0084 Between analysis and transformation: technology, methodology and evaluation on the SPLICE project

Authors

Mark Johnson, Peter Lager, Bill Pollard, Graham Hall, Miranda Edwards, Pete Whitfield, Rupert Ward

University of Bolton, Deane Road, Bolton BL3 5AB

mwj1@bolton.ac.uk

Abstract

This paper concerns the ways in which technological change may entail methodological development in e-learning research. The focus of our argument centres on the subject of evaluation in e-learning and how technology can contribute to consensus-building on the value of project outcomes, and the identification of mechanisms behind those outcomes. We argue that a critical approach to the methodology of evaluation which harnesses technology in this way is vital to agile and effective policy and strategy-making in institutions as the challenges of transformation in a rapidly changing educational and technological environment are grappled with.

With its focus on mechanisms and multiple stakeholder perspectives, we identify Pawson and Tilley's 'Realistic Evaluation' as an appropriate methodological approach for this purpose, and we report on its use within a JISC-funded project on social software, SPLICE (Social Practices, Learning and Interoperability in Connected Environments). The project created new tools to assist the identification of mechanisms responsible for change to personal and institutional technological practice. These tools included collaborative mind-mapping and focused questioning, and tools for the animated modelling of complex mechanisms. By using these tools, large numbers of project stakeholders could engage in a process where they were encouraged to articulate and share their theories and ideas as to why project outcomes occurred. Using the technology, this process led towards the identification and agreement of common mechanisms which had explanatory power for all stakeholders.

In conclusion, we argue that SPLICE has shown the potential of technologically-mediated Realistic Evaluation. Given the technologies we now have, a methodology based on the mass cumulation of stakeholder theories and ideas about mechanisms is feasible. Furthermore, the summative outcomes of such a process are rich in explanatory and predictive power, and therefore useful to the immediate and strategic problems of the sector. Finally, we argue that as well as generating better explanations for phenomena, the evaluation process can itself become transformative for stakeholders.

Introduction

E-learning research as a subset of educational research has adopted many of the methodological approaches of the social sciences towards evaluation. Typically these tend to be characterised by evidence-based approaches to identifying project performance against indicators, outcomes and baselines (JISC, 2007). Despite an increasing awareness of the importance of continual evaluation throughout a project's duration—particularly with regard to agile and iterative projects, such approaches remain close to the classic model of quasi-experimental research and evaluation drawn up by Campbell and Stanley (1963). Such techniques rely on traditional social science methods of data collection including questionnaires, focus groups and interviews, with little use made of specific technological innovations to support the evaluative methodology.

The situation is different in fields outside e-learning, notably Operational Research, where much work on methodological development in evaluation and research has been conducted in recent years, with not only concepts and techniques but new technologies arising. These have included various techniques for problem specification deriving multiple stakeholder viewpoints from Soft Systems (Checkland, 1990) approaches (for example, the Strategic Options Development and Analysis (SODA) technique which is designed to aid participative problem definition with the use of modelling software (Eden and Ackermann, 1989)). Different kinds of software innovation helped Beer's Team Syntegrity (1994) take an alternative approach to Operational Research which seeks to gain consensus about problem definitions and solutions through guiding stakeholders through a highly structured set of activities to explore different viewpoints. Mingers argues that these techniques can be combined using multi-methodological approaches to deal with highly complex problems on condition that the different world-views associated with each technique are surfaced and evaluated (Mingers, 2004).

This paper argues that technology creates not only a context for research in e-learning, but also a context for understanding outcomes. As such, technological development necessitates methodological development as institutions seek better ways to understand and adapt to the transformational effects that technological and social change has on them. To this end, we report on the use of a methodologically innovative approach to evaluation, Realistic Evaluation (Pawson and Tilley, 2002), on a JISC-funded project on social software, SPLICE (Social Practices, Learning and Interoperability in Connected Environments), and on the technological innovations that accompanied the use of this methodology.

The SPLICE project and realistic evaluation

The SPLICE project focused on the ways individuals and institutions change their technological habits with social software, and the things that can be done to effect these changes. Over the course of 18 months, learners, teachers and industrial partners in the creative industries were exposed to

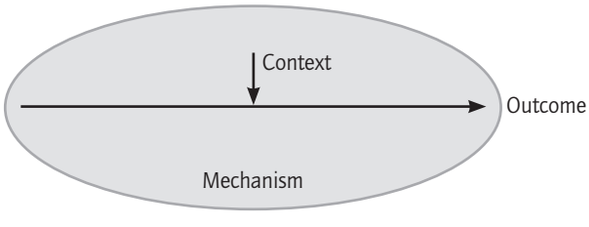
various interventions to encourage adoption of online social habits. The interventions included the use of social forums, micro-blogging, and specific learning activity designs. The project was focused on those interventions which produced change in which circumstances. In some cases, the interventions 'worked' (e.g. learners or teachers transformed their habits); in some, they didn't. In most cases, the interventions can only be said to have 'sort of' worked. In addition to identifying what worked and where, the evaluation was concerned with identifying 'why' things did and didn't work, and particularly what 'sort of' working meant. This entailed an ongoing process of modelling outcomes throughout the project, with the ultimate aim of producing realistic models of change in habit which could then be used by institutions to plan strategy and policy with social software.

Realistic Evaluation was chosen as a methodology to support this. It asserts that discoverable mechanisms are responsible for social phenomena, and that better knowledge of these mechanisms can give greater control to practitioners, whether teachers, administrators or learners. In asserting the role of mechanisms in the social world, Realistic Evaluation is rooted in the philosophy of Critical Realism (Bhaskar, 1977). Pawson and Tilley argue that the job of evaluation is to uncover those mechanisms through a process which they (following Bhaskar) call 'retroduction'. In essence, retroduction involves describing the context within which a possible mechanism might be responsible for producing a particular outcome. The relationship between Context, Mechanism and Outcome is shown in Figure 1. In line with the Critical Realist position, Pawson and Tilley argue that, whilst the experience of a project to any particular observer (or stakeholder) might be different (or relative to the observer), those experiences are not *that* different. In other words, they may be the product of a common mechanism working within each individual context. Thus, in encouraging individual participants to articulate the mechanisms that they feel to be responsible for what they experience, it may be possible to consider overarching explanatory frameworks which describe mechanisms which are common to each. Such overarching mechanisms can then be considered for their explanatory and predictive power with regard to each individual outcome.

In SPLICE, the Realistic Evaluation approach was geared around identifying the causal mechanisms behind the impact of the interventions on the different stakeholders in different contexts in the project. With the identification of these mechanisms, institutions could be equipped with models that would predict the likely consequences of interventions around social software, and thus be in a better position to guide policy and strategy. In other words, the 'value' of the SPLICE project could be realised in the form of knowledge that was practically useful to other institutions.

The practicalities of identifying mechanisms between the large variety of stakeholders on the project was challenging. In addition to basic questions like "how can a mechanism be captured or expressed?" or "how can common mechanisms be agreed?", there were organisational problems concerning how all the different stakeholders together with the variety of different project activities could be represented. Unlike Operational Research techniques, Pawson and Tilley's Realistic Evaluation has not given rise to particular technologies. However, given that the central approach

Figure 1: Context, mechanism and outome in Realistic Evaluation



of the methodology is the identification of mechanisms, and its emphasis is on multiple stakeholder engagement, technology would appear to have something to contribute. In SPLICE technology, as well as being the object under investigation, also proved to be an important factor in the evaluation process, with specific tools developed to aid the realistic evaluation process.

The SPLICE context

Realistic Evaluation relies on sharing the contexts and outcomes from multiple perspectives. These different contexts and outcomes reflect the variety of stakeholder perspectives on a project. Like most learning technology projects, there were a large number of stakeholders in SPLICE, which included:

- Technical developers.
- Project managers.
- Teachers.
- Accounting managers.
- Institutional administrators.
- Funding body programme managers.
- Creative Technology practitioners.
- Learners.

The stakeholders within SPLICE had different experiences of it. Those learners with whom the project interventions 'worked' reported significant personal transformations. These were, however, the minority. For most, the picture was more complicated, with some continuing to feel uncomfortable with social software, and others 'dabbling' without feeling they wanted to engage at a deep level. Amongst other stakeholders, institutional administrators varied in their experiences of the project, from simply managing the project money, to identifying key synergies between project outcomes and institutional objectives. Individual teachers varied in their experiences, from overcoming reticence to engage in new technologies, to transforming their teaching practices. Despite the variety of these experiences however, common patterns of experience were discernable. The purpose of the realistic evaluation approach was to elicit the nature of these common patterns between the groups of different stakeholders.

From the Realistic Evaluation perspective, each stakeholder could report an outcome (or a number of outcomes) from the project. With each outcome, a context for that outcome could also be established. Given these reports of outcomes, and identification of contexts, the evaluative task was to explore possible mechanisms which might be responsible in each case. The starting point for identifying mechanisms was to ask the stakeholder "what do you think is going on?". Sometimes, this would produce interesting results, as individual stakeholder theories could be gathered, compared, shared and tested. At other times, theoretical explanations could be presented to the stakeholders to see if they felt that such explanations were meaningful to them.

This process of engaging stakeholders and encouraging theorising continued throughout the project. The project was designed around an iterative action research cycle of Plan → Do → Reflect → Review (similar to that articulated by Argyris and Schön (1974)). This regular theorising helped to build models of "what's going on?" as the project progressed. These models were then used to design the next cycle of interventions. As a starting point, the project began with some 'ideal' models and mechanisms which were used to plan the initial interventions. These were largely drawn from work on the Personal Learning Environment (Johnson and Liber, 2008), and involved a model of personal organisation drawn from Stafford Beer's Viable System Model (1981). In the light of the results of actions taken against the backdrop of this model, both the model and action plans were modified. One advantage

of the iterative method was that, with many project partners, it was often appropriate to let individual partners shape their contribution to the project in the light of local conditions within their institutions. Within the Realistic Evaluation framework, this was simply to provide different 'contexts' for actions, and to report on outcomes observed in the light of those contexts. This process was repeated over the course of the project.

As the project progressed, low-level mechanism descriptions and theories needed to be cumulated into over-arching explanatory frameworks. For this, technology became an important means by which complex mechanisms could be articulated, and stakeholder opinions shared. Through the use of tools, a higher level of synthesis and cumulation of experiences, results, ideas and theories about "what was going on" could be shared and explored.

SPLICE tools for evaluation

In this second stage of the evaluation, stakeholders were consulted together and led through a process of sharing and exploring project outcomes and mechanisms. The basic structure of this process was:

1. Focused questioning to explore principle challenges.
2. Interactive modelling to explore possible explanatory mechanisms.
3. Reflection in the light of models, and repeating the process from 1.

These stages tended to be quite large-scale affairs owing to the number of stakeholders, and the initial step was conducted over the course of an 'evaluation day' with most of the project partners gathering to share their ideas. Owing to the large number of participants, and variety of different stakeholder roles, the coordination of this process required technological mediation. For a), a tool was developed for collaborative mind-mapping which allowed many participants to contribute their thoughts and theories in response to a particular question and then to share and reflect and ultimately vote on which question would then follow. To do this, an existing open-source mind-mapping tool was adapted to allow it to receive input from Twitter (www.twitter.com), a popular message sharing service.

The technology afforded agility and capacity to record and organise the views of stakeholders which 'low-tech' approaches to collaboration and brain-storming (for example, the use of post-it notes) did not. This capacity to organise allowed for the structured drilling-in on particular emerging themes, which aided consensus-building within the stakeholder group.

The basic design of the use of the technology was:

1. Allow for structured input of opinion and thoughts by stakeholders in a way which captured as much as possible.
2. Enable synchronous participation in the process for those not physically present
3. Allow for coordination and steering of the discussion towards consensus between stakeholders.
4. Through using other mind mapping tools, provide mechanisms for sharing the resulting mind map.

The results of this exercise and details of the technology are described in more detail below.

The second stage of the evaluation at 2 involved the creation of software to bring possible mechanisms 'to life' so that all stakeholders could grasp the emerging possibilities for mechanisms that might explain both their and others' experiences. This software allowed for the creation of dynamic processes which could relate both to the ideas, theories and categories emerging from the mind mapping exercise at 1 and to broader established theoretical constructs from the social sciences. These mechanisms were

interactive, allowing participants to explore the results of particular actions. This deeper level of engagement by stakeholders served to lead conversations about causal mechanisms for the project to a deeper and more focused level. For example, where in the mind map exercise, a distinction was made by one stakeholder between different types of people within institutions ("energy creators"—the instigators of innovation, "energy neutrals"—those who are receptive to innovation, "zappers"—those who actively resist, and sometimes sabotage, innovation), using the modeller, more probing questions emerged ("what *exactly* happens when a zapper is introduced to a new idea?")

Collaborative mind mapping through Twitter and FreeMind

The mind-mapping exercise itself was conducted over the course of the SPLICE 'evaluation day' and was used to hone-in on possible mechanisms through an iterative 3-stage process over the course of the day:

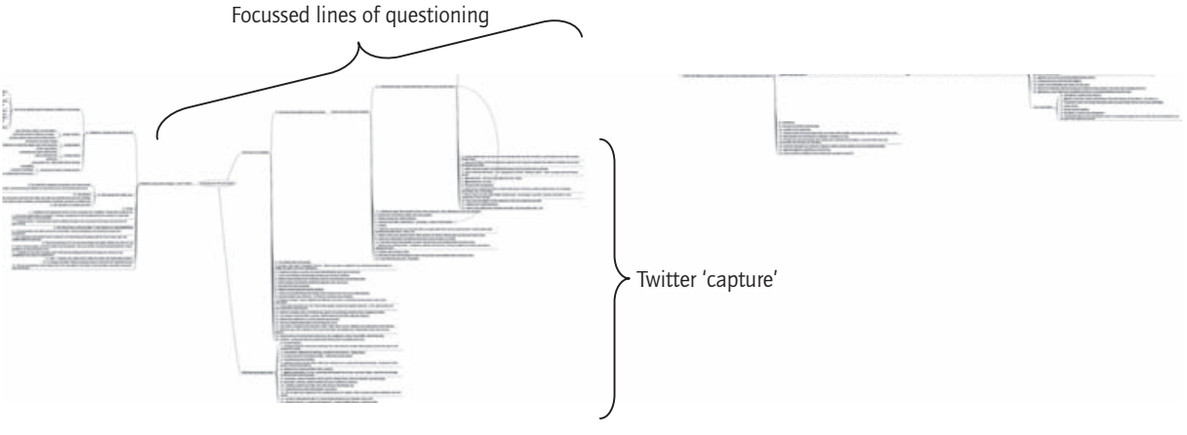
1. brain-storming and capturing possible answers to a question;
2. reflecting on results and voting for most effective answers;
3. drilling into chosen issues and repeating the process.

The process was repeated many times during the course of the day, coordinated by a facilitator whose job it was to ensure fair representation of all stakeholder views. Stakeholders submitted ideas and voted through Twitter.

By the end of the day, this exercise resulted in a large mind map whose basic structure can be seen in Figure 2. From this structure the questions that emerged from participant feedback can be identified as the successive roots of previous feedback. The initial starting questions were "How have you changed in your technological habits over the course of the project?" and "How have your institutions changed over the course of the project?" The day was divided between exploring these two questions. To give an insight into how the technology worked, it is useful to demonstrate how the questioning developed during the day, and how the results of the questioning fed into the deeper identification of mechanisms.

The initial responses to the first questions included positive and negative comments from those present, depending on their experiences. On the positive side, some reported that they had changed through "following other professionals on Twitter" or being "more willing to let students dictate the agenda in the classroom" or by "connecting real life practice with the online environment". On the negative side, some worried that "technophobes were getting left behind", or were concerned about an "over-dependence on technology". After capturing responses to the initial question, all stakeholders reflected on the responses gained and a vote was taken to decide which of these different responses would be pursued at the next

Figure 2: The basic structure of the twitter-generated mind map session



iteration of the investigation. The top-ranked 'indicator of personal change' was the realization that "I became more relaxed about what I put online". This was then pursued by repeating the data-gathering exercise and asking about the causes of this 'increased relaxation', or indeed what it meant.

Answers to this revolved around the emerging realisation that there was a large community of practice engaged in online social activity, with an increasing awareness that participation in online activity was an indicator of the social capital of an individual ("starting to judge other people by their online exposure"). The top-rated response in this iteration was that increased relaxation in putting things online was due simply to "realizing the value of online engagement".

This raised the issue of "what is the value and when do you see it?", since identifying 'value' appeared to be the principle cause for engaging with the technology. The iteration under this question produced responses suggesting that value lay in getting feedback and building relationships online. As the questioning progressed, theories were suggested, and at this point, there was an interesting correspondence with Bhaskar's Transformational Model of Social Activity where it was discussed whether the emergence of social connectivity online drives the increase in online habit. Such theoretical correspondences led themselves to deeper consideration in the second stage of the evaluation. Deeper discussion at this point led to the consideration that what was 'valuable' was often what was *not* put online, and the group questioning continued down this path. This raised the question of the distinction between that which is deeply personal and that which people are happy for others to see, and following this, the question of whether the boundary between 'public' and 'private' life is changing in the light of technology. In turn, the differences between those who are disposed positively towards technology and those who aren't became the focus of the next iteration.

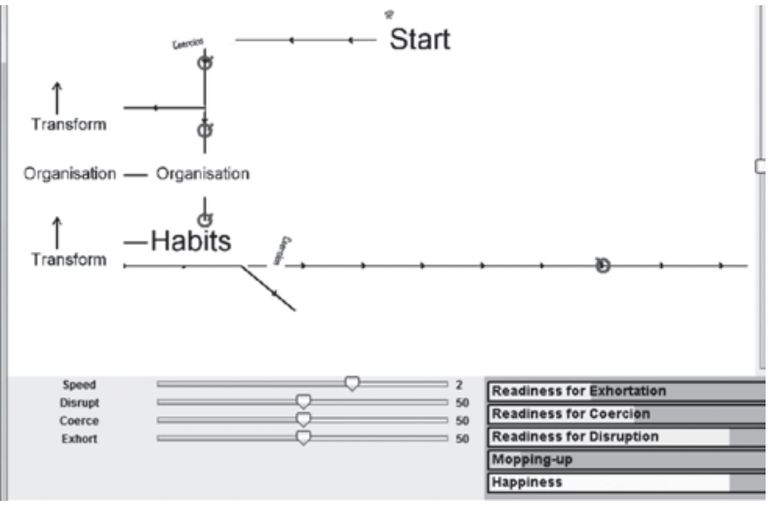
As the process continued and the territory of the questioning became deeper, the mind map facilitated navigation back to where the questioning had started, and so helped make connections between the deep level (and increasingly theoretical) discussion, and the basic questions that it had begun with. As things progressed, the questioning focused on understanding the 'relevance' (as opposed to the 'value') of technology, together with the variation in the individual ability to change habits in the light of new developments. These issues of personal difference distilled to the differences between individuals who explored future scenarios in the light of new technological developments, and those who detected threats in technology to personal life. Finally, this led to a focus on the mechanisms whereby individuals organize themselves, with differentiation between those for whom priority was given to 'future gazing' and experimentation, and those who sought to remain in touch with embodied human experience and felt the need to 'protect' it from technology. This led to a discussion around the fact that the discussion itself was part of what technology does: that whether technology does or doesn't work; whether users like or dislike it, there is still something to talk about. These discussion further suggested possible mechanisms, with some relevance to different mechanisms of the personal organization suggested by Harré (1984), Beer or Archer (2000).

Exploring possibilities: animated modelling of mechanisms

Stage 1 of the evaluation resulted in rich descriptions of things which happened to people, but did not go so far as to suggest common possibilities for mechanisms which might produce this. All stakeholders had at least some opinions about why things happened (some had more fully worked-out theories). These theories and opinions were developed to different levels of sophistication. Thus the task began to identify those theories

which would fit the outcomes and mechanisms described by the mind-mapping. However, the risk with taking this approach is that a particular theoretical standpoint can be privileged over the real experience of stakeholders. Within the SPLICE evaluation process, since all theories are effectively descriptions of mechanisms, there was clearly a role for some technology to make the action of mechanisms more apparent and understandable so that all stakeholders could relate their experience to the mechanisms.

Figure 3: Inner State mechanism



A number of options are available for dynamic animated system models, including widely-used tools which are often employed in social simulations (e.g. STELLA (www.iseesystems.com) or VenSim (www.vensim.com)). However, in terms of producing rapid graphical representations of dynamic processes these tools can be over-complex. In our evaluation process we developed a simple java-based tool to animate mechanisms in the project called 'InnerState'. The purpose of InnerState was to unite some of the theoretical models behind the project with the categories and new mechanisms that emerged from stage 1 of the evaluation.

InnerState allows for the construction and description of interactive mechanisms through the combination of a series of components. These include 'conveyor belts', 'transformers', 'amplifiers', 'drop-points', 'collectors' and 'generators'. Across these components are passed 'burdens'—or 'things to be processed'. These components may be arranged in any form, configurable through an XML file.

One use-case of InnerState in the evaluation of SPLICE was in the distinction-making between different responses to technology. In figure 3, the mechanism shown is a suggestion for the ways in which different sorts of interventions might be handled by individuals. The categories identified through the first stage of the evaluation were that people could be disrupted in their practice, or they could be coerced into doing something, or a new technological practice could be 'exhorted'. Building on the organizational model suggested by the VSM, three different levels of personal 'regulation' were identified; 'habits', 'organisation' and 'future planning'. The model developed showed the relationship between different types of intervention and the different ways in which those interventions would affect 'habits', 'organisation' and 'future planning'.

The model was a starting point for a deeper discussion which also drew in the distinction between 'zappers', 'energy creators' and 'energy neutrals' which emerged in stage 1 of the evaluation. Questions arising from this included: "what do you do to a zapper to get them to change?"; "what do you do to an energy creator?"; "what are the problems with energy creators?", etc. This led to more distinctions which mapped the level of personal regulation regarding 'habit', 'organisation' and 'future planning' with the identity of a person as a 'zapper', 'energy neutral' or 'energy creator': for example, the characterization of a 'zapper' as being low in 'future planning', but high in habitual responses, or an 'energy creator' as high in 'future planning' but less high in 'day-to-day' habitual responses. Discussions continued in the group at this deeper level.

In this stage of the evaluation process, a number of InnerState mechanism descriptions were produced and made available to the group, including animations of Bhaskar's Transformation Model of Social Activity, Beer's Viable System Model applied to the person, and Harré's theory of selfhood and positioning. By doing this, we were able to relate concepts reported by project stakeholders with concepts derived from theoretical descriptions, whilst engaging rich discussion with project stakeholders.

Teaching and learning in the evaluation process

The use of InnerState for animating mechanisms took the group discussion of mechanisms to a different level. From talking generally about 'zappers' in the institution, a more specific discussion could be had relating to the differences between the ways 'zappers' and 'energy creators' might be organized (in this case, by referring to Beer's Viable System Model). As these discussions progressed, the group dynamics were also important. Unlike phenomenological methods of social research, in Realistic Evaluation the researcher is not considered to be neutral. Nevertheless, it was important to ensure that the explanatory power of the mechanisms that evolved satisfied all stakeholders in the process. Thus in place of theoretical neutrality, there was a dynamic of 'mutual teaching and learning'. Usually, this dynamic was led by the researcher who taught other stakeholders about possible mechanisms using InnerState as a vehicle, those stakeholders situated their experiences, and through the interaction with InnerState taught their own theories confirming or challenging what they understood to be the mechanism the researcher was suggesting. Over time, the mechanism and categories within InnerState were refined.

This teaching and learning model for research is different from those positions advocated in phenomenological or evidence-based research. In the phenomenological case—notably in the popular Grounded Theory method of Glaser and Strauss (1967), the researcher aims to 'bracket-out' any initial presuppositions, with theory emerging from the categories identified by the phenomenological process (usually employing questionnaires, text analysis, coding, etc). In evidence-based research, often a particular theoretical model is considered against the evidence gathered from the research. The problem for Pawson and Tilley with phenomenological approaches is that they regard it as unreasonable to try to ensure that researchers have no initial theory as to what causes particular outcomes. Drawing on Bhaskar's claim that 'reasons are causes', Pawson and Tilley argue that the theoretical views of not only researchers but all stakeholders are causal in the phenomena which result, and must therefore be surfaced through the evaluation process. Furthermore, by engaging in the evaluative process, those theoretical views may change, and thus the evaluation process is itself transformative as well as analytical.

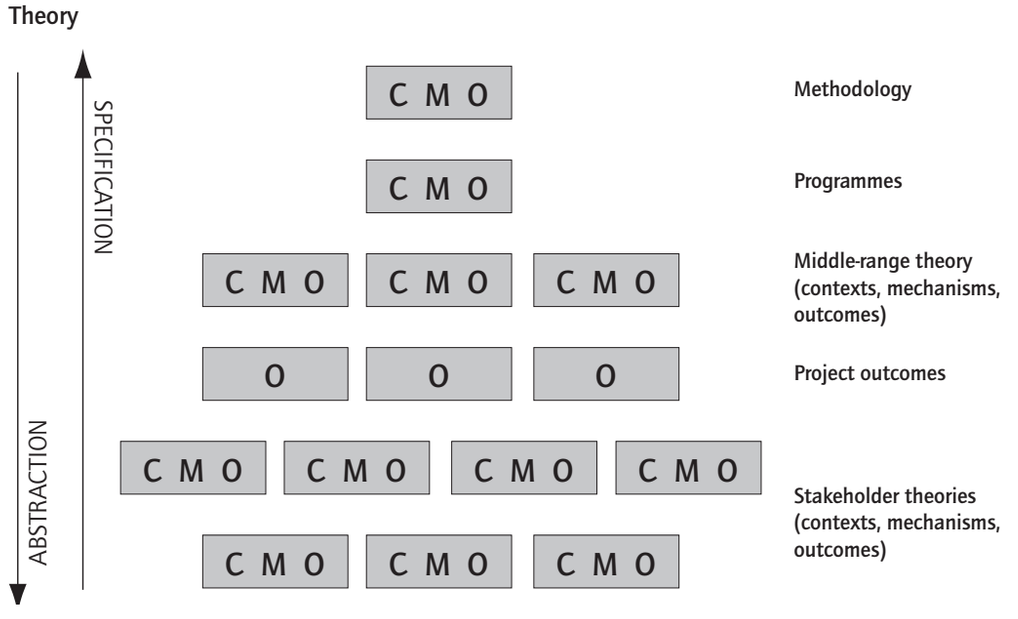
Evidence-based methodological approaches, for Pawson and Tilley, suffer from being overly prescriptive in their specification of 'evaluation criteria' at the beginning of a project. Drawing again on Bhaskar, they argue that the act of identifying evaluation criteria is to assert a particular view of the world which may or may not be grounded in reality. Since the purpose of the evaluation process for them is to identify the mechanisms at work in social systems, this imposing of a world-view before the project starts can be both causal in the project results and also not conducive to a critical engagement with the real mechanisms of the project.

The process of gaining agreement through the SPLICE evaluation was itself transformative on the ways in which stakeholders viewed the project, and had some causal effect in changing habits of those who engaged in the evaluation day. This causal effect was partly due to the way in which many different ideas for mechanisms and outcomes could be cumulated

and shared. This transformational aspect of the evaluation process is key to the Realistic Evaluation approach, since it is agued that the 'value' in the project rests on the efficacy of actions taken as a result of it. As stakeholders become better informed about the mechanisms at work in the project, their actions should become more efficacious as a result. Were this not to be the case, it would simply mean that the mechanisms identified were wrong.

Figure 4 below is an adapted version of a diagram Pawson and Tilley use to explain how different stakeholder perspectives may be cumulated. Basically, the process described is one which relates Abstraction (upwards movement) to Specification (downwards movement). At the 'specification' end, individual stakeholder perspectives identify Contexts, Mechanisms and Outcomes. This results in larger scale general project outcomes which are then used as evidence for the establishment of middle-range theories. Using middle-range theories, new programmes are established. Finally, the methodological approaches used to govern the design of programmes may also be reviewed in the light of developments within projects.

Figure 4: Cumulation of evaluation results on SPLICE (adapted from Pawson and Tilley, 2004)



In the evaluation process on SPLICE, stakeholders were engaged in a process which took them from the low-end specification stage to middle-range theory, and through this process agreement could be established as to possible mechanisms in the project.

Conclusions: between analysis and transformation and the role of technology in evaluation

This paper has described the evaluation process on the SPLICE project. In particular, it has focused on the tools used for the evaluation on the project. The construction of these tools was deemed necessary as a way of taking a different approach to evaluation which avoided the pitfalls of phenomenological inquiry and evidence-based evaluation. By adopting Realistic Evaluation as an approach, we needed to find a way of dealing with descriptions of mechanisms as the principle 'data' of the evaluation process. This has entailed the use of technologies for collaboration and technologies for the sharing, teaching and exploration of possible mechanisms. The question therefore remains: "to what extent does technological advancement affect methodological practice in evaluation?"

In answer to this question, some key distinctions need to be made between methodological practices in evaluation. Traditional evaluation practices tend to be analytical in their treatment of the thing to be evaluated. They treat that thing as something which exists independently of the evaluation process, which is not affected by that process in any way. Realistic Evaluation is transformational as well as analytical. It regards 'value' as being inherent in the actions which arise from a better understanding of the mechanisms of a project. If Pawson and Tilley's arguments about evaluation are correct, then the role of technology in the methodological evolution of evaluation practice is fundamental.

Technology, as well as being (in the case of SPLICE, as with other e-learning projects) the 'thing evaluated' is also often the key medium through which 'things are understood'. Social technology in particular provides a means by which stakeholder perceptions and theories may be surfaced as a project progresses. If 'value' is inherent in the actions that arise from a better understanding of what is going on, then the role of the means by which that understanding is gained is absolutely entwined in the process. As we have shown in the SPLICE evaluation, it makes possible processes that lead to greater shared understanding of common mechanisms: a true synthesis of different perspectives and experiences.

'Synthesis' is perhaps one of the greatest current challenges facing e-learning development. Funding bodies, both national and international, are awash with lengthy reports from thousands of projects from across the world. Many of these projects are doing similar things, many of them 'sort of' work, and yet establishing any common consensus about 'what is going on' seems an elusive goal. Partly, this may be a failure of technology, or at least a failure to address the challenge of describing project outputs in ways which can be assimilated and synthesized into consistent descriptions of mechanisms. Sometimes these evaluative failures result in expensive 'wrong turnings' in funding programmes due to poor middle-range theory, or (more often) simply to produce programme-level outputs which avoid higher-level analysis altogether, preferring to cite specific instances of 'good practice'. With both of these, the institutional manager asking for specific programme-level advice "If I do this too, what's likely to happen?" can often get either an uncertain answer, or (worse) one which glosses over either negative or positive outcomes.

This might lead us, perhaps not unreasonably, to ask "what is the value in doing projects?" However, the knowledge of "what's likely to happen" exists, but often does not get conveyed to programme-level. It is likely that if that same institutional manager asked their question to one of the key stakeholders involved in a project which has done something similar, the answer might be more realistic: "Well, a few learners will like it, a few won't and for most it will 'sort of' work. And as for the teachers, you must watch out for your 'zappers'!" Behind such a reply is an awareness of the real outcomes, contexts, and some idea of a mechanism that lies behind the understanding of the original project. The value for the institutional manager lies in the realism of the advice they are given. Armed with an accurate prediction of what's likely to happen, they can plan accordingly.

In the SPLICE project, we have investigated ways of identifying the value in the project through developing technologies to identify mechanisms at work. The value of the project lies in the effectiveness of the mechanisms identified. The actual mechanisms of SPLICE will be reported elsewhere, but the process of identifying them, and particularly the role of technology in that process, points to a way of addressing the problems of synthesis and value in programme-level outputs.

References

Archer, M. (2000) *Structure, Agency and The Internal Conversation.* Cambridge

Argyris, C. and Schön, D. (1974) *Theory in practice: Increasing professional effectiveness,* San Francisco: Jossey-Bass.

Beer, S. (1981) *Brain of the Firm Second Edition*, John Wiley, London and New York.

Bhaskar, R. (1979) *A Possibility of Naturalism* Sage

Bhaskar, R. (1975) *A realist theory of science* Sage

Beer, S. (1994) *Beyond Dispute: the invention of team syntegrity* Wiley

Campbell, D. T. and Stanley, J. C. (1963) *Experimental and Quasi-Experimental Designs for Research Chicago:* Rand McNally.

Checkland (1990) *Soft Systems Methodology in Action* Wiley

Eden, C. and Ackermann, F. (1989) *Strategic options in development and analysis* (SODA)—using a computer to help with the management of strategic vision Knowledge based management support systems

Glaser, B.G. and Strauss, A. (1967) Discovery of Grounded Theory. *Strategies for Qualitative Research*. Sociology Press

Harré, R. (1984) *Personal Being: A Theory for Individual Psychology* Harvard University Press

JISC (2007) *Six Steps to Effective Evaluation: a handbook for programme and project managers* www.jisc.ac.uk/media/documents/programmes/digitisation/SixStepsHandbook.pdf [accessed on May 22nd 2009]

Johnson, M. and Liber, O. (2008) *The Personal Learning Environment and the Human Condition: from Theory to Teaching Practice Interactive Learning Environments*, v6, no. 1

Luhmann, N. (1995) *Social Systems* Stanford University Press

Mingers J. (2004) *Realising Systems Thinking: Knowledge and Action in Management* Science Springer

Pawson, R. and Tilley, N. (2004) *Realistic Evaluation* Sage