

## Telling Context from Mechanism in Realist Evaluation: The role for theory

<p><b>Hannah Jolly</b></p> <p>University of Southern Queensland</p> <p>hannah.jolly@usq.edu.au</p>	<p><b>Lesley Jolly</b></p> <p>Strategic Partnerships Research Consultants</p> <p>ljolly@bigpond.net.au</p>
--	--

**Keywords:** Realist evaluation, context and mechanisms, curriculum evaluation, curriculum model, realist program theory, formal theory in realist evaluation, curriculum model

### Abstract

Realist evaluation is based on the premise that aspects of context trigger particular mechanisms in response to an intervention, which result in observable outcomes. This is often expressed in the formula  $C+M=O$ . Contexts are defined as the conditions that an intervention operates in (often but not exclusively sociocultural), while mechanisms are understood to be the future action that people take in response to the intervention. There is much debate, however, about the definitions and because distinctions are not clear-cut it can be difficult to decide which is which, particularly when the intervention concerns some program of curricular intervention. In this paper we discuss how we resolved this dilemma in an evaluation of a curriculum change in 13 universities in Australia and New Zealand. In that case we found a cascade of contexts and mechanisms, whereby what was a mechanism from one point of view (such as the decisions involved in course design) became a context triggering later mechanisms (such as teacher and student behaviours). The scholarly literature defining curriculum helped us to organise our thinking and subsequent analysis in a rational way, but in many evaluations there may not be a handy body of work that discusses how to understand the topic of the intervention in this way, nor do many consultant evaluators have the luxury of long hours in the library. We consider some ways in which evaluators might decide on defining contexts and mechanisms in principled ways and some of the consequences of those decisions.

## Introduction

The contribution of Pawson and Tilley's (1997) realist approach to program evaluation has constituted a significant shift from available methods. It is most simply understood as a method for evaluating "what works for whom in what circumstances" (Pawson & Tilley, 1997). Rather than focus on global judgements about the worth of a program, it seeks to identify the varieties of success and failure that any program experiences and the factors that contribute to all of the eventual outcomes. The basic premise is that there will be a range of conditions, often sociocultural, that affect the outcomes of any program. These are referred to as Contexts (C). In addition the ways in which people respond – their reasoning about what they should do and the resources they can bring to bear (Pawson & Tilley, 1997, p.67) – will also vary. In the realist approach this is referred to as the Mechanism (M). Hypotheses about how the program results in observed outcomes (O) is often expressed in the formula  $C + M = O$  (CMO). The attraction of this approach lies in the fact that it notes real life programs are rarely entirely successful or entirely unsuccessful, but have patches of success and failure. Also, it is common to find that a program judged to have worked well in one place fails in another or in subsequent years. Realist Evaluation (RE) not only focusses on underlying factors behind outcomes but the various ways in which they can combine and recombine to cause outcomes.

Since its publication, the approach has been widely taken up and applied with varying methodological success (Pawson & Manzano-Santaella, 2012), suggesting that the application of the method is not so simple. Pawson and Manzano-Santaella (2012, p. 176) have now published a discussion of some of the challenges of the "practice on the ground," including the oft expressed problem of "I am finding it hard to distinguish Cs from Ms and Os, what is the secret?" (Pawson & Manzano-Santaella, 2012, p. 188). Whilst their paper discusses this issue in some detail, we will also address the subtleties of this challenge, and attempt to explore their recommendation that "which property falls under which category is determined by its explanatory role" (Pawson & Manzano-Santaella, 2012, p. 187).

### The challenges of applying the realist evaluation approach

Whilst much of the discussion of the difficulties of applying the realist approach is given over to understanding the differences in function between Contexts and Mechanisms, this may be premature if a suitable understanding of the *function* of a CMO configuration as a *whole* is not applied to the process of evaluation. In their 2012 "workshop" on the method, Pawson and Manzano-Santaella (2012, p. 188) emphasise that "the function of CMO configurations...is that they are rather narrow and limited hypotheses, which attempt to tease out specific causal pathways, as pre-specified mechanisms, acting in pre-specified contexts spill out into pre-specified and testable outcome patterns." That is to say, these configurations are sensitive to the actual moment in the intervention process being considered. They need to be used at appropriate times and in appropriate ways during the data analysis if they are to help us to make meaningful evaluations.

In our case, we had an idea of what the intervention was meant to achieve and how it was meant to achieve it, and we began analysis by trying to define contexts and mechanisms directly from the data. When we took this approach we found that it led us in circles. This is because the *function of variables* in a moment of analysis (that is, whether a variable acts as a C (Context) or as an M (Mechanism) is very much dependent on the *focus of explanation* at a given point in the analysis. Something which is a mechanism at one stage of an intervention, such as the reasoning leading to particular decisions about how to design and implement a program, may then produce a fresh context for a later stage, such as the way subjects strategise in response to the program design.

This situation was complicated in the example evaluation by the fact that the program of intervention was taking place in multiple sites, and with differing purposes and methods of implementation in each site. We knew that the focus of explanation needed to vary from site to site, but had not yet pinned down how. Add to this that the program in question concerned a curricular innovation (the notion of curriculum being notoriously slippery), and we quickly discovered that analysis of the data we had collected was creating more questions than answers.

As Pawson and Manzana-Santaella (2012, p. 178) reiterate, “realist evaluation is [or should be] avowedly theory-driven; it searches for and refines explanations of program effectiveness.” While it can be daunting to be told that more theory is needed, in our case it turned out that the theory that helped us to define the specific causal pathways to be investigated was a quite practical one about the nature of curriculum. While this is a highly debated topic, once we had settled on an understanding of what “curriculum” encompasses and how the various elements interact, the evaluation task became much easier.

### **The example evaluation**

The evaluation in question was of a program of curricular innovation that had taken place at a variety of universities across Australia and New Zealand. The program involved the introduction of the Engineers Without Borders (EWB) Challenge into the first year engineering curriculum. The EWB Challenge was conceived as a means of exposing students to the principles of engineering design and problem solving, by providing a design challenge based on the requirements of a real, third-world community who have worked with EWB on sustainable development projects. This program of innovation constituted a “widespread curriculum renewal in engineering education”, because:

The first year in engineering had traditionally focussed on basic science and maths and the introduction of the Challenge and its associated team-based project work allowed for development of the so-called “soft skills” amongst the graduate attributes: communication and teamwork and an understanding of the need for sustainable development. The Challenge has been in operation since 2008 and every engineering school in Australia has made some use of them at one time or another. This [evaluation] project was carried out with the co-operation of 13

universities from Australia and New Zealand who have maintained their use of the projects, albeit in widely divergent types of student cohort and courses. (Jolly, 2014, p. 3)

Thus, the evaluation was seeking to understand both how the program had been applied differently in different sites, and for different purposes, and what contributed to local success and failures. As such, the evaluation was focused on both process and outcome, in that it sought to discover both how the intervention worked, and with what effect. Realist Evaluation (RE) is ideal for this kind of multi-site, multi-context situation where correlations between variables are unlikely to apply in all cases and an understanding of the range of generative causation that can apply is required. In CMO terms, the ideal, desired operation of the intervention could be expressed in a highly compressed form (Table 1).

It needs to be noted that there are dangers in such shorthand representations of CMO configurations (Pawson & Manzana-Santaella, 2012), which we will discuss further below. For now we acknowledge that this hypothesis about how the program should work includes many finer grained levels of CMO configuration. In fact it was the task of the evaluation to find out just what those finer-grained configurations were.

Table 1: **The ideal CMO configuration for the program (based on Jolly, 2014)**

<b>Context (C)</b>	+	<b>Mechanism (M)</b>	=	<b>Outcome (O)</b>
<ul style="list-style-type: none"> <li>• First year engineering curricula emphasise technical and theoretical subjects and pay little attention to practical “real-world” engineering.</li> <li>• Need to develop so-called “soft skills” such as communication and teamwork in engineering students.</li> <li>• Need to respond to widespread concern for environmental issues, especially sustainable development.</li> </ul>		<ul style="list-style-type: none"> <li>• The use of EWB projects with real-world clients will expose students to project-based design work in engineering and this exposure will influence the reasoning of students about how they develop as students and as engineers.</li> </ul>		<ul style="list-style-type: none"> <li>• Students will develop the targeted teamwork and communication skills; start to become familiar with engineering project and design methods; and will learn to incorporate sustainability considerations into engineering design.</li> <li>• These developments will be maintained and built on throughout their time at University.</li> </ul>

With a high volume of data and differing focus and desired detailed outcomes across sites, the task of evaluation was difficult and complex. Fortunately, the data collection process had already been planned and executed on an explicit theoretical basis; that is, on a principled notion of “curriculum”:

Since the topic [of the evaluation] was curriculum change, [we had begun the process] with a review of what could be understood to be included in the term “curriculum” and this resulted in taking a broad scope, including aspects of institutional and course context as well as what happened in the classroom. Program logic analysis was carried out in each site to gain insight into how staff understood what they were doing and how outcomes should be achieved. Data was collected in a broadly ethnographic manner with the help of research assistants recruited and trained at each site. Data sources included documents produced by institutions, staff and students, observations of classroom activities, interviews and focus groups and an online survey for students. (Jolly, 2014, p. 3)

Jolly (2014) discusses the process of basing the data collection in program logic analysis in more detail. Here we are interested in how a theory of curriculum helped us to refine our understanding of the effectiveness of this program by identifying appropriate focusses of explanation within which to define Contexts and Mechanisms.

### **Defining and dealing with the notion of “curriculum”**

Dillon (2009, p. 3) points out that review of the literature on curriculum shows more than 120 definitions of the term, “presumably because authors are concerned about either delimiting what the term means or establishing new meanings that have become associated with it...we need to be watchful, therefore, about any definitions that capture only a few of the various characteristics of curriculum”. He approves the classic definition of Schwab, which is also amongst the more comprehensive in the literature:

Curriculum is what is successfully conveyed to differing degrees to different students, by committed teachers using appropriate materials and actions, of legitimated bodies of knowledge, skill, taste, and propensity to act and react, which are chosen for instruction after serious reflection and communal decision by representatives of those involved in the teaching of a specified group of students who are known to the decision makers. (Schwab, 1983, quoted in Dillon, 2009, p. 343)

This definition begins by saying that curriculum is about outcome, ‘what is successfully conveyed to differing degrees’. Curriculum in this definition also includes the more usual elements of materials, actions, knowledge and skills, but the phrase ‘chosen for instruction’ draws our attention to the fact that the activities that will develop knowledge and skills have to be designed. People with authority and those who ‘represent those involved’ need to make choices about what and how they teach. In order to make those choices they need knowledge of the prospective students.

To recast this statement into a realist form, we could assume that all curricula operate in contexts that begin at the institutional level. Institutions such as schools, training organisations, workplaces and industries set the conditions under which required knowledge and acceptable practice can be defined. As a result of those conditions, people who are designing syllabi and going into classrooms make certain decisions about how to develop the required knowledge and practice in their students. At this point in the curriculum process, mechanisms within the ‘course design’ process – the reasoning about how to use resources – have the power to explain outcomes in relation to the nature of the program. The result of those decisions and resource allocations is a set of learning environments that are the context in which students decide what they will do with the learning opportunities on offer. At this point “Course Design” has the power to explain learning outcomes when considered as a context in which a range of mechanisms are triggered. The outcomes of those mechanisms constitute ‘what is successfully conveyed’ and sometimes that can include things that educators had no intention of conveying. Thus an element of the underlying theory or model of curriculum, such as Course Design, cannot be labelled either as Context or as Mechanism until we know exactly what we are attempting to explain.

### **Defining and dealing with Contexts and Mechanisms**

In the example evaluation discussed here, in attempting to sort out what the focus of evaluation at each site should be, we recognised the need to take a broad theory of curriculum but to organise it in line with Realist thinking. Table 2 shows a map of the curricular landscape that was devised accordingly. Each of the boxes in this map details the relevant focus of explanation at a given point in the curricular landscape. These points could then be examined independently for different sites, according to the data we were dealing with from those sites. While the headings used in the table were derived from the wider literature on curriculum, the descriptors were derived from our data, and the structure of the table reflects our belief in the “generative causation” that is at the basis of all Realist Evaluation (Pawson & Tilley, 1997, p. 56).

This curricular landscape map became an analytical device allowing the data to be unpacked into a single explanatory proposition about how the program of intervention was working in any given site or at any given level (macro to micro). This explanatory proposition is as follows:

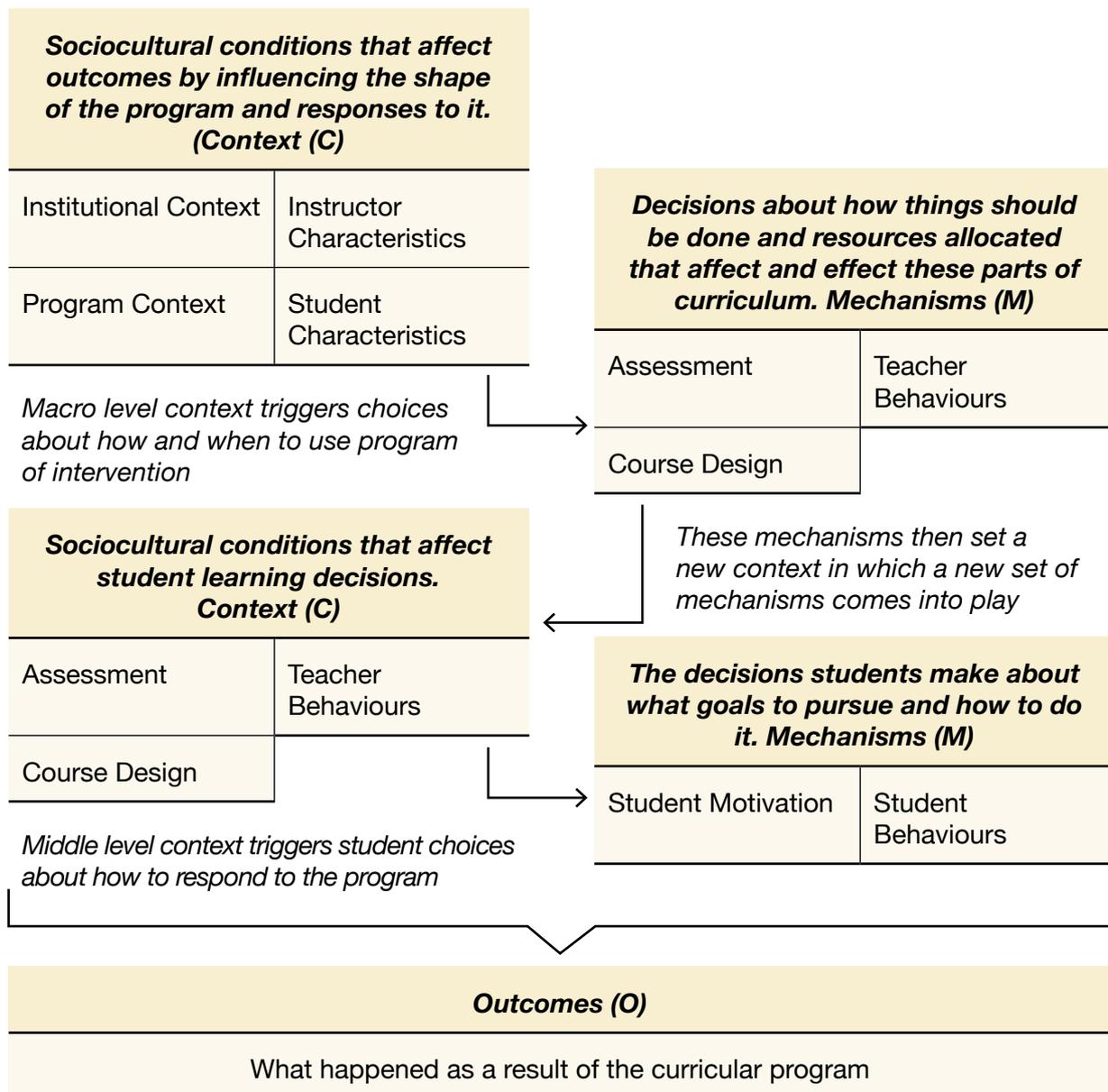
A high level curriculum context pertaining to institutions and programs [triggers] choices people [instructors and administrators] made about how and when to use the EWB challenge projects. The actions that resulted from those choices, such as what pedagogy would be used in the classroom or how assessment would be undertaken, then set a new context in which students made choices about how they would respond to the projects and what was being asked of them. (Jolly, 2014, p. 14)

Table 2: **The Curricular Landscape**

←—————→ <i>Potentially context</i>			
←—————→ <i>Potentially mechanism</i>			
<p><b>Institutional Context</b></p> <p>The institutional factors affecting the way in which the course is implemented</p>	<p><b>Instructor Characteristics</b></p> <p>The experience, beliefs and attitudes of instructors before responding to this instance of the course</p>	<p><b>Teacher behaviours</b></p> <p>How teachers respond to this instance of the course, including decisions, attitudes, interactions</p>	
<p><b>Program Context</b></p> <p>The nature of the course within the program</p> <p>Factors affecting status, purpose and perceptions of the course within its program context</p>	<p><b>Student Characteristics</b></p> <p>The experiences, beliefs and attitudes of students prior to responding to the course</p>	<p><b>Student Motivation</b></p> <p>The factors affecting the kinds and level of effort/ interest, etc., that students put in to the course</p>	<p><b>Student behaviours</b></p> <p>The nature and amount of student participation in the course, including how and how much they focus on topics, processes and products</p>
	<p><b>Assessment</b></p> <p>The nature of assessment tasks, the nature of criteria, the weightings of criteria</p>		
	<p><b>Course Design</b></p> <p>The nature, amount and sequence of organisation of the course, the topic focus, the resources and learning processes it incorporates and the level and nature of prior knowledge it assumes.</p>		
<p><b>Outcomes</b></p> <p>What happened or changed as a result of teacher and student behaviour</p>			

This means that the curriculum map was able to be further unpacked into an evolving set of Cs and Ms, as seen in Figure 1.

Figure 1: **The CMO Cascade**



In one way, this cascade complicates the task of deciding what counts as context and what as mechanism since many aspects of curriculum might be both. Pawson and Manzana-Santaella (2012, p. 187) remind us that ultimately the choice about “which property falls under which category is determined by its explanatory role”. As already noted, any ingredient of the model may operate as either context or mechanism depending on what it is we are trying to explain about the curriculum and how it works. If, for instance, we want to explain the variations in the ways the 13 universities designed their courses, we need to see course design as a set of reasoning, decisions

and allocation of resources (Mechanisms). But if we are examining why some students managed group work well and others did not, we might need to consider the course design as a relevant part of the context. In other words, it depends on whether we aim to explain the next point in the curriculum cascade (in this case the production of a course design) or the ultimate outcomes of the whole program (here a set of learning outcomes). In line with the requirement that realist evaluations need to produce many CMO configurations instead of a catalogue of characteristics (Pawson & Manzano-Santaella, 2012), we need to treat the curriculum model as tool for identifying potential CMO configurations. The model identifies potentially relevant moments in the working of curriculum, but the direction and detail of relevant CMO configurations need to be established empirically. While we cannot go into the whole analysis here (for the final report of the project, see Jolly, 2014) we will take one example to illustrate what we mean.

### **A detailed example**

#### *Identifying what had to be explained*

One of the first steps in the evaluation was a series of program logic interviews with the teaching staff at each of the 13 sites (University of Wisconsin, 2010; Markiewicz, 2010). As well as clarifying objectives for each site, this process allowed us to explore participants' understandings of how the EWB Challenge program was thought to work. This gave us an empirical way to start generating hypotheses in CMO form (although admittedly sketchy at this stage) which also guided our data collection activities. Although many staff identified a better understanding of sustainability as a desirable outcome, most appeared to be relying on the content of the projects alone to bring this about (Jolly, Crosthwaite, Brodie, Kavanagh & Buys, 2011). This led us to pay attention to what students learned about sustainability and how they learned it, but in the process of investigating that, our attention was drawn to a wide range of student goals and motivations – what they wanted to get out of the course – and how those things affected learning outcomes differently in different contexts. Sustainability-related mechanisms are separated from other student-goal related mechanisms in our final analysis because participating universities had requested information specifically about the sustainability goal.

Data were gathered through a range of techniques including interviews, participant observation, surveys and documentary analysis. Recurrent themes or category of event were identified through the constant/comparative method (Richards, 2005) and labelled using *in vivo* category labels. Categories of Cs and Ms were able to be grouped according to whether they were observed to have a positive or negative impact on the overall outcomes identified in the program logic interviews. Certain contexts were deemed to enable positive mechanisms, while others disabled them. Mechanisms either supported the attainment of the desired outcomes or they inhibited them (Sochacka, 2011). They were also grouped by similarity into higher level clusters which relate to both wider educational theory and observed reality and could be expected to apply

across contexts. For instance, a set of contexts relating to the ways in which student actions were driven by the ways in which the course presented activities and goals were grouped together under the Context *C2 Alignment of assessment with learning goals*. This is the well-known principle of constructive alignment (Biggs, 1996). The category level factors illustrate what the principle looks like for this particular set of interventions and the labelling of ‘clusters’ allows comparison with other educational debates and thus transferability of the findings. The entire analysis is presented in diagrammatic form in Figures 2 and 3. From these figures a potentially very long list of CMO configurations could be generated and workshops were run with all participants to help them work through their own cases. The most common hypothesis held by educators about the sustainability outcome is represented in Table 3.

Table 3: **How the program delivered improved understanding of sustainability**

<b>Context (C)</b>	+	<b>Mechanism (M)</b>	=	<b>Outcome (O)</b>
<ul style="list-style-type: none"> <li>Course projects were real-world projects (C5) with real clients from developing nations such as Cambodia.</li> </ul>		<ul style="list-style-type: none"> <li>Students reason that these communities need help and that engineers have a responsibility to come up with design solutions to identified problems which are appropriate to the culture and environment and able to be operated by local communities beyond initial implementation.</li> </ul>		<ul style="list-style-type: none"> <li>Sustainability is automatically built into the project designs AND students acquire attitudes that take account of client need and project specificity in always striving for a sustainable solution.</li> </ul>

Figure 2: Clusters and Categories of Context

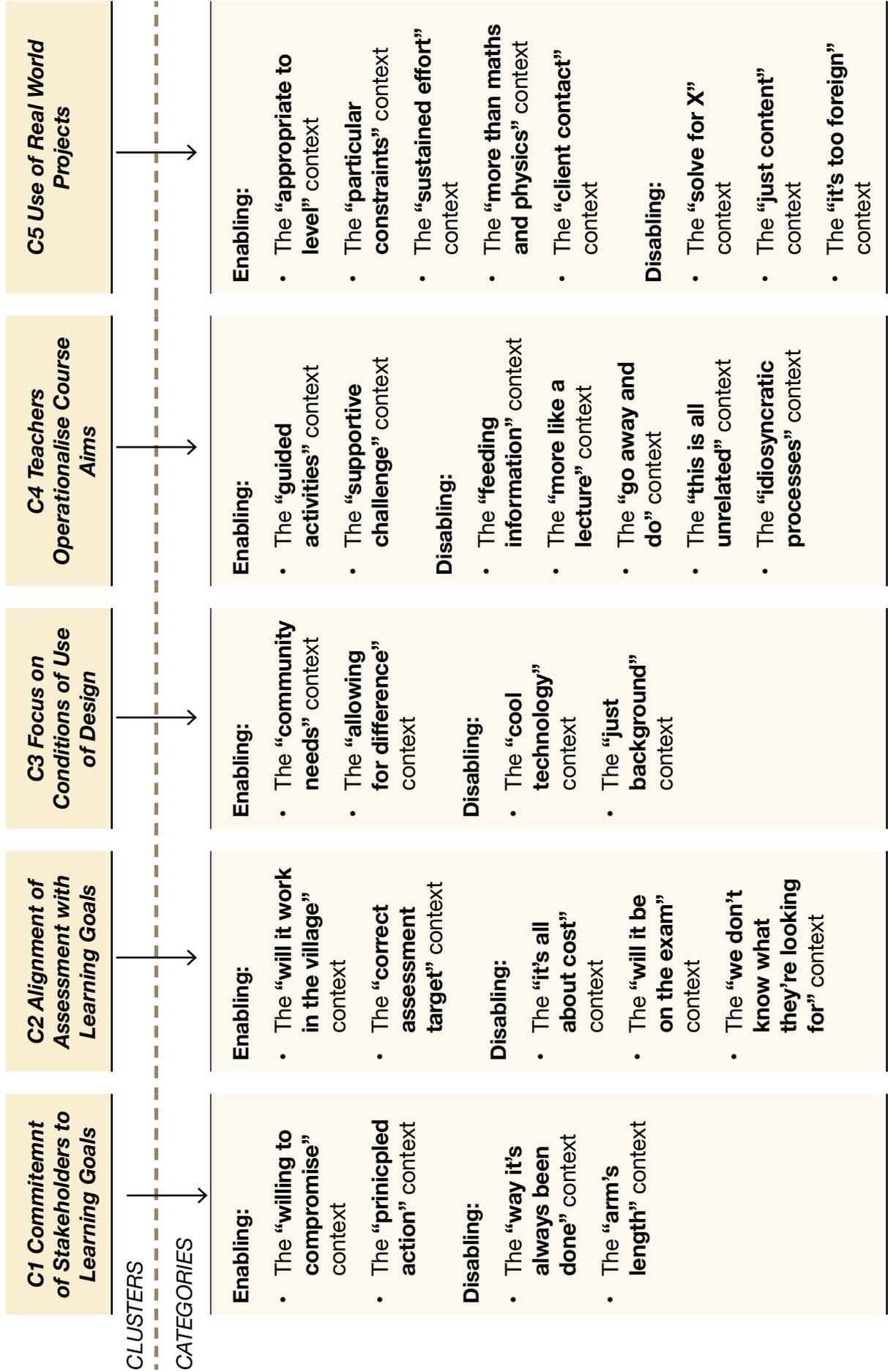
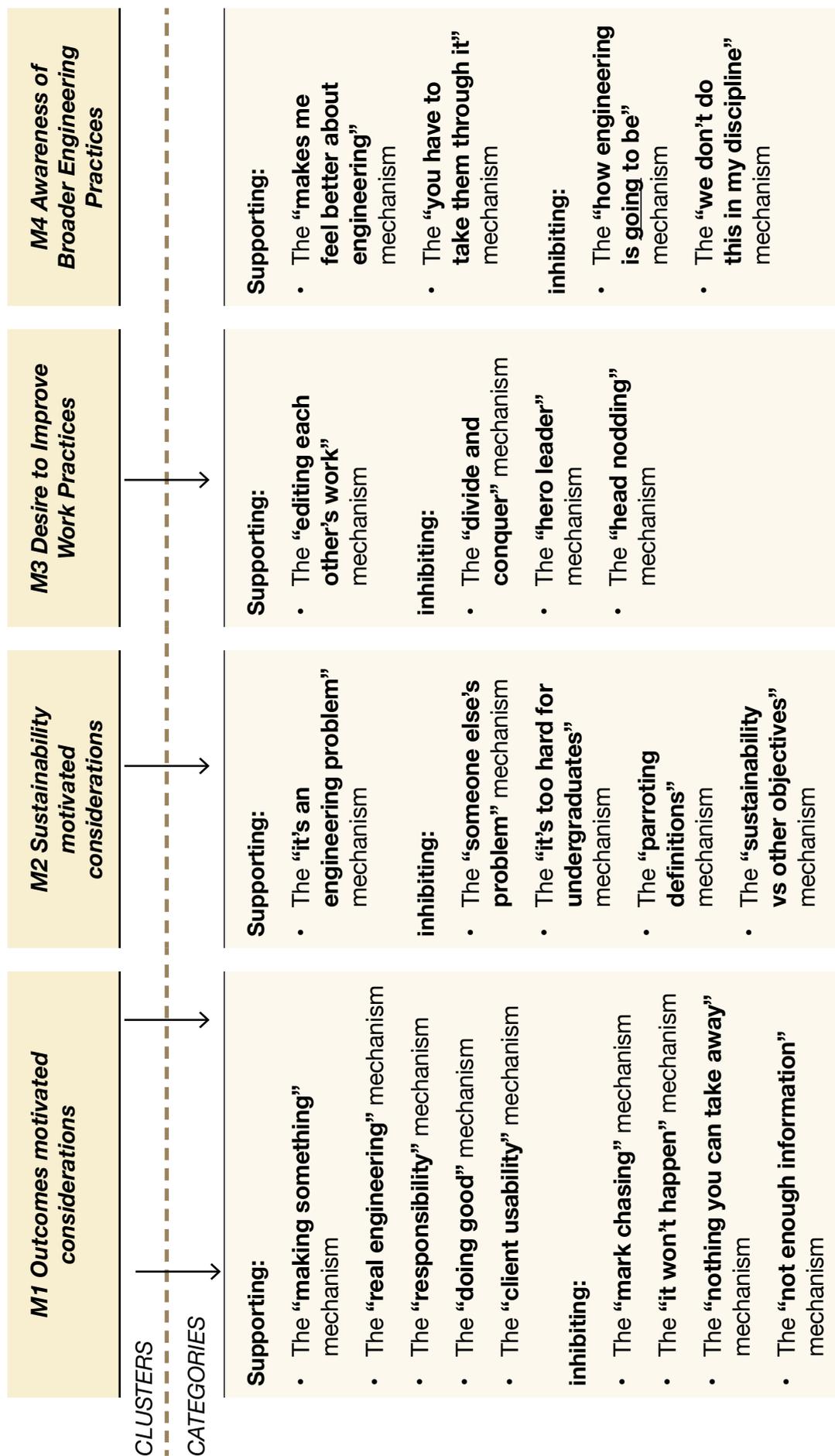


Figure 3: Clusters and Categories of Mechanisms



### The nature of the data

We examined a wide range of data in considering the merits of the hypothesis represented in Table 3 about how this intervention improved students' understandings of sustainability, but we have room for only a very short selection of that data.

Firstly, we note that although course designers and instructors were inclined to include sustainability in their aspirations for student work, it was not always made explicit in course learning objectives and classroom activities. Where published course objectives did mention sustainability, it was more often than not in terms such as, "Appreciate the socio-cultural, political, environmental and economic contexts in which engineering is practised". Thus, it did not seem that students were expected to gain a very deep understanding of sustainability, and the most explicit classroom discussions of the concept we observed were reiterations of standard definitions such as that from the Brundtland Report: "Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs" (World Commission on Environment & Development, 1987) without extended discussion of the concept. In fact one expert in the field told us that sustainability was too hard to teach to undergraduates.

Nevertheless, when asked in a survey to rate their confidence in accounting for issues of sustainability in their decision-making 81% of students professed themselves confident or very confident. At interview, however, most students we spoke to found it hard even to reiterate the Brundtland definition and offered formulations such as:

Before I only knew about environmental sustainability and I thought that that was the only sustainability there was, but now I know the triple bottom line thing and that it has to be socially sustainable as well, and financially and environmentally, because you can't just make something environmentally sustainable and make it really, really expensive because people aren't going to use it. (Focus group, Go8A)

There was evidence of superficial approaches to sustainability in the project reports also, with many students reiterating the same stock phrases in the background section of the report and then not appearing to pay very much attention to it in the detailed design solution. We also observed in classroom observations that tutors and instructors were inclined, perhaps not surprisingly, to emphasise more technical and traditional aspects of the project, as reflected in the following exchange between an observer and a student:

"Researcher: Why did you choose to build in material with a high embodied energy?"

Student: Yeah well we should think about sustainability and all that but it's all about cost really." (Observation notes, Go8B)

Students appeared to be strongest in their understanding of and commitment to sustainability where they could see it as a legitimately engineering (that is, technical) problem.

The observed outcome therefore seemed to be that at least some students learned something about sustainability but applied no critique to the concept (and therefore may not be able to transfer it to other problems) and may be inclined to consider it something to think about after all other factors (such as cost) were served.

### **Deciding the context/mechanism question**

Empirical investigation not only articulated actual outcomes but also described some of the reasoning (on behalf of both students and instructors) that produced them, and described the contexts (socio-cultural conditions) in which they happened. However, Contexts and Mechanisms cannot just be observed, and different kinds of theory are needed to help us in analysis. Here we had a number of different theories in play for different purposes.

First, there were the participants' theories of how the intervention should work and this is important for explaining why they did the things they did and for indicating potential reasons for success and failure. With this in mind, we used in vivo labels for our categories to capture something of this level of theory, although this presented some difficulties. For instance, we saw many instances of classroom teaching, tutoring and resources which de-emphasised sustainability concerns in favour of factors such as cost. We chose to label these situations the "all about cost" context, although it could be argued that when a student tells us that, it is evidence of their reasoning about how to respond and thus should be a mechanism.

However, we also saw that almost universally students went through a process of weighing up sustainability against other objectives and could speak fluently about that process. We therefore chose to identify the "sustainability vs other objectives" as a significant mechanism producing the observed outcomes. This allows for a challenge to and modification of the participant theory represented in Table 3 that the project setting alone will result in attention being paid to sustainability. At the same time, our choice of context and mechanism was driven by higher order theories of curriculum and teaching such as Biggs' (1996) constructive alignment, which states that all of the elements of curriculum from course design and objectives to classroom activities to assessment criteria need to be coherent to provide the desired result.

Each university could now consider their use of the EWB projects according to the contexts and mechanisms in operation there. Table 4 represents just two possibilities.

### **Implications**

While our theory of curriculum derived ultimately from educational theory, it matched well with the views of stakeholders revealed through program logic analysis, about how the intervention ought to work in their setting. However, program logic approaches (Chen, 1990; Den Heyer, 2001; Rogers, Petrosino, Heubner & Hacsí, 2000) and the

closely related program theory formats (Funnel, 2000; Shadish, Cook & Leviton, 1991; Weiss, 1996) tend to concentrate on inputs and outputs and on identifying moments when a project might be monitored, rather than making explicit the factors that people involved understand to be influential in making the project work or not. It is this understanding that our model of curriculum embodies; it includes every factor that was mentioned in literature and by participants as part of the operation of the curriculum, and it thus provided a set of focusses which helped to identify contexts and mechanisms, and where and how they applied. The implication is then that such a model of operation of the intervention can be helpful to any CMO analysis. As we have shown, it does not however necessarily help in the allocation of phenomena to either context or mechanism.

Table 4: **Actual CMO configurations (CMOCs) with respect to sustainability**

<b>Context (C)</b>	+	<b>Mechanism (M)</b>	=	<b>Outcome (O)</b>
<ul style="list-style-type: none"> <li>Assessment and classes are presented in such a way as to reward results that have little to do with sustainability (the <b>“all about cost”</b> context C2) and which represent the engineer’s job as devising new widgets (the <b>“cool technology”</b> context C3).</li> <li>Assessment weights sustainability against other criteria appropriately (the <b>“correct assessment target”</b> context C2) and engineering is represented as <b>“more than maths and physics”</b> context C5.</li> </ul>		<ul style="list-style-type: none"> <li>Students allocate time and effort to acquiring technical skills (the <b>“real engineering”</b> mechanism M1) before those associated with sustainability (the <b>“sustainability Vs other objectives”</b> mechanism M2).</li> <li>Students allocate time and effort to understanding and incorporating sustainability considerations (the <b>“its an engineering problem”</b> mechanism M2).</li> </ul>		<ul style="list-style-type: none"> <li>Superficial treatment of sustainability considerations.</li> <li>Realistic understanding of role and requirements of sustainability in design.</li> </ul>

## Conclusion

What stands out most to us from Pawson and Manzano-Santaella's (2012) discussion of the difficulties of applying the realist approach, in light of our own experience of this method, are the following issues. First, formal theory has a role to play in refining the program theories held by participants and exemplified as a CMO configuration (CMOc) in Table 3 above. Formal theory can help articulate *what* is to be explained (Tilley, 2009). In our case a theory of curriculum helped us focus on aspects of process as well as content. However, using this theory in concert with participants' own theories allowed a methodological flexibility for understanding the data (and how it functions in a causal CMOc) according to different moments of analysis. Thus, the understanding of instructors about how improved sustainability considerations could come about were influential in our decision to call "all about cost" a context rather than a mechanism. It allowed us to draw attention to contextual factors they could change fairly readily. In short, we agree with Pawson and Manzano-Santaella (2012, p.189) that Realist Evaluation needs program theory but we have argued that formal theory can help us make decisions about where in the CMO sequence any particular phenomenon fits. However, they are also right to point out that "programmes do not come in pre-ordained chunks called contexts, mechanisms and outcomes" (Pawson & Manzano-Santaella, 2012, p. 189) and it is not helpful to use formal theory to make it seem as though they do.

## Acknowledgements

The evaluation discussed in this paper was a project of the Australian Learning and Teaching Council (PP2010-1647) called *Curriculum Renewal in Engineering Through Theory Driven Evaluation*, headed by Professor Caroline Crosthwaite of the University of Queensland. We also wish to thank two anonymous reviewers for thoughtful comments which helped refine our thinking.

## References

- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347-364.
- Chen, H. (1990). *Theory Driven Evaluation*. Newbery Park: Sage.
- Den Heyer, M. (2001). *The Temporal Logic Model*. Ottawa, Canada: International Development Research Centre.
- Dillon, J. T. (2009). The questions of curriculum. *Journal of Curriculum Studies*, 41(3), 343-359.
- Funnell, S.C. (2000, Autumn). Developing and using a program theory matrix for program evaluation and performance monitoring. *New Dimensions for Evaluation*, 2000(87), 91-102.

- Jolly, L. (2014). *Curriculum renewal in engineering through theory-driven evaluation*, Sydney: Office for Learning and Teaching. Available at: <http://www.olt.gov.au/resource-curriculum-renewal-engineering-through-theory-driven-evaluation>
- Jolly, L., Crosthwaite, C., Brodie, L., Kavanagh, L., & Buys, L. (2011). *The impact of curriculum content in fostering inclusive engineering: data from a national evaluation of the use of EWB projects in first year engineering*. Paper presented at the National Conference of the Australasian Association for Engineering Education, Fremantle, Western Australia.
- Markiewicz, A. (2010). *Monitoring and Evaluation Core Concepts*. Professional Development materials used in training workshops for the Australasian Evaluation Society.
- Pawson, R. (2006). *Evidence-based policy: A realist perspective*. London: Sage.
- Pawson, R., & Manzano-Santaella, A. (2012). A realist diagnostic workshop. *Evaluation*, 18(2), 176-191.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: Sage.
- Richards, L. (2005). *Handling Qualitative Data*. London: Sage.
- Rogers, P.J., Petrosino, A., Heubner, T.A., & Hacsí T. A. (2000, Autumn). Program theory evaluation: practice, promise and problems. *New Dimensions for Evaluation*. 2000(87), 5-13.
- Shadish, W., Cook, T.D., & Leviton, L.C. (1991). *Foundations of Program Evaluation*. Newbury Park: Sage.
- Sochacka, N. W. (2011). *Realistic Analysis of Socio-Technical Interventions in the Context of Urban Water Management*. PhD dissertation at the University of Queensland, Australia.
- Tilley, N. (2009). What's the "What" in "What Works? Health, Policing and Crime Prevention. *Crime Prevention Studies*, 24, 119-143.
- University of Wisconsin. (2010). *University of Wisconsin-Extension, Program Development and Evaluation Model*. Retrieved from <http://www.uwex.edu/ces/pdande/>
- Weiss, C.H. (1996). Theory based evaluation, past, present and future, Paper presented at the Annual Meeting of the American Evaluation Society, Atlanta, CA.
- World Commission on Environment & Development, (1987). *Our Common Future*. Oxford, England: Oxford University Press.

