations from funders (e.g. Christensen 2024).

8.4 A proposed research agenda to advance theory development and empirical research on SIR

Suppose the scientific community is to respond to funders' and other stakeholders' calls for research to create societal impact. In that case, a sound understanding of when, how, and why such changes are brought about is needed. This requires both theory development and empirical research.

As stated above, more rigor needs to be put into theorizing why and how research contributes to societal impact. A first call to action is to make explicit the assumptions underlying different SIR conceptualizations. Clarity is needed for each assumption made insofar as its origin in social science research, stakeholder perceptions, or emergent theory.

Second, one avenue to accommodate may be to identify and label generic impact pathways and how they are connected through social agents in wider systems (Rasmussen et al., 2024). By unveiling impact pathways' underlying assumptions and systematically testing these assumptions in empirical studies, we can build our knowledge base about how, when, and why these pathways work.

Third, depending on the theoretical orientation, one may identify and label configurations of contextual conditions wherein research is introduced and brings about societal change. Evidence use is complex (Boaz and Nutley 2019), and we need to know the conditions for research use, let

alone societal change. In sum, such approaches focusing also on context may aid both generative and configurational assumptions about how research contributes to societal change.

In doing so, we also echo Morton et al., (2015) call for identifying unintended consequences and risks for the societal impact of research. Not all uses may be intended, let alone legitimate.

Put together, an increased focus on the underlying assumptions and theorizing, by making such theoretical assumptions explicit, will help researchers and policymakers alike to understand the conditions necessary for research to have an impact, understand the pathways, and refine strategies to maximize societal benefits and enhance the robustness and credibility of research impact assessments/SIR models.

9. Conclusion

This article identified and surveyed 24 societal impact of research models through an umbrella review. The majority of these models were developed in the health domains. The SIR models were mapped based on the theory of change components included and the evaluation tools applied. The study found that logic models were dominant. Only six models had explicit causal links, and only two made explicit assumptions about why research contributes to societal change. Further theorizing and conceptual clarity are needed to advance the science of research impact.

Supplementary data

Supplementary data is available at *Research Evaluation Journal* online.

Funding

Conflict of Interest

There are no conflicting interests.

References

Aksnes, D. W., Langfeldt, L., and Wouters, P. (2019) 'Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories', *Sage Open*, 9: 1–17. https://doi.org/10.1177/2158244019829575

Archibald, T. et al. (2016) 'Assumptions, Conjectures, and Other Miracles: The Application of Evaluative Thinking to Theory of Change Models in Community Development', *Evaluation and Program Planning*, 59: 119–27. https://doi.org/10.1016/j.evalprogplan.2016.05.015.

Banzi, R. et al. (2011) 'Conceptual Frameworks and Empirical Approaches Used to Assess the Impact of Health Research: An Overview of Reviews', *Health Research Policy and Systems*, 9: 10. https://doi.org/10.1186/1478-4505-9-26.

Belbasis, L., Bellou, V., and Ioannidis, J. P. A. (2022) 'Conducting Umbrella Reviews', *BMJ Med*, 1: e000071. https://doi.org/10.1136/bmjmed-2021-000071.

Belcher, B. M., Bonaiuti, E., and Thiele, G. (2024) 'Applying Theory of Change in Research Program Planning: Lessons from CGIAR', *Environmental Science & Policy*, 160: 103850. https://doi.org/10.1016/j.envsci.2024.103850.

Bernstein, A. et al. (2006) 'A framework to measure the impact of investments in health research'. Ottawa.

Boaz, A., and Nutley, S. (2019) 'Using Evidence'. In: *What Works Now? Evidence-Informed Policy and Practice*, pp. 251-77. Bristol, UK: Policy Press.

Buckwalter, K. C. et al. (2017) 'Iowa Model of Evidence-Based Practice: Revisions and Validation', *Worldviews on Evidence-Based Nursing*, 14: 175–82.3. https://doi.org/10.1111/wvn.12223.

Buxton, M., and Hanney, S. (1996) 'How Can Payback from Health Services Research be Assessed?', *J Health Serv Res Policy*, 1: 35–43. https://doi.org/10.1177/135581969600100107.

Cacari-Stone, L. et al. (2014) 'The Promise of Community-Based Participatory Research for Health Equity: A Conceptual Model for Bridging Evidence with Policy', *American Journal of Public Health*, 104: 1615–23. https://doi.org/10.2105/Ajph.2014.301961.

CAHS (2009) '*Making an Impact. A Preferred Framework and Indicators to Measure Returns on Investment in Health Research*'. Ottawa: Canadian Academy of Health Sciences.

Chen, H.-T. (1990) *Theory-Driven Evaluations*. Newbury, CA: SAGE Publications.

Christensen, T. A. (2024) '*Theory-of-Change Used in Research-for-Impact-in-Society*'. Denmark: Novo Nordisk Foundation.

Coara (2022) *Coalition for Advancing Research Assessment: Agreement on Reforming Research Assessment*. July 20th, 2022. https://www.coara.org/agreement/the-agreement-full-text/, accessed 12 Dec. 2025.

Cohen, G. et al. (2015) 'Does Health Intervention Research Have Real World Policy and Practice Impacts: testing a New Impact Assessment Tool', *Health Research Policy and Systems*, 13: 3. https://doi.org/10.1186/1478-4505-13-3.

Coryn, C. L. S. et al. (2011) 'A Systematic Review of Theory-Driven Evaluation Practice from 1990 to 2009', *American Journal of Evaluation*, 32: 199–226. https://doi.org/10.1177/1098214010389321.

Cruz Rivera, S. et al. (2017) 'Assessing the Impact of Healthcare Research: A Systematic Review of Methodological Frameworks', *Plos Medicine*, 14: e1002370. https://doi.org/10.1371/journal.pmed.1002370.

D'Este, P. et al. (2018) 'How Do Researchers Generate Scientific and Societal Impacts? Toward an Analytical and Operational Framework', *Science and Public Policy*, 45: 752–63. https://doi.org/10.1093/scipol/scy023.

Deeming, S. et al. (2017) 'Measuring Research Impact in Australia's Medical Research Institutes: A Scoping Literature Review of the Objectives for and an Assessment of the Capabilities of Research Impact Assessment Frameworks', *Health Research Policy and Systems*, 15: 13. https://doi.org/10.1186/s12961-017-0180-1.

Dhillon, L., and Vaca, S. (2018) 'Refining Theories of Change', *Journal of MultiDisciplinary Evaluation*, 14: 64–87. https://doi.org/10.56645/jmde.v14i30.496.

Duryea, M.,M.,H., and Parfitt, A. (2007) 'Measuring the impact of research'.

Elster, J. (1998) 'A Plea for Mechanisms'. In: *Social Mechanisms an Analytical Approach to Social Theory*, pp. 45-73. Cambridge, UK: Cambridge University Press.

Engel-Cox, J. A. et al. (2008) 'Conceptual Model of Comprehensive Research Metrics for Improved Human Health and Environment', *Environmental Health Perspectives*, 116: 583–92. https://doi.org/10.1289/ehp.10925.

Funnell, S. (2011) *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. U.S.: Jossey-Bass Publishers.

Graham, K. E. R. et al. (2012) 'Evaluating Health Research Impact: Development and Implementation of the Alberta Innovates—Health Solutions Impact Framework', *Research Evaluation*, 21: 354–67. https://doi.org/10.1093/reseval/rvs027.

Grant, J. et al. (2010) 'Capturing Research Impacts A review of international practice'. RAND Europe.

Greenhalgh, T. et al. (2016) 'Research Impact: A Narrative Review', *BMC Med*, 14: 78. https://doi.org/10.1186/s12916-016-0620-8.

Guijt, I. (2013) '*ToC Reflection Notes 3: Working with Assumptions in a Theory of Change Process'*. The Hague: Hivos.

Guinea, J. et al. (2015) 'Impact Oriented Monitoring: A New Methodology for Monitoring and Evaluation of International Public Health Research Projects', *Research Evaluation*, 24: 131–45. https://doi.org/10.1093/reseval/rvu034.

Harper, L. M., Maden, M., and Dickson, R. (2020) 'Across Five Levels: The Evidence of Impact Model', *Evaluation (London, England 1995)*, 26: 350–66. https://doi.org/10.1177/1356389019850844.

Janssens, F. J. G., and de Wolf, I. F. (2009) 'Analyzing the Assumptions of a Policy Program an Ex-Ante Evaluation of "Educational Governance" in The Netherlands', *American Journal of Evaluation*, 30: 411–25. https://doi.org/10.1177/1098214009341016.

Joly, P. B. et al. (2015) 'ASIRPA: A Comprehensive Theory-Based Approach to Assessing the Societal Impacts of a Research Organization', *Research Evaluation*, 24: 440–53. https://doi.org/10.1093/reseval/rvv015.

Kellogg, W. (2006) '*WK Kellogg Foundation Logic Model Development Guide'*. East Battle Creek, MI: WK Kellogg Foundation.

Kok, M. O., and Schuit, A. J. (2012) 'Contribution Mapping: A Method for Mapping the Contribution of Research to Enhance Its Impact', *Health Research Policy and Systems*, 10: 21. https://doi.org/10.1186/1478-4505-10-21.

Kramer, D. M. et al. (2013) 'Did You Have an Impact? A Theory-Based Method for Planning and Evaluating Knowledge-Transfer and Exchange Activities in Occupational Health and Safety', *International Journal of Occupational Safety and Ergonomics*, 19: 41–62. https://doi.org/Doi https://doi.org/10.1080/10803548.2013.11076965

Kubisch, A. C. et al. (2010) *Voices from the Field III: Lessons and Challenges from Two Decades of Community Change Efforts*. Washington, D.C.: Aspen Institute.

Kuruvilla, S. et al. (2006) 'Describing the Impact of Health Research: A Research Impact Framework', *BMC Health Serv Res*, 6: 134. https://doi.org/10.1186/1472-6963-6-134.

Lemire, S., Christie, C. A., and Nielsen, S. B. (2021) '*Mending the Theory Gap in Evaluation*'. In: M Palenberg, A Paulson (eds.) *The Realpolitik of Evaluation Why Demand and Supply Rarely Intersect*, pp. 136–50. London: Routledge.

Lemire, S., Porowski, A., and Mumma, K. (2023) '*How We Model Matters: Visualizing Program Theories*'. Rockville, MD: Abt Global.

Leone, V., Modica, L., and West, S. (2017) 'The Melbourne Children's Knowledge Translation and Research Impact Project: Final Report'. https://doi.org/10.25374/MCRI.5440228.v1.

Logan, J., and Graham, I. D. (1998) 'Toward a Comprehensive Interdisciplinary Model of Health Care Research Use', *Science Communication*, 20: 227–46. https://doi.org/10.1177/1075547098020002004

Mayne, J. (2017) 'Theory of Change Analysis: Building Robust Theories of Change', *Canadian Journal of Program Evaluation*, 32: 155–73. https://doi.org/10.3138/cjpe.31122.

Mayne, J. (2019) 'Revisiting Contribution Analysis', *Canadian Journal of Program Evaluation*, 34: 171–91. https://doi.org/10.3138/cjpe.68004.

Mayne, J., and Johnson, N. (2015) 'Using Theories of Change in the CGIAR Research Program on Agriculture for Nutrition and Health', *Evaluation*, 21: 407–28. https://doi.org/10.1177/1356389015605198.

Meagher, L., Lyall, C., and Nutley, S. (2008) 'Flows of Knowledge, Expertise and Influence: A Method for Assessing Policy and Practice Impacts from Social Science Research', *Research Evaluation*, 17: 163–73. https://doi.org/10.3152/095820208x331720.

Mertens, D. M. (2016) 'Assumptions at the Philosophical and Programmatic Levels in Evaluation', *Evaluation and Program Planning*, 59: 102–8. https://doi.org/10.1016/j.evalprogplan.2016.05.010.

Milat, A. J., Bauman, A. E., and Redman, S. (2015) 'A Narrative Review of Research Impact Assessment Models and Methods', *Health Research Policy and Systems*, 13: 7. https://doi.org/10.1186/s12961-015-0003-1.

Morton, S. (2015) 'Progressing Research Impact Assessment: A 'Contributions' Approach', *Research Evaluation*, 24: 405–19. https://doi.org/10.1093/reseval/rvv016.

Nkwake, A. M. (2013) *Working with Assumptions in International Development Program Evaluation*. New York: Springer.

Pawson, R., and Tilley, N. (1997) *Realistic Evaluation*. London: SAGE Publications.

Pedersen, D. B., Grønvad, J. F., and Hvidtfeldt, R. (2020) 'Methods for Mapping the Impact of Social Sciences and Humanities-A Literature Review', *Research Evaluation*, 29: 4–21. https://doi.org/10.1093/reseval/rvz033.

Penfield, T. et al. (2014) 'Assessment, Evaluations, and Definitions of Research Impact: A Review', *Research Evaluation*, 23: 21–32. https://doi.org/10.1093/reseval/rvt021.

Phipps, D. et al. (2016) 'The Co-Produced Pathway to Impact Describes Knowledge Mobilization Processes', *Journal of Community Engagement and Scholarship*, 9: 10. https://doi.org/10.54656/gokh9495.

Raftery, J. et al. (2016) 'Models and Applications for Measuring the Impact of Health Research: update of a Systematic Review for the Health Technology Assessment Programme', *Health Technology Assessment*, 20: 1–254. https://doi.org/10.3310/hta20760.

Rasmussen, C. D. N., Nielsen, S. B., and Sørensen, O. H. (2024) 'Policy Utilisation of Occupational Safety and Health Research: Results from a Tripartite Unicameral Parliamentary System in Denmark', *Evidence & Policy*, 20: 460–85. https://doi.org/10.1332/17442648y2024d000000025.

Razmgir, M. et al. (2021) 'Exploring Research Impact Models: A Systematic Scoping Review', *Research Evaluation*, 30: 443–57. https://doi.org/10.1093/reseval/rvab009.

Reale, E. et al. (2018) 'A Review of Literature on Evaluating the Scientific, Social and Political Impact of Social Sciences and Humanities Research', *Research Evaluation*, 27: 298–308. https://doi.org/10.1093/reseval/rvx025.

Redman, S. et al. (2015) 'The SPIRIT Action Framework: A Structured Approach to Selecting and Testing Strategies to Increase the Use of Research in Policy', *Social Science & Medicine*, 136-137: 147–55. https://doi.org/10.1016/j.socscimed.2015.05.009.

Rishma, M., Sandra, M.-J., and Josephine, B. (2018) 'How to and How Not to Develop a Theory of Change to Evaluate a Complex Intervention: Reflections on an Experience in the Democratic Republic of Congo', *BMJ Global Health*, 3: e000617. https://doi.org/10.1136/bmjgh-2017-000617.

Rogers, P. J. (2008) 'Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions', *Evaluation*, 14: 29–48. https://doi.org/10.1177/1356389007084674.

Scott, C. S. et al. (2014) 'Expanding Assessments of Translational Research Programs: Supplementing Metrics with Value Judgments', *Evaluation & the Health Professions*, 37: 83–97. https://doi.org/10.1177/0163278713500984.

Searles, A. et al. (2016) 'An Approach to Measuring and Encouraging Research Translation and Research Impact', *Health Res Policy Syst*, 14: 60. https://doi.org/10.1186/s12961-016-0131-2.

Smit, J. P., and Hessels, L. K. (2021) 'The Production of Scientific and Societal Value in Research Evaluation: A Review of Societal Impact Assessment Methods', *Research Evaluation*, 30: 323–35. https://doi.org/10.1093/reseval/rvab002.

Strahan, K., Keating, A., and Handmer, J. (2020) 'Models and Frameworks for Assessing the Value of Disaster Research', *Progress in Disaster Science*, 6: 100094. https://doi.org/10.1016/j.pdisas.2020.100094.

Sørensen, O. H. et al. (2022) 'Measuring Societal Impact of Research—Developing and Validating an Impact Instrument for Occupational Health and Safety', *Research Evaluation*, 31: 118–31. https://doi.org/10.1093/reseval/rvab036.

Taplin, D. H., and Clark, H. (2012) *Theory of Change Basics. A Primer on Theory of Change*. New York City, NY: ActKnowledge.

Van Eerd, D., Moser, C., and Saunders, R. (2021) 'A Research Impact Model for Work and Health', *American Journal of Industrial Medicine*, 64: 3–12. https://doi.org/10.1002/ajim.23201.

Weiss, C. H. (1997) 'Theory-Based Evaluation: Past, Present, and Future', *New Directions for Evaluation*, 1997: 41–55. https://doi.org/10.1002/ev.1086.

Williams, V. L. et al. (2009) *Demonstrating and Communicating Research Impact: Preparing NIOSH Programs for External Review*. Santa Monica, CA: RAND Corporation.

Wyatt Knowlton, L., and Phillips, C. C. (2008) *The Logic Model Guidebook: Better Strategies for Great Results*. Thousand Oaks, CA: Sage.

Young, J. et al. (2014) '*ROMA a Guide to Policy Engagement and Influence*'. London: Overseas Development Institute.