

Developing a Recursive Evaluation Plan of a Complex Interprofessional Healthcare Education Initiative

**J. Scott Parrott, PhD, Patricia Findley, DrPH, MSW,
Matthew Rosenthal, MS, Pamela Rothpletz-Pugila, EdD, RD**

Abstract

Background: A university interprofessional education (IPE) and interprofessional practice (IPP) initiative is a complex undertaking: incorporating multiple system levels (administration, faculty, students, patients), integrating many theoretical perspectives, and coordinating a host of individual IPE research projects. Guidance for evaluating such an IPE initiative is lacking.

Methods and Findings: We describe five key challenges to evaluating the effectiveness of such an initiative, and the processes and tools we have developed to meet those challenges. We draw from recent developments in evaluation science to theoretically ground our description. Additionally, we share concrete tools we have developed in the process. By tacking between theoretical and concrete aspects of our efforts, we hope to both provide ideas for other IPE initiatives, as well as provide a basis for future research comparing cases (complex university IPE initiatives).

Conclusions: While all complex IPE university initiatives are unique, we suspect that they share many common evaluation challenges. By framing these common practical challenges as common theoretical challenges, we seek to offer a description of our concrete case as well as a basis for future comparison of similar initiatives.

Keywords: Interprofessional education; University education; Complexity; Evaluation

Introduction

How do we evaluate a complex, multifaceted, and evolving interprofessional education initiative in a large U.S. university? Where do we start? Is there a roadmap relevant to a large school of biological and health sciences in a major U.S. university? While models for designing interprofessional healthcare education (IPE) initiatives exist, models for the systemic and systematic evaluation of entire initiatives (rather than assessment of limited student outcomes) are less frequently reported. Evaluations of IPE initiatives have more frequently begun to draw from a complex-

ity framework [1,2]. There is a good reason for this: evaluating a complex, evolving, and multipronged initiative involves much more than reporting findings about students “outcomes” for discrete IPE events [3].

Our purpose here is to describe the evaluation framework and tools we developed to meet the challenges of instituting and understanding an Interprofessional Education (IPE)/Interprofessional Practice (IPP) initiative at Rutgers University over the past two years. We do not attempt to define a general path to evaluating a complex initiative. Indeed, we doubt such a path exists [4]. Rather, we focus on the process of developing the evaluation framework to meet the particular needs of our context (though we suspect that many of the issues we identify are encountered by other institutions). This process has been iterative, negotiated, and contested, and we expect that it will continue to be, due to the changing nature of health care delivery under policies guided by the Affordable Care Act. The key, though, is that we have developed an approach that, while rigorous, provides the ability to incorporate divergent perspectives on the question: what, exactly, is going on here?

The goal of this article is twofold: (1) provide an outline of the evaluation challenges, and (2) describe the emerging evaluation framework and tools used at our institution for this particular initiative. So, our effort is both general and concrete. Our hope is to engage our program delivery problem in such a way that it solves our immediate evaluation challenges, but also contributes to the larger theoretical issues involved in the evaluation of complex initiatives.

The Case: A brief overview of the Rutgers Interprofessional Education Initiative

Before describing the specific evaluation challenges and approaches we have developed to address those challenges, we offer a schematic description of the Rutgers Interprofessional Education Initiative.

The setting: Rutgers Biomedical and Health Sciences (RBHS) is a campus within the larger Rutgers, The State University of New Jersey. RBHS is itself comprised of eight separate schools (including two medical schools, pharmacy, dental, health related professions, nursing, biomedical sciences, and public health). Each school comprises multiple departments.

The initiative: In 2013 a chancellor-level Office of Interprofessional Programs was established within RBHS. Uneven local interest and efforts in IPE/IPP were in place prior to 2013, due, in part, to different professional licensing requirements. Cooperative efforts across RBHS schools and with Rutgers schools outside of RBHS (e.g., Rutgers School of Social Work) were in place prior to 2013. However, there was no concerted attempt to coordinate across programs, departments, or schools. The formation of centralized administrative units within RBHS and its component schools was, therefore, a response to disparate though uneven grassroots level interests and efforts. The charge of the new organizational units is to facilitate the development, implementation, and evaluation of interprofessional educational, clinical, and research programs across RBHS.

Challenges of evaluating IPE initiatives at Rutgers RBHS

Evaluation and communication of results were integral components of the vision for instituting a university IPE initiative from the beginning. A review of current research on the evaluation of complex interventions, as well as our experience with some very early IPE events, led us to conclude that an evaluation plan for a *comprehensive initiative* needed to have characteristics quite different from the assessment of limited outcomes of discrete IPE events—while at the same time providing an evaluation infrastructure in which separate IPE-related research projects could be investigated. Our approach is best described as *bricolage*, drawing from different evaluation approaches and perspectives that all have one thing in common: they provide principles to guide evaluation under conditions of innovation, change, and, most importantly, complexity [5-8].

Typical research projects face similar challenges: clear definition of the question, adequate resources, rigorous methodology, et cetera. However, when trying to understand how, whether, or to what degree individual IPE-related research projects mesh into the progress of a larger initiative, additional challenges emerge. Five key initiative-level evaluation characteristics posed substantial challenges and shaped our evaluation approach. Not unlike the process described by Pippa Hall, Lynda Weaver, and Pamela Anne Grassau [2], the search for practical solutions to these concrete challenges leads us to draw on aspects of three different though closely related evaluation approaches: complexity theory [5,9], realist evaluation [6,10,11], and developmental evaluation [8]. Because the majority of our stakeholders (administrators, faculty, staff, students, and, ultimately, patients) are not experts in either interprofessional education or methods of program evaluation, framing the methodological approaches as solutions to concrete problems has been vital.

In the following section we describe:

1. Concrete examples of each of these challenges for evaluation,
2. The theoretical grounding of each of these challenges as they relate to initiative-level evaluation, and
3. The implication for our evaluation effort in terms of concrete tools and processes.

The tack between the concrete and the theoretical is intentional. Any tool or process created to help direct and evaluate the larger initiative needed to both solve concrete challenges faced by our stakeholders and be grounded in relevant theory.

Challenge 1: Multiple entangled systems

Example: Our IPE initiative seeks to effect change at student, faculty, university, and eventually, patient and health system levels [3]. Outcomes of the initiative at one level can affect other levels: “up” (students responses to IPE events change faculty behaviour) as well as “down” (administrative changes can alter faculty choice architecture and thus faculty behaviour).

Theoretical background for evaluation: An interprofessional education initiative involves multilevel and interpenetrating systems—comprising different organiza-

tional units, stakeholder networks, and relationships with different external institutions (e.g., professional associations, funding and accrediting organizations, etc.). However, for most of our stakeholders, knowledge and the focus of concern has historically been highly local—confined largely to a single discipline (e.g., medicine) or program. Broadening both the level of concern as well as knowledge necessitated the development of an orienting framework that would break out of an oversimplified duality of “us” (*my* discipline, or *my* program) versus a reified “Rutgers.” We needed an explicit framework that prompted local stakeholders to view “the university” as a dynamic system of systems—composed of any number of interrelated components [5,12], involving human as well as non-human elements [13,14], and that focused attention explicitly on “linkages, relationships, feedback loops and interactions among the system’s parts” [15]. To describe the initiative as a *dynamic* or *complex* system highlights the nonlinear, emergent, and self-organizing properties [16]. Explicitly framing the initiative in this manner is a strategic decision. It both prompts initiative evaluators to identify the likely systems and subsystems implicated in any particular evaluation event, but also problematizes stakeholder assumptions that “what happens over there” has no effect on “my local circumstance.” For instance, a change in accreditation standards for the physician assistant profession may or may not have any implication for educating dietetic interns. By adopting a complex systems frame of reference, that question is at least now on the table.

Concrete evaluation needs

From an evaluation perspective, the reality of the initiative as multiply entangled systems gives rise to the need to: *Create a representation of the system/subsystems that is both collaborative and easily edited.* This system map can serve as an orienting framework for all stakeholders and provide a way to “locate” where particular events, individual research projects, or changes fit within the larger initiative.

Challenge 2: Context dependence

Example: Our initiative has set goals for different system levels: improving collaboration among students, faculty, departments, and schools. Whether a particular effort (IPE event, change in curriculum, faculty development effort) “works” or not depends in large measure on the context. But, for any given effort there are multiple contexts. For instance, faculty efforts take place within the context of teaching loads, departmental “cultures,” school policies, existing networks among faculty, et cetera. From an evaluation perspective, identifying which contexts matter is anything but straightforward.

Theoretical background for evaluation: The focus of an *initiative*, in contrast to a specific IPE event or a particular course, is to make changes to the larger system so that effects we expect to result from the concerted efforts that fall under the heading of “IPE” are more pervasive and lasting. How might we expect this to happen? In the words of Ray Pawson [17], “If the right processes operate in the right conditions, then the programme will prevail.” The obvious evaluation implication is that we need to gather information on both the processes as well as the *conditions*. In other words, we need an evaluation plan that directs the faculty carrying out the actual interpro-

professional education tasks (i.e., the actual instruction) to theorize about the context(s) of their activities, and how these may increase or decrease the likelihood of the success of their efforts. This requires helping our stakeholders—many of whom operate within the randomized-controlled trial paradigm of simple cause-effect relationships purified of environmental noise—adopt a different way of thinking about how complex interventions work. Gathering vital information on conditions of successful IPE means shifting to what Charles Ragin [7] calls “configurational thinking.” In other words, adopting the perspective that IPE events, curricular changes, collaborations, etc. will likely only result in lasting changes under certain combinations of conditions. Thus, evaluation planning requires the close consideration of the different contexts of IPE efforts that alternatively enable or inhibit specific efforts to have their intended effect.

Concrete evaluation needs

From an evaluation perspective, this sensitivity to context gives rise to two concrete needs:

1. **Create an IPE research-planning template that both solicits theory and prompts the researcher to identify context.** When planning specific research projects within the larger evaluation effort, solicit input from stakeholders not only about their interventions and expected outcomes, but also capture perspectives on *why* they expect the particular effort to work and what the contextual conditions of success or failure might be.
2. **Create initiative-evaluation tools to identify and compare cross-subsystem relationships.** Adopt the perspective that all IPE *planning* activities are data collection events. For instance, planning for changes in school policy, discussions of curricular innovations, and meetings to plan an IPE event are all instances in which theories are assumed about the effectiveness of particular mechanisms and the conditions under which they might succeed or fail.

Challenge 3: Theoretical pluralism

Example: Theories that explain how, why, and under what conditions students learn may be very different from theories of effective faculty development. Theories of organizational behaviour may be beneficial for understanding the effects of changes in structure and processes, but may give little insight into concrete student learning outcomes.

Theoretical background for evaluation: We do not believe a single theory is able to encompass the range of efforts and stakeholders in a complex, multilevel initiative [18]. Depending on which subsystem(s) within the initiative is the focus, different types of theory may be more useful for practice and evaluation. For instance, stakeholders focusing on curriculum and student learning may draw on different learning theories [19]. Stakeholders whose interest lies more with interdisciplinary student or

faculty teams may be more likely to draw on psychosocial theories. Stakeholders interested in more global aspects of the initiative or primarily interested in cross-system linkages may find sociological, organizational, or systems theories [20] more useful. That theory is vital is assumed [3,21,22]. Theoretically driven evaluation enhances the initiative goal efforts by clarifying what outcomes (goals) are expected, as well as by assessing how and to what extent the initiative has realized those goals. The major challenge lies in actually weaving [2] these different theories and perspectives into a coherent and workable evaluation plan.

Concrete evaluation needs

From an evaluation perspective, the reality of multiple explanatory theories (sometimes complementary, sometimes competing) points to the following needs:

1. **Capture researcher explanatory theories at both research and evaluation levels.** Evaluation tools should be designed to help explicate and integrate theories at both the initiative and concrete IPE intervention level.
2. **Sensitize researchers to causal mechanisms and context dependencies.** Stakeholders should be prompted to identify the explanatory mechanisms and context dependencies (explicit or implicit) for why they think an IPE effort will result in or contribute to the desired outcomes.
3. **Analyses of the range of theories should be planned.**

Challenge 4: Research coordination

Example: Evaluations of faculty and student outcomes for different IPE events use different assessment tools. Data are captured in different formats in different data collection platforms. Coordination among administrative units for the collection of student information across courses and events can pose a serious challenge.

Theoretical background for evaluation: Related to the challenge of theoretical pluralism is the challenge of coordination and sharing of evaluation resources. Creating different measurement tools, data collection resources, data platforms, et cetera from scratch not only risks wasting limited resources, but also substantially complicates the IPE effort. Theoretical, methodological, and statistical expertise is, likewise, not an infinite resource, and lack of coordination and prioritization at an initiative level can hamper evaluation efforts. While different research questions and tools are needed, time is wasted constructing and vetting tools and data collection platforms already used in other parts of the university. Additionally, incompatible methods of assessing the same outcomes may limit the ability for data in one part of the organization to be combined with data in other parts. This does not mean gathering homogeneous data, however. The Institute of Medicine [3] recommends triangulating multiple data collection approaches, which could include the use of administrative data. Thus, procedures are needed for combining administrative, faculty, student, and patient data from university administrative and academic datasets, while ensuring stakeholder privacy.

Concrete evaluation needs

From the above, we identified the following evaluation needs:

1. **Create an IPE research template to facilitate cross-event IPE evaluation.** Create a tool to aid stakeholders to carry out IPE-related research projects, while at the same time coordinating these efforts and capturing data on them.
2. **Create an evaluation tool library.** A set of data collection tools, vetted for use in populations such as ours, should be combined into an evaluation tool library.
3. **Create a common, secure data platform.**

Challenge 5: Emergence and recursivity

Example: As goals are achieved, what was innovative can become normalized. At the same time, efforts have unintended consequences and conditions in the larger university context change. Methods of identifying and measuring both the new goals and new processes for achieving them are needed. Unintended results, processes, and relationships can affect the ongoing success of the initiative both positively and negatively. For example, individuals initially neutral toward the initiative may create a coalition to actively resist implementation. Interdisciplinary teaching events may facilitate cooperation on evaluation and research projects across departments, enhancing scholarship or opening new research venues.

Theoretical background for evaluation: Pawson [6] calls emergence “complexity’s final and most devilish twist” (p. 42). From a methodological perspective, this “devilishness” has three faces: (1) you can only measure what you know to measure, thus it is easy to miss unanticipated emergent properties that may be critical for understanding the changing trajectory of the initiative, (2) “outcome variables,” in the form of initiative goals, will alter over time, and (3) new theories and expectations of “what is happening and why” may develop among stakeholders. And the challenge is not merely a matter of developing a methodological approach that can flexibly detect, prioritize, and explain. The approach must also facilitate negotiation among stakeholder’s contested visions of the initiative, including clinical versus research-relevant outcomes. Even fairly simple physical and mathematical dynamic systems can give rise to complex and surprising results [18,23]. However, when the focus is on human behaviour (as it is with IPE/IPP) the situation for the evaluation team can become even more difficult.

Different stakeholder groups can “see” quite different results (e.g., what is happening in the dental school appears to be something quite different than what is happening in the allied health school), they may have different or even conflicting notions of what the outcomes are or should be, and they may develop different folk theories to understand the changes (or lack thereof). This leads to what David Byrne [4] calls “complex complexity,” where the complex processes of changes to entangled systems are filtered through the different interpretive and evaluative frameworks of a range of local stakeholders. Methods for identifying and measuring new system realities imply that the evaluation framework has to evolve with the systems. In addition, pro-

cedures need to evolve to identify emergent outcomes and processes (both anticipated and unanticipated), while allowing for competing stakeholder explanations of the process. Additionally, because changing perceptions and understandings of the evolution of the initiative over time can reframe stakeholder perceptions of what happened at earlier points [24-26], it is vital to not rely purely on the memory of the participants.

Concrete evaluation needs

The reality of emergence leads us to identify the following evaluation needs:

1. **Capture stakeholder perceptions of change.** The evaluation methodology for the IPE initiative must be designed from the beginning to adapt to changing system conditions and stakeholder goals and understandings.
2. **Collect data across time.** The evaluation process cannot be a “one off,” but must collect data across time and across different subsystems, incorporating, but not relying on, human memory [3].
3. **Identify evolving evaluation questions and methods.** If possible, records of stakeholder negotiations about what has changed, what needs to be measured, and why this is happening should be captured in documentary format so that competing understandings of the initiative can be captured across time.

In the face of these challenges, we sought to outline a coherent process, concrete tools, and research infrastructure (our “evaluation plan”) for evaluating both individual IPE efforts (e.g., changes in curriculum, specific IPE events, individual research projects) as well as the broader IPE initiative.

In less abstract language, these five challenges can roughly be organized into three general stakeholder questions:

1. Where does this research fit within the larger scheme and what does it depend on? (Understand entangled systems and plan for context dependence.)
2. Why do I believe this intervention/effort will work and what do I need to find out? (Capture plural theories and coordinate research activities.)
3. What happened and how might this change the IPE initiative going forward? (Identify emergence and facilitate recursivity.)

In the following section we describe our approach (thus far) to answering these three questions and addressing the five challenges of initiative evaluation.

Mapping the system and context dependencies

A rough map of entangled systems

The first step in defining the evaluation plan was to develop a general, initial map of the IPE system and subsystems involved in our initiative. Because multiple IPE activ-

ities and research projects were operating at the same time and at different levels, an *initiative* level planning and evaluation framework needed to provide us with a way to “see” all of these independent though related efforts together; and not only see them, but to plan for how to mobilize school-level resources to support these efforts and evaluate their place within the larger initiative. This working tool allows us to identify the key systems involved, and provides a way to visualize how activities and research efforts in one part of the larger initiative relate to other parts of the system. This ability to visualize the moving parts (e.g., changes in school resources, creation of case scenarios, engaging and training faculty facilitators, implementing curricular changes, etc.) of the initiative is key. Faculty involved in particular IPE events are often heavily focused in efforts that directly concern them, but may be unaware of efforts in other areas of the university. This localized focus can lead to (a) competition over resources (e.g., space, statistical support, etc.) and (b) an uneven evaluation focus.

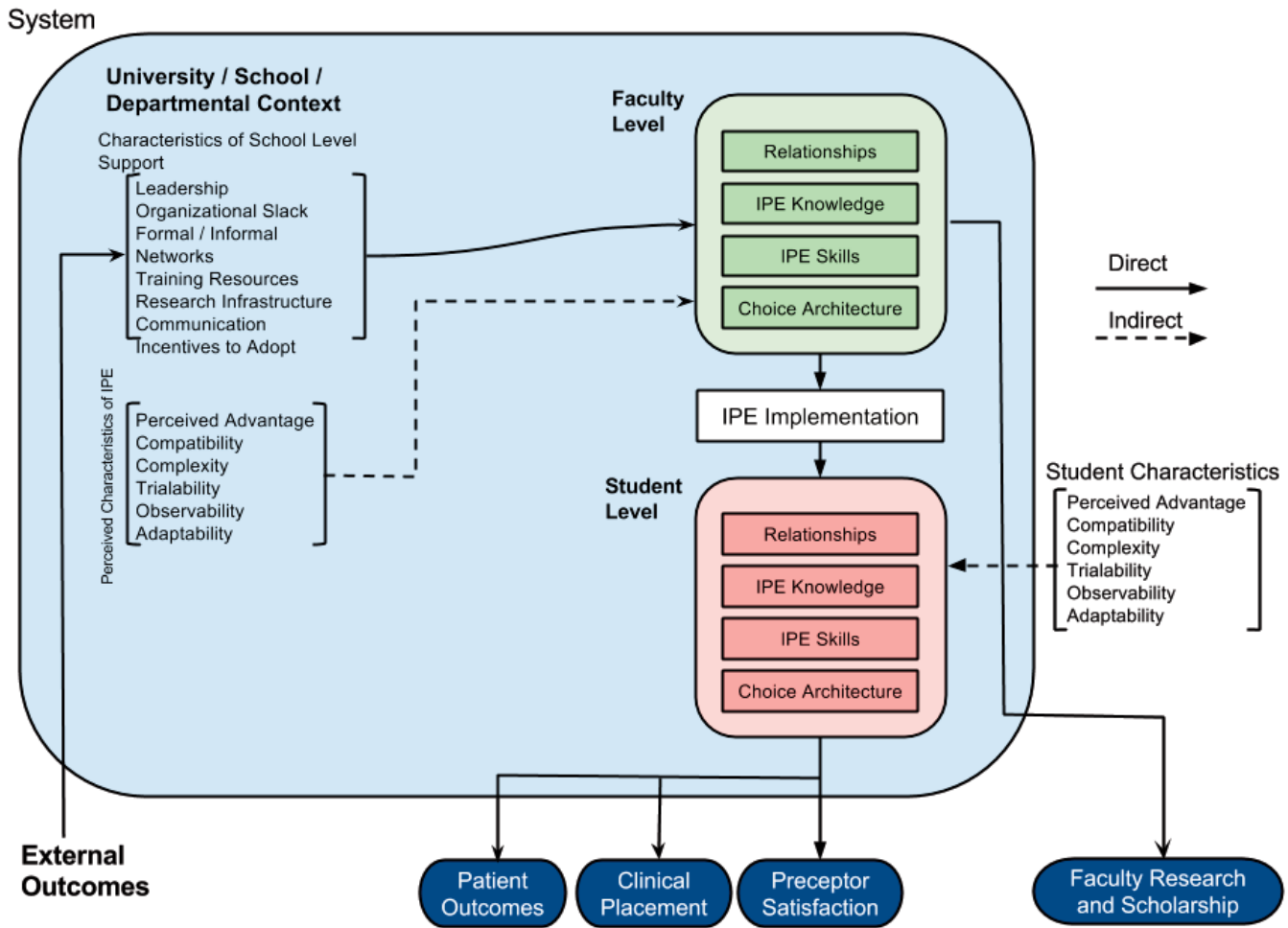


Figure 1. Map of IPE initiative entangled systems

Additionally, the relationships depicted in the map provide direction for identifying cross-system linkages to be accounted for when understanding the events at lower system levels. Because higher-level systems (e.g., the university or faculty levels) provide the context for the operation of causal mechanisms [9] at lower levels,

specifying cross system relationships is vital. For instance, changes in administration activities provide the context for faculty behaviour, which, in turn, provides the context for student activities, and any one of these changes may or may not have the intended effect. Therefore, an adequate framework for evaluation is needed to be able to provide a coherent map in which more specific context, mechanism, and outcome linkages could be identified and examined [10].

The current version of the map of the major entangled systems involved in the IPE initiative is presented in Figure 1.

How do we use it?

The map provides an image of the initiative components and relationships (and, so, directs attention toward some aspects and not others). However, any given version of the map is not meant to be definitive. Though the initial map was developed in discussion among stakeholders, it quickly became apparent that a static picture was not going to meet our evolving needs and perspectives in regards to our initiative. The map needed to be fluid and provide the basis for discussion and negotiation. In short, the map has been and continues to be subject to ongoing revision. Over the past two years, the map has gone through several revisions, as faculty with different backgrounds and theoretical perspectives have become involved in the initiative.

In order to facilitate the developing (and sometimes contested) nature of the initiative, the map was recreated in an online platform (Google Drawings) in order to facilitate collective comments, editing, and revision. An unanticipated benefit of using the online platform was the ability to hyperlink to sub-algorithms, study descriptions, et cetera. For instance, embedding the student-level box with a hyperlink allows us to immediately jump to a page listing all current studies examining student outcomes. Thus, the map as figured above [Figure 1] is only the top level of a much deeper map of the system.

The map does more than simply provide a picture of the initiative as a system of systems. It helps with both initiative planning and initiative evaluation.

Initiative planning: Targeting contexts

The map allows us to plan for strategic interventions to change contexts that would enable (trigger) relationships at lower system levels. If we discover that faculty interest is tending to focus on a particular aspect of IPE (e.g., integrating interprofessional practice concepts into individual course curricula), then the map prompts us to ask what department or school resources might facilitate that interest. For instance, altering the organizational context to incentivize faculty to increase their IPE knowledge may increase the link between the strength of the curricula and student knowledge. In other words, students may learn better *within* classes when the classes are set within faculty- and university-level contexts that support the class-level mechanisms. The map can help us prioritize what characteristics of the context need to be addressed to enhance a link between mechanism and outcome at a lower system level.

So, for instance, in Figure 1, particular characteristics of the school-level system have been identified that have been demonstrated to be associated with the diffusion

of innovations within a health service context [27] (e.g., leadership, communication, training resources, etc.). Whether or to what degree school-level changes actually do provide supportive or enabling contexts for faculty- or student-level IPE efforts can be identified, planned, and evaluated in its own right.

Evaluation planning: What questions can be answered and how?

Identify answerable research questions: The map is designed to help focus specific IPE research projects as well as organize the larger initiative evaluation. (We draw on the distinction described by James Fain [28], where the purpose of research is to identify causal relationships generalizable beyond the particular program or project, while the goal evaluation is to make a judgment about the effectiveness of a particular program.) For instance, we may be interested (from an evaluation perspective) in “the effects” of forming an IPE task force (university level) on student team-building skills (student level). But this is unlikely to be a feasible question for a specific research project.

The map helps sensitize stakeholders to one of the hard facts of complex systems: causal explanations in social research are local [4,5]. Whether this or that student improves their teamwork skills during a particular case simulation (or even over the history of a course) will generally be traceable in any clear manner to individual, class, instructor, and to some lesser degree department characteristics. The specific actions of a university-level organizational unit are unlikely to be traceable as causes in this case. Attempting to research highly mediated cross-system relationships may simply not be feasible, even if, from an evaluation perspective, this mapping is a primary goal. The map helps sensitize stakeholders to this reality.

Method and tool selection: The map additionally helps to identify research and evaluation methods appropriate for the particular relationship of interest. Because the map helps identify level-crossing relationships, stakeholders can identify situations in which a mixed-methods approach is used to link quantitative relationships (e.g., increased departmental participation leads to increased student exposure) to a qualitative evaluation of the ways in which different aspects of the context may affect the relationships. For instance, the degree to which “a student is exposed to IPE” and the “level of faculty enthusiasm” for IPE is mediated by the educational theories instructors implicitly or explicitly adhere to when organizing their curriculum [29].

Moving from map to theory

The map does not, however, tell us *why* particular interventions work or not. For instance, we may identify recurrent patterns between the use of standardized patients with faculty facilitators and improved students communications skills, but it is not clear why this should be the case (i.e., What is it that standardized patients add to an IPE case scenario? Are faculty facilitators needed? Under what conditions?). To understand the explanatory connections between simulated multidisciplinary team problem solving and the use of those skills in actual clinical situations, some rationale (theory) is needed that connects the two.

Integrating research and evaluation theories

As we noted above, stakeholders have different perspectives or explanations (theories) for why different IPE interventions might work (research questions) or how and why some aspects of the initiative are working while others are not (evaluation questions). In order to capture these theories, we needed to first enable our stakeholders to identify and articulate them. Because our stakeholders come from different clinical backgrounds, they do not necessarily share a common lexicon or framework for articulating behavioural or organizational change theories. So, in order to try to capture and possibly integrate the range of research and evaluation theories, we needed to (1) equip our stakeholders with a common terminology and (2) enable ourselves to disentangle (a) research theories from (b) evaluation theories.

Clinical versus IPE research: Establishing a common lexicon

We approach the first problem by drawing from the lexicon three related evaluation perspectives (none of which alone provided us with the conceptual tools we believed we needed): realist analysis [6], configurational comparative analysis [7,30], and complexity analysis [5]. We do not attempt a detailed explication or critique of these perspectives here. Rather, our purpose is to define the terms we use to help us understand how to use theory for our IPE initiative. Table 1 details our working definition for each of the terms.

Table 1. Definitions of terms used in the evaluation framework

Term	Working definition
Characteristics	Aspects of a situation identified as being salient for an adequate description of the situation.
Mechanism	A configuration of characteristics that are proposed to have a “logic”, that is, the relationship between the characteristics can be stated in general terms and thus identified as present/absent or more/less across instances.
Case	A concrete configuration of mechanisms as enacted by actual individuals within the system.
Context	A configuration of characteristics at an analytic level above the case, which are hypothesized to affect the operation of the mechanisms within a particular case.
Outcome	A state of the system or sub-system present at a particular time after the occurrence of the purportedly causal event and which results from the operation of mechanisms within cases within contexts.

How do the different terms in Table 1 relate to theories? Pragmatically, a “theory” is an explanation of how aspects of configurations (characteristics, mechanisms, contexts) “work” to bring about a particular system state (outcomes). Theories can be more or less structured, more or less explicit, “folk” theories or formal theories, et cetera. The point is that they are verbal expressions that capture different answers to “why does this occur?” And the purpose of these explanations is that they direct attention to (a) which combinations of characteristics are of interest [31], (b) the mechanism or logic of the

structured relationships among characteristics [11], and (c) how these configurations of mechanisms should be related to outcomes [6]. Theories are vital because they both direct planning as well as determine the evaluation approach [21].¹

From the perspective of the initiative evaluation, our goal is to catalogue and understand the theories (explanations) of why particular IPE efforts are associated with particular IPE-related results. Within their own disciplines, our stakeholders are adept at providing explanations (identifying the theories that inform their practice and research). But, when we move to interdisciplinary education research and evaluation, we have found that our stakeholders benefit from support in developing explanations and are often unfamiliar with the theories that underlie their efforts [19]. Consider the following three questions in Table 2. While our stakeholders need no support to identify relevant theory for research questions in their own domain of clinical expertise (the first type of question), we have found that support is needed for questions of the second (IPE research) and third type (IPE evaluation). In the sections below, we describe how we both prompt and capture information in regards to the competing theories encompassed by both IPE research and IPE initiative evaluation.

Table 2: Stakeholder theoretical familiarity by domain of research and evaluation inquiry

Domain of inquiry	Typical stakeholder familiarity
Disciplinary research	
Why should differentiating treatment of cystic fibrosis patients based on their pre-treatment inspiratory muscle strength level improve treatment outcomes?	Consistent deep content knowledge in the area of expertise. Stakeholders can easily explain how something works and why it works (easily identify all items listed in Table 1).
Interdisciplinary research	
Why should using TeamSTEPPS* training across departments decrease the number of errors in interprofessional team acute care simulations?	Moderate to shallow and inconsistent knowledge of educational theories in general and IPE theory in particular. Stakeholders can generally offer informed “folk theories” if prompted.
Initiative evaluation	
Why should comparing the educational theories behind different IPE training instances help us understand how the IPE initiative is working?	Little to no familiarity with theories of organizational or social change or principles of program evaluation. Stakeholders need a great deal of support.

Notes: * TeamSTEPPS® is an evidence-based teamwork system available from the Agency for Healthcare Research and Quality (AHRQ). Its purpose is to optimizing patient care through improving communication and teamwork skills among health care professionals as well as frontline staff such as registration desk workers.

Prompting, articulating, and capturing IPE research theories

While the initiative map keeps before us the multilevel and context-dependent

nature of our IPE initiative, promoting theoretically driven evaluation can help us manage the challenges of theoretical pluralism (multiple theories are needed at various levels) and coordination (what is being evaluated? What are the competing explanations? How are resources allocated among discrete IPE research efforts?). As a practical matter, this means collecting key pieces of information on both specific IPE events, as well as integrating these into the larger initiative evaluation plan.

Since faculty often consult with university administrative offices for data collection and logistical and statistical support, our approach uses the process of providing university support for particular research efforts as a way to collect information and data for initiative level evaluation and strategic planning. In the normal course of providing methodological, data collection and statistical support for research on IPE events, certain common pieces of information are gathered. Typically, this information is discussed in research planning meetings, but details were rarely kept. We realized that by capturing information on the research objectives of specific IPE projects, we could begin to piece together a picture of the larger initiative. So, we created an event evaluation template (based roughly on the W.K. Kellogg Foundation Logic Model Development Guide [32]) that allows us not only to capture information from the faculty to help direct the particular analysis plan, but that can also be used as a vehicle for collecting information across the entire IPE initiative. Figure 2 provides the template of the information collected as well as how the information is used to direct methodological approach, data collection and analysis.

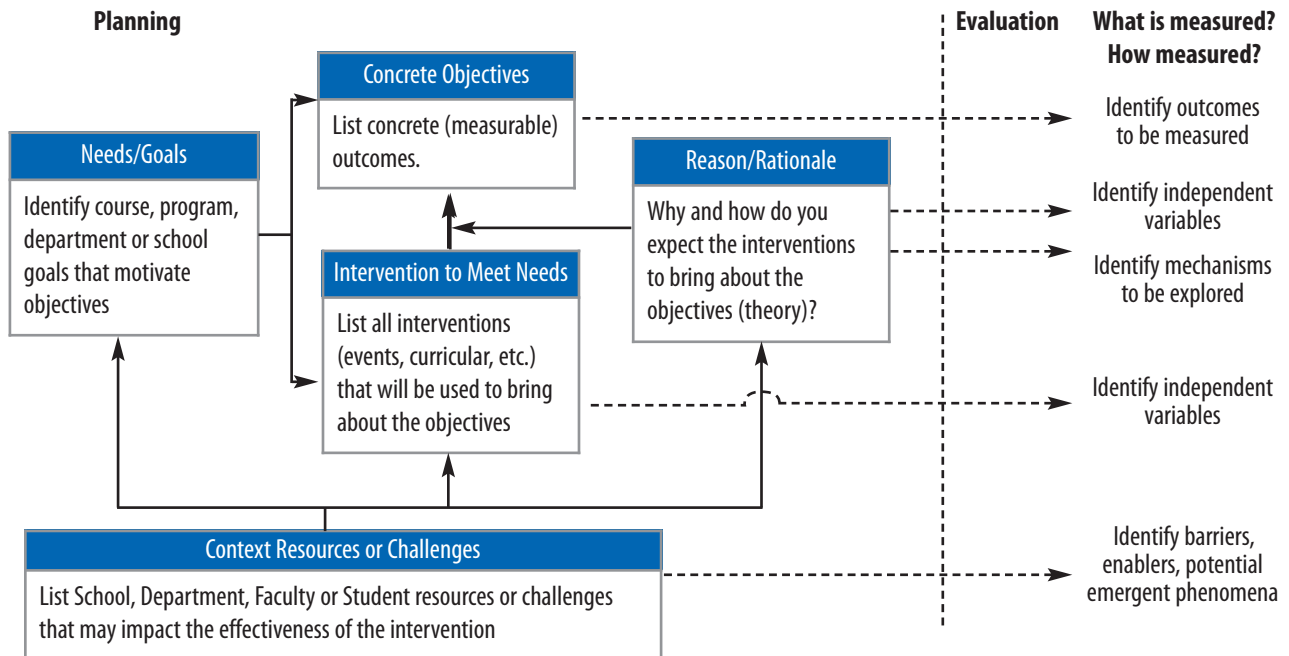


Figure 2. IPE research project planning template

Two key pieces are included in the above research planning template that are vital for our evaluation of the initiative, and that are not commonly considered when planning research on limited IPE efforts or events: reason/rationale and context resources or challenges. By encouraging researchers to provide some explanation for why and

how they expect the particular intervention to bring about the target outcomes, we can begin to build a representation of the different theories (explanations) used or assumed by stakeholders. This provides the basis for an initiative-wide assessment of competing explanations of why IPE “works” and will allow us to identify potential areas of elaboration and exploration. Encouraging researchers to think about the higher-level conditions for success (or the challenges they face) can provide direction to the initiative for changes at higher system levels. A regular review by the major stakeholders can then provide the basis for identifying which competing theories [2] to prioritize in order to focus research efforts as well as identify larger system changes.

Prompting and capturing IPE Initiative Evaluation

While our stakeholders are well acquainted with the methods of analysis typically used when answering IPE *research* questions, they are often less familiar with the combinatorial comparative methods that are more relevant for identifying patterns for evaluating initiative-level questions. Figure 3 provides a schematic to help our stakeholders begin to specify theories for initiative-level implementation and evaluation. The basic unit of analysis is the *mechanism*, which, for clarity, we have simplified into “being present” (M) and absent (~M). So, to use a concrete example, a statement of the mechanism might look something like “the presence of faculty IPP team facilitators” (M) should be associated with “students actively soliciting case-relevant information from other students” (O).

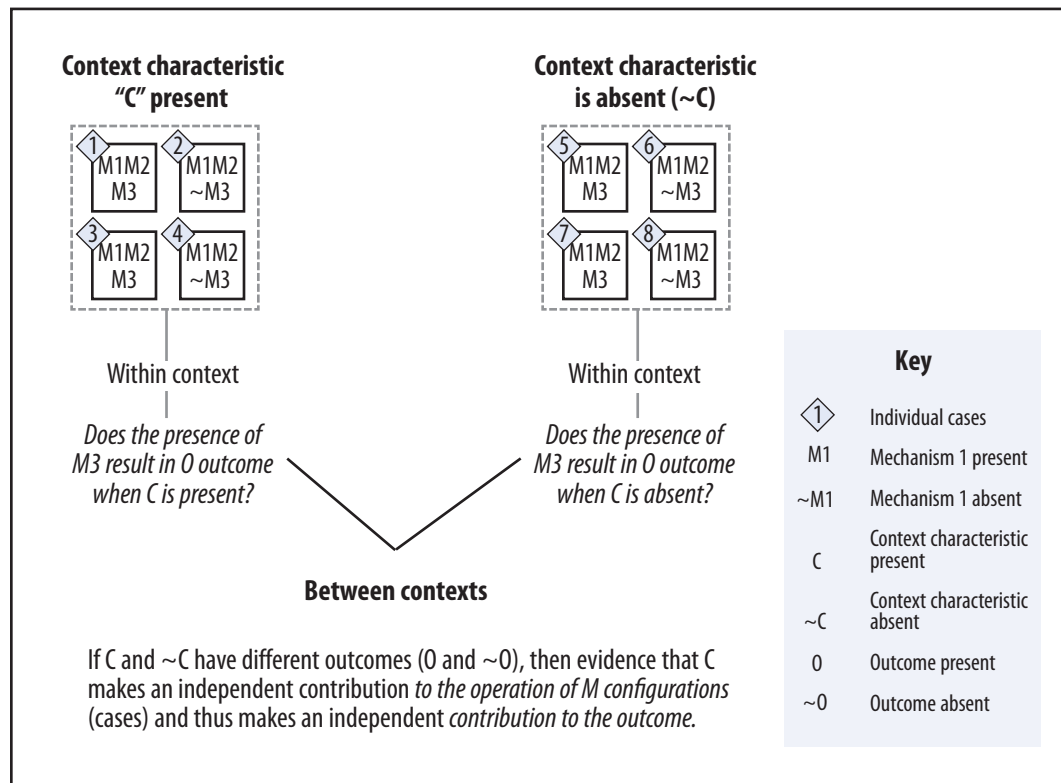


Figure 3. Relations between mechanisms, cases, and contexts: Logic for configurational comparative analysis

In Figure 3 we have included multiple mechanisms (M1, M2, M3) within each case to indicate that within any concrete IPE event multiple causal relationships are operating at the same time. In this example, the presence of a faculty facilitator (M3) is only one possible reason why student communication patterns take the form they do. There are other possible mechanisms at work shaping student communication patterns within the same event (for instance, M1 could stand for “presence of a medical school student” and M2 could stand for “students have clinical experience”).

Cases (indicated by numbered diamonds) differ from each other in their configuration of mechanisms. So, in Figure 3, cases 1 and 3 are alike in that they both have all three mechanisms present (M1, M2, and M3). Cases 1 and 2 differ in that mechanism M3 is present in one but not the other. Case 2 has M1 and M2, but not M3 (\sim M3).

The benefit of an initiative focused (rather than case focused) analysis is that we can collect information across cases and then carry out a comparative analysis. Notice, however, that at this point we have only talked about a comparative analysis *within* a particular context. Revisiting the above example (M3 in Figure 3), the context (C) might be “when faculty are trained in small group facilitation skills.” The question this approach to analysis allows us to answer is: within a particular context (C), is the presence or absence of M3 associated with the desired outcome (or, in the language of combinatorial case analysis, is the presence of M3 consistent with the presence of the outcome (O) under condition (C)? [7]).

If we have cases where case-level mechanisms have the same configurations but *context* differs, then we can ask the *between context* question: are differences in particular context characteristics (e.g., C versus \sim C) associated with differences in how mechanisms “work”? In our concrete example, if the presence of faculty facilitation (M3) is consistently associated with the desired student communication patterns (O), but only when faculty have had prior training in facilitation skills (C) but not when faculty have had no training (\sim C), then we have evidence that spending the time and effort to train faculty in small-group facilitation skills is likely to help bring about the desired student communication patterns.² On the other hand, if we found that there was no difference in an association between faculty facilitation (M3) and student communication skills (O), depending on whether the faculty were trained or not (C and \sim C), then we could determine that expending school resources on faculty facilitation training was not necessary.

In short, adopting a case-based comparative approach helps us to formulate theories for how different types of IPE-related mechanisms operate within and across different system and subsystem contexts. This allows us to gain a larger perspective on how, under what conditions, and for whom the initiative is working.

Our presentation of how we capture and use theory in our IPE initiative above has been unavoidably abstract. More concretely, how does this help us with either initiative planning or evaluation? We answer that question by way of a description of two school-level efforts where implementation and evaluation goals were hampered by a lack of theory (see Table 3).

Table 3: IPE implementation and evaluation efforts hampered by lack of theory: Two vignettes

<p>Vignette 1: Increasing IPE participation via time buy-out</p>
<ul style="list-style-type: none"> • Aim: Enlist faculty champions across departments • Action: Dean's office to buy out or allocate up to 10% of a faculty member's time to participate on IPE planning activities and events. • Implicit theory: Individuals are prevented from participating by barrier (time). Remove the barrier and participation will increase. • Result: Extremely low participation in the buy-out. • How theory might improve implementation: The mechanism was never explicitly stated. It is likely that research into and considered discussion of the causes of champion participation would have identified multiple "barriers," "facilitators" or "conditions" for champion recruitment. Something like a force field analysis might have prompted a more nuanced theory and hence multipronged intervention. • How theory would improve evaluation: Some faculty participated, but without using the time buy-out (i.e. volunteered). Other faculty already involved in local IPE efforts did not participate in the initiative. A multipronged theory (specifying multiple mechanisms) would have provided a basis to collect data to construct cases to allow for determining under what conditions faculty would be likely to become champions.
<p>Vignette 2: increasing IPE research by overcoming data analysis barriers</p>
<ul style="list-style-type: none"> • Aim: Increase research on individual IPE events and, thereby, increase faculty publications in this area. • Action: Vet a set of data collection tools and develop a common data platform for IPE-related research questions. • Implicit theory: Individuals are prevented from participating by barrier (time and knowledge limitations on data collection tools and methodologies). Remove the barrier and participation will increase. • Result: Data from common set of tools was collected across four IPE team simulation events captured into central database. Work began on several manuscripts and then stalled. • How theory might improve implementation: While information was useful for answering initiative feasibility questions, lack of theoretical grounding in IPE hampered construction of manuscripts that would answer pressing questions in the larger IPE field.[3] • How theory would improve evaluation: Research was layered onto existing events. The events were treated as ends in and of themselves often without any clear notion of why or how the structure of the event would lead to any particular outcome. Thus, generic data was gathered, but there was no clear notion of how it might answer questions about targeted initiative outcomes.

In both of the school-level interventions described in Table 3, *explicit and more developed theory* may have improved the likelihood of successful *implementation*. Notably, both interventions worked off of the same implicit theory (mechanism): humans are more likely to exhibit a particular behaviour when barriers to that behaviour are removed. In retrospect, the supposition that this single mechanism would be sufficient to bring about behaviour change was clearly inadequate. By drawing on more sophisticated theories of organizational and individual behaviour change [20,33], multi-mechanism approaches could be specified and implemented.

Likewise, *evaluation* of why the implementation efforts were unsuccessful is hampered by a lack of explicit theory. In the first vignette of Table 3, we do not know why individual faculty did not take advantage of the time buy-out. It is still possible, of course, to go back and interview select faculty to identify reasons (and thus, identify

other possible mechanisms that could be mobilized), but a year after the implementation, individual memories are likely to be both fuzzy and have been potentially reframed by intervening initiative-related events. In the second vignette, the implementation of the data collection tools and common platform were successful, but faculty members found themselves awash in data that did not provide answers to questions that were largely implicit at the beginning. By starting the process with clearly defined statements of why a particular event (or aspect of an event) should lead to the desired outcome, appropriate data could have been collected. The cases demonstrate two different hazards of trying to evaluate an intervention without a theory: the evaluators risk either having no data and relying on post hoc reflection or they are faced with a glut of data for which they have no clear use.

Identifying emergence and planning for the future

The final IPE evaluation challenge we sought to address involves two closely related issues: the identification of *emergent characteristics* and *recursivity*. Briefly, the concept of emergence points to the formation of new system properties that did not exist previously. They are “configurational regularities” [5] of relationships, policies, practices, et cetera that both may become self-organizing and have broad systemic effects—literally shaping the nature of the system because of their extensive effects [15]. These may be regularities at lower levels that ramify upward to higher levels of the system (e.g., school response to faculty demand for research resources by creating IPE research infrastructure) or structures created administratively at higher levels that change the context for the operation at lower levels (changes in faculty tenure and promotion criteria).

At its simplest, recursivity can be thought of as an “output” of the system (intended or unintended, desired or undesired) becoming an “input” or context for the operation at later points in time. The potential recursivity of emergent phenomena are important for the evaluation of an IPE in two ways:

1. The contexts of discrete IPE research projects change, thus potentially changing the link between interventions and outcomes as well as potentially necessitating new explanations. This may mean that research on IPE future events should collect data on different variables, utilize different methodologies, or even identify alternative outcomes.
2. Planning for emergence and taking into account recursivity can inform the strategic planning and evaluation of the IPE initiative. Newly available university resources may influence both the methods by which we implement IPE as well as the goals of the initiative. Or, if we know that faculty have participated in opportunities to develop their small-group facilitation skills, then student IPE events can be organized to take advantage of these skills. From the perspective of an initiative-level evaluation, this means that the evaluation plan will need to be flexible enough to account for changing conditions, methods, outcomes, and explanation.

The challenge is how to identify emergent phenomena. Emergent phenomena that take the form of explicit goals (e.g., building a school-level infrastructure for IPE research) are more straightforward to evaluate. We know what to look for. But how do we identify emergent characteristics that are unanticipated? How can we determine when a network of relationships, an event, or a resource has become self-organizing (that is, is now the new normal)?

In order to plan for and anticipate emergent phenomena that have the potential to recursively restructure the system or subsystems, we created a tool to help IPE stakeholders at our university identify emergent phenomena that could impact the initiative (see Table 4 with examples).

Table 4. Example tool for identifying emergent phenomena and assessing potential IPE impact

Level	Aspect	Challenges	Opportunities	Prompt new goals or research?
University/ School <i>Consider how existing or new school or university level characteristics may affect IPE efforts either directly or by changing the faculty or student level context</i>	Policies			
	<i>New promotion guidelines to increase emphasis on scholarship</i>	Lack of research infrastructure for educational research	Motivation for faculty to carry out research on new and existing IPE efforts	Increase the number of IPE research projects and manuscripts
	Structure/Processes <i>(university offices, committees, task forces or formal procedures that may impact IPE)</i>			
	<i>IRB requirements</i>	Many new, limited scope IPE research projects means increased IRB paperwork	Publishing results of IPE experiences in peer-reviewed publications	Explore possibility of umbrella IRB for university IPE research
	<i>IPE Task force dedicated staff</i>	Unclear what demands will be made on new staff person's time	Potential for coordination with Research Office	
	Resources			
<i>Budget</i>	Multiple schools are involved with varying sizes of student bodies, need to balance budget with needs and goals by school	Multiple sources for funding and in kind resources (i.e., administrative staff support)	Potential to develop a collaborative budget for IPE	

Table 4 (continued)

Level	Aspect	Challenges	Opportunities	Prompt new goals or research?
Faculty <i>Consider how existing or new faculty level characteristics may affect IPE efforts “up” (to school) or “down” to student level</i>	Structures (<i>formal and informal networks of faculty that may impact IPE</i>)			
	<i>Manuscript teams resulting from IPE Stroke Event</i>	Logistical (different campuses) and schedule challenges, different levels of interest and ability in drafting manuscripts	Potential model for future multi-faculty research teams	Can use results from publications to contribute to funding applications
	Resources			
	<i>Creation of Redcap research platform</i>	Low faculty awareness	Ability to build out evaluation tools and data platform very quickly	
	Events			
	<i>Faculty small group facilitation training</i>	Varying levels of faculty facilitation experience, and demand for training	Development of a faculty training team on facilitation	Can lead to training professionals in the field
Student	Structures			
	<i>Scheduling of IPE events to allow maximum student participation</i>	Curricular calendars vary in terms of start/end dates, exam periods, and internship times. Finding “best” time to hold events is difficult	Faculty and university administrators need to think more creatively about when and how (i.e., online, in person) IPE events, as well as frequency	
	Resources			
	<i>Over 10 distinct disciplines can be represented by student participation at IPE events within the university</i>	Some disciplines can require entire classes to attend, others rely on volunteer participation at IPE events	Create interdisciplinary courses that are cross-listed to meet multiple needs	Enhanced use of technology to increase the numbers and schools of students for participation in events
	Relationships			
	<i>Faculty are learning different student learning styles and requirements</i>	Students are learning about educational requirements for each other’s disciplines and forming relationships	Smaller and more focused IPE activities can be developed longitudinally to grow in depth with the students’ education	Evaluation of the longer term relationships and the potential impact on IPP

Note: *IPE-related structures or events indicated in italics; aspects of assessment in regular text

The emergent phenomena tool is primarily a planning tool to be used in discussions among IPE stakeholders. The goals are to help stakeholders identify what the potential new conditions of IPE practice at our university are, and to identify potential challenges and opportunities associated with these emergent phenomena. Some phenomena may prompt new goals (or the retirement of prior initiative goals) based on the priorities of the stakeholders. Structural changes may also require revising the overall initiative map (see Figure 1) or theories used to explain new phenomena.

The outcome of the periodic review of the continually changing conditions of IPE practice will affect the larger initiative evaluation. If new goals, methods, or conditions are identified, then the ongoing initiative evaluation plan will need to evolve to meet the needs of the new circumstance.

Discussion

Our interprofessional education initiative is still in its formative stages. While interprofessional events and activities are not new in our university, we made the decision to prioritize and organize these efforts as a way to meet the changing needs of the healthcare profession [34]. Evaluation planning was integral to this initiative from the beginning. However, the broad scope of the initiative presented several evaluation challenges. The initiative operates at several different levels, so changes that occur at one level effect changes at other levels. We needed to be able to explain (rather than merely trace) why changes occurred at the same time as we coordinated the evaluation at different levels. Finally, we needed an evaluation plan that would evolve as the initiative evolved, and to use this information to plan for the future.

To meet these daunting challenges we developed an evaluation approach that draws on principles of both complex systems and realist analysis. And, though the evaluation plan was necessarily complex, it also needed to be practical. Stakeholders needed concrete tools to help them prioritize, negotiate, and evaluate the operation of this ambitious initiative. We summarize the mapping of the challenges to the evaluation tools and processes in Table 5.

Initially, we attempted to identify evaluation approaches we could draw on for our initiative. However, most university IPE evaluation descriptions available in the literature either focused on specific projects or programs within initiatives [35-37]. When they did appear to be in a situation similar to ours [38,39], there was little description of what tools were used or, more specifically, what approaches could be used to understand how theories of IPE research or evaluation might inform the evaluation process. Alternatively, some evaluations of complex initiatives were carried out in a clinical setting rather than in a university setting [2,40,41], and, while these approaches informed our university initiative evaluation, none provided a clear framework for our setting. In short, detailed evaluation plans for complex university IPE initiatives that focus on outcomes at multiple levels and across levels are in short supply [42]. Thus, we were forced to turn to a handful of more general evaluation approaches and methodologies to construct our own.

Going forward, there are many unknowns. In just the two years since the official start of our IPE initiative, our evaluation plan and techniques have already evolved,

Table 5: Summary of challenges, evaluation needs and tools/processes to meet needs

Challenges	Evaluation needs	Tools	Process
Challenge: Multiple systems	Create a representation of the system/sub-systems that is both collaborative and easily edited	Initiative Map	Each evaluation event happens with reference to “where” it falls on the map
Challenge: Context dependence	Create IPE research planning template that solicits both theory and prompts researcher to identify context	Initiative Map with Event Data Collection Template	Specific context dependencies are identified
	Create initiative evaluation tools to identify and compare cross-sub-system relationships	Initiative Map with Combinatorial Case Analysis	Successive refinements captured over time
Challenge: Capture theories	Capture researcher explanatory theories at both research and evaluation levels	Combinatorial Case Analysis Approach and Event Data Collection Template	When providing methodological, statistical and data collection support, capture all information into a secure dataset
	Sensitize researcher to causal mechanisms and context dependencies	Event Data Collection Template	Prompt researchers for perceived barriers, resources needed
	Analyses of the range of theories should be planned	Combinatorial Case Analysis Approach	IPE interventions and planning events are all considered data collection events
Challenge: Coordinate research activities	Create IPE research template to facilitate cross-event IPE evaluation	Event Data Collection Template	Create a database of IPE event research protocol data
	Create Evaluation Tool library	IPE Research Infrastructure	Update library as necessary as new tools are used or developed
	Create a common, secure data platform	IPE Research Infrastructure	Establish common platforms for IPE-related data collection
Challenge: Identify emergence recursively	Capture stakeholder perceptions of change	Stakeholder Planning/Evaluation Template	Solicit stakeholder input and feedback regularly and capture this information using a common template
	Collect data across time	Stakeholder Planning/Evaluation Template	Establish regular IPE evaluation review sessions
	Identify evolving evaluation questions and methods	Stakeholder Planning/Evaluation Template	Consider all decision-making events as data collection opportunities

and the description provided here represents “lessons learned” during that period. We anticipate that there are many more lessons to be learned. We also anticipate that as our faculty becomes more familiar with educational theories underlying IPE, the nature of our IPE research will change—new questions asked, new data sources needed. Indeed, new developments (e.g., the creation of an IPE team simulation video archive) are already underway. We expect that our initiative evaluation will adapt, but our hope is that the reflexive evaluation approach we have described above will allow for this. The proof of the pudding, though, will be in the eating.

While obvious, it still bears stating: our evaluation efforts are neither objective nor disinterested. As with all human behaviour, our IPE evaluation efforts are insinuated within multiple fields of interests, biases, and relations of power. While we do not hold to an ideological view of interprofessional education (i.e., we do not, in fact, know whether IPE/IPP are answers to the “quality chasm” [34] in healthcare or will “mitigate the world health workforce crisis” [43]), we recognize there is a vested interest in the success of our initiative. Our hope, however, is that we have developed a principled evaluation plan for our comprehensive IPE initiative that will make it more difficult to delude ourselves when things go differently (as they inevitably do) than we would like.

Acknowledgements

The authors would like to thank Alexis Fulks and Alma Merians for thoughtful comments on earlier drafts.

Abbreviations

IPE	Interprofessional Education
IPP	Interprofessional Practice
AHRQ	Agency for Healthcare Research and Quality

Notes

1. The reader will have no doubt noticed that there is a fair amount of slippage in these terms. For instance, “mechanism” can refer to a configuration of phenomena (what we are generically calling “characteristics”) within a classroom, within a program, among faculty across programs, at the level of the school, etc. The same applies to the other terms as well. Adopting a multi-system (multi-level entangled systems) perspective means that what counts as the “context” for one question may be the unit of analysis (the case) for a different question. In other words, these terms are purely formal and have no absolute reference, and this can be confusing for stakeholders. Regardless of the analytic level at which the analysis is couched, we can ask similar questions: What theory are we drawing on? How is it applied? How do we know if the theory is supported (i.e., how likely are we to see the desired result when the hypothesized relationships are present)?
2. When the focus is squarely on evaluating a limited and lower level system theory, the evaluation approach is different as there is generally no comparison of context configurations. For instance, a faculty member seeking to determine whether using video debriefs of IPP team simulation exercises is related to a change in the use of mnemonics in team communication strategies will probably use classic correlational analyses to test this relationship. Using a configurational analysis is unlikely to be of use.

References

1. Cooper, Helen, & Spencer-Dawe, Eileen. (2006). Involving service users in interprofessional education narrowing the gap between theory and practice. *Journal of Interprofessional Care*, 20, 603–617.
2. Hall, Pippa, Weaver, Lynda, & Grassau, Pamela Anne. (2013). Theories, relationships and interprofessionalism: Learning to weave. *Journal of Interprofessional Care*, 27, 73–80.
3. Committee on Measuring the Impact of Interprofessional Education on Collaborative Practice and Patient Outcomes. (2015). *Measuring the impact of interprofessional education on collaborative practice and patient outcomes*. Institute of Medicine. URL: <http://iom.nationalacademies.org/Reports/2015/Impact-of-IPE.aspx> [December 7, 2015].
4. Byrne, David. (2005). Complexity, configurations and cases. *Theory, Culture & Society*, 22, 95–111.
5. Byrne, David, & Callaghan, Gillian. (2013). *Complexity theory and the social sciences: The state of the art*. New York, NY: Routledge.
6. Pawson, Ray. (2013). *The science of evaluation: A realist manifesto*. Los Angeles, CA: Sage Publications Ltd.
7. Ragin, Charles C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond*. Chicago, IL: University of Chicago Press.
8. Patton, Michael. (2010). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York, NY: The Guilford Press.
9. Westhorp, Gill. (2012). Using complexity-consistent theory for evaluating complex systems. *Evaluation*, 18, 405–420.
10. Astbury, Brad, & Leeuw, Frans L. (2010). Unpacking black boxes: Mechanisms and theory building in evaluation. *American Journal of Evaluation*, 31, 363–381.
11. Westhorp, Gill. (2013). Developing complexity-consistent theory in a realist investigation. *Evaluation*, 19, 364–382.
12. Rickles, Dean, Hawe, Penelope, & Shiell, Alan. (2007). A simple guide to chaos and complexity. *J Epidemiol Community Health*, 61, 933–937.
13. Harman, Graham. (2009). *Prince of networks: Bruno Latour and metaphysics*. Melbourne, Australia: re.press.
14. Syväjärvi, Antti, Stenvall, Jari, Harisalo, Risto, & Jurvansuu, Hanna. (2005). The impact of information technology on human capacity, interprofessional practice and management. *Problems and Perspectives in Management*, 1, 82–95.
15. Hawe, Penelope, Shiell, Alan, & Riley, Therese. (2009). Theorising interventions as events in systems. *American Journal of Community Psychology*, 43, 267–276.
16. Glouberman, Sholom, & Zimmerman, Brenda. (2002). *Complicated and complex systems: what would successful reform of Medicare look like?* Canadian Research Index. URL: http://c.ymcdn.com/sites/www.plexusinstitute.org/resource/collection/6528ED29-9907-4BC7-8D00-8DC907679FED/ComplicatedAndComplexSystems-ZimmermanReport_Medicare_reform.pdf [December 7, 2015].
17. Pawson, Ray, & Manzano-Santaella, Ana. (2012). A realist diagnostic workshop. *Evaluation*, 18, 176–191.
18. Mowles, Chris. (2014). Complex, but not quite complex enough: The turn to the complexity sciences in evaluation scholarship. *Evaluation*, 20, 160–175.
19. Craddock, Deborah, O'Halloran, Cath, Mcpherson, Kathryn, Hean, Sarah, & Hammick, Marilyn. (2013). A top-down approach impedes the use of theory? Interprofessional educational leaders' approaches to curriculum development and the use of learning theory. *Journal of Interprofessional Care*, 27, 65–72.
20. Suter, Esther, Goldman, Joanne, Martimianakis, Maria Tina, Chatalalsingh, Carole, Dematteo, Dale J., & Reeves, Scott. (2013). The use of systems and organizational theories in the interprofessional field: Findings from a scoping review. *Journal of Interprofessional Care*, 27, 57–64.
21. Reeves, Scott, Boet, Sylvain, Zierler, Brenda, & Kitto, Simon C. (2015). Interprofessional education and practice guide No. 3: Evaluating interprofessional education. *Journal of Interprofessional Care*, 29, 305–312.
22. Reeves, Scott, & Hean, Sarah. (2013). Why we need theory to help us better understand the nature of interprofessional education, practice and care. *Journal of Interprofessional Care*, 27, 1–3.
23. Gleick, James. (1987). *Chaos: Making a new science*. New York, NY: Viking.
24. Holyoak, Keith J., & Cheng, Patricia W. (2011). Causal learning and inference as a rational process: The new synthesis. *Annual Review of Psychology*, 62, 135–163.
25. Keil, Frank C. (2006). Explanation and understanding. *Annual Review of Psychology*, 57, 227–254.
26. Markman, Arthur B., & Gentner, Dedre. (2001). Thinking. *Annual Review of Psychology*, 52, 223–247.
27. Greenhalgh, Trisha, Robert, Glenn, MacFarlane, Fraser, Bate, Paul, & Kyriakidou, Olivia. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *Milbank Q*, 82, 581–629.

28. Fain, James A. (2005). Is there a difference between evaluation and research? *Diabetes Educator*, 31, 155–156.
29. Wong, Geoff, Greenhalgh, Trisha, Westhorp, Gill, & Pawson, Ray. (2012). Realist methods in medical education research: What are they and what can they contribute? *Medical Education*, 46, 89–96.
30. Rihoux, Benoît, & Ragin, Charles C. (2008). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques*. Thousand Oaks, CA: Sage.
31. Cilliers, Paul. (2010). Complexity, difference and identity. In P. Cilliers & R. Preiser (Eds.), *Complexity, difference and identity* (pp. 3–18). Dordrecht, NL: Springer Netherlands.
32. W.K. Kellogg Foundation. (2006). *W.K. Kellogg foundation logic model development guide*. URL: <https://www.wkkf.org/resource-directory/resource/2006/02/wk-kellogg-foundation-logic-model-development-guide> [December 7, 2015].
33. Barr, Hugh. (2013). Toward a theoretical framework for interprofessional education. *Journal of Interprofessional Care*, 27, 4–9.
34. Committee on Quality of Health Care in America, Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academy Press.
35. Cooper, Helen, Spencer-Dawe, Eileen, & McLean, Elspeth. (2005). Beginning the process of teamwork: Design, implementation and evaluation of an inter-professional education intervention for first year undergraduate students. *Journal of Interprofessional Care*, 19, 492–508.
36. Abramovich, Alex, Espin, Sherry, Wickson-Griffiths, Abigail, Dematteo, Dale, Baker, Lindsay, Egan-Lee, Eileen, & Reeves, Scott. (2011). Translating collaborative knowledge into practice: Findings from a 6-month follow-up study. *Journal of Interprofessional Care*, 25, 226–227.
37. Attack, Lynda, Parker, Kathryn, Rocchi, Marie, Maher, Janet, & Dryden, Trish. (2009). The impact of an online interprofessional course in disaster management competency and attitude towards interprofessional learning. *Journal of Interprofessional Care*, 23, 586–598.
38. Baker, Lindsay, Egan-Lee, Eileen, Leslie, Karen, Silver, Ivan, & Reeves, Scott. (2010). Exploring an IPE faculty development program using the 3-P model. *Journal of Interprofessional Care*, 24, 597–600.
39. Bandali, Karim, Niblett, Bradley, Yeung, Timothy Pok Chi, & Gamble, Paul. (2011). Beyond curriculum: Embedding interprofessional collaboration into academic culture. *Journal of Interprofessional Care*, 25, 75–76.
40. Trojan, Lana, Suter, Esther, Arthur, Nancy, & Taylor, Elizabeth. (2009). Evaluation framework for a multi-site practice-based interprofessional education intervention. *Journal of Interprofessional Care*, 23, 380–389.
41. Egan-Lee, Eileen, Baker, Lindsay, Tobin, Stasey, Hollenberg, Elisa, Dematteo, Dale, & Reeves, Scott. (2011). Neophyte facilitator experiences of interprofessional education: Implications for faculty development. *Journal of Interprofessional Care*, 25, 333–338.
42. Thistlethwaite, Jill Elizabeth, Kumar, Koshila, Moran, Monica, Saunders, Rosemary, & Carr, Sandra. (2015). An exploratory review of pre-qualification interprofessional education evaluations. *Journal of Interprofessional Care*, 29, 292–297.
43. World Health Organization. (2010). *Framework for action on interprofessional education and collaborative practice*. URL: http://www.who.int/hrh/resources/framework_action/en/ [December 7, 2015].
44. Brock, Douglas, Abu-Rish, Erin, Chiu, Chia-Ru, Hammer, Dana, Wilson, Sharon, Vorvick, Linda, Blondon, Katherine, Schaad, Douglas, Liner, Debra, & Zierler, Brenda. (2013). Interprofessional education in team communication: Working together to improve patient safety. *Postgraduate Medical Journal*, 89, 642–651.