Evaluating problem structuring methods

Leroy White
University of Bristol
Introduction

• This seminar will discuss the challenges associated with the evaluation of problem structuring methods.
• PSMs are seen as complex interventions that seek change and action at many levels, both individual and system.
• There is now a widespread acceptance that the traditional evaluation approaches are inappropriate for the evaluation of PSMs. The difficulty is compounded when PSMs are used in multi-agency area-based initiatives.
• It is proposed that evaluation, while pragmatic and situated, must be a theory based exercise.
• Part of the challenge of evaluation is to provide a narrative of the intervention as well as an agreed assessment.
Problem structuring methods (PSM) are characterised as a family of methods for supporting decisions for groups of diverse composition within a complex environment to agree a problem focus and make commitments to a series of actions. They are usually applied to unstructured problems characterised by multiple actors, multiple perspectives, conflicting interests, and high levels of uncertainty. To tackle problems with these characteristics, PSMs must:
- enable alternative perspectives to be considered with each other
- be transparent to a range of participants
- operate iteratively
- allow contingent solutions

A key feature of these approaches is the use of a model to represent alternative scenarios or versions of the situation of interest, combined with facilitation to help group members make constructive improvements.
Evaluation: feasible or desirable?

- Evaluation has not been a strong feature in the published reports.
- It's not clear the extent to which one can evaluate PSMs (i.e., measure their contribution), in that they are complex technologies (is it feasible?).
- Mingers & Rosenhead said that their wide use might explain their efficacy, but not whether PSMs are better than other processes, or whether one is better than the other.
- There is very little to go on with regards to their utility or usefulness, or whether they are better than doing nothing.

- It is not clear whether evaluation is desirable... because
  - Too difficult to show anything different to doing nothing... or other process
  - Each situation is unique.. (Checkland, Eden)
Lessons from the evaluation of GDSS

• The debates closely echo positivist versus interpretivist approaches
• The 1st often involves experimental or survey approaches, where generalising claims are explicit and constitute an explanation that is the basis of scientific reasoning
• The second is where explanatory claims are less explicit (some even deny any intention towards explanations)
Lessons from the evaluation of GDSS

• Looking at the GDSS literature the emphasis seems to be on the approaches to evaluation. I.e. the methods of evaluation to verify internal validity (explanation causality) or external validity (generalisations) see Eden, Finlay

• Those that are advocates of a more *experimental approach* are concerned with measurement... eg process gain or loss (productivity) and measures of participants perception of the process or facilitation

• Many of the studies have shown that experimental evidence is weak... rarely show significant differences.. however the approaches remain popular (content free)

• Eden argued that *interpretivist approach* will offer at least as much, if not more insights into performance and their future development..

• Others have argued that grounded theory is better because it reveals subtlety and richness of experience. Although some explanations published on why PSM worked in a particular context are weak. E.g. Mckay... could only conclude that PSM ‘helpful’
Validity ad reliability

- Advocates of the positivist say that they should be evaluated and further still to do so either in lab conditions or in the field... the desire is for robust framework.
- This is rejected on the basis that the approaches deviates from the real-life context.
- Advocates of interpretivist appear to emphasis internal validity and proceed as though what really matters is the quality of the original piece of work providing the platform on which explanation can be built. Checkland further argues theoretical validity ... through thick description.
- Finley states that the differences in approach is on experimental vs theoretical validity and that choice between the two is too simplistic.
- No one talks of external validity although case study approach, advocated by some, depends on thick description so that findings from one settings could be seen as transferable to other settings.
- Emphasis on external reliability even rarer (ie repeated study would produce same results).
- Debates distracted us from explanation... into concerns about quality of evaluation and an emphasis on the method of evaluation.
Generalisations

• Eden states because of variety … any evaluation of a GDSS is unlikely to be generalisable across all GDSS … but must be generalisable across all circumstances in which GDSS is used…

• M&R conclude that whilst broadly accepting the interprevist view, it is still possible to assemble general evidence about the perceived effectiveness of PSMs Connell and Omerod are examples.

• Desire for explanation …..but debates have focused on being critical of approaches to evaluation on methodological grounds .. None have really thought that intervention theory is crucial
Summary

Summary:

- In past debates the emphasis was on the design of an evaluation and on emphasis on internal validity.
- M&R claim that most PSM pioneers would claim an interpretivist position...that an evaluation is difficult in that each situation is unique and can only be evaluated on their own terms.
- However, it is claimed (although not shown) that for theoretical generalisation, plausibility depends on theoretical rigour rather than typicality of case.
- Eden suggest grounded theory ... however, even grounded theory calls for rigorous retesting (Denzin and Lincoln).
Alternative?

- Positivist approach is inappropriate for evaluating a PSM intervention. The complexity of the intervention in using PSM is too difficult to be captured by one overarching theory.
- With the interpretivist approach each evaluation of an intervention may bring out different realities across which it may be difficult to judge in any way the value of the different interventions.
- The focus should be on what works better for whom in what circumstances and why? (Pawson and Tilley)
- Adopt a different attitude to methods …
- Focus on interventions given the problem context, are ‘working’ or not and how this may be explained based on the evidence from use.
- Here, all methods have merit if put the theories that can explain an intervention at the centre of an evaluation design..
- So theory should be explicit, evaluation steps built round them, elaborating on assumptions revealing causal chains, and engaging all parties concerned .. this helps build capacity and is also a learning process.
<table>
<thead>
<tr>
<th><strong>Author(s)</strong></th>
<th><strong>Title</strong></th>
<th><strong>PSM used</strong></th>
<th><strong>Evaluation</strong></th>
<th><strong>explanations</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Phahlamohlaka and Friend</td>
<td>Community planning for rural education</td>
<td>SCA, NGT</td>
<td>Single case-study. Reflections by facilitator and satisfaction survey to 30-40 people</td>
<td>Full group engaged through combination of SCA and NGT. Learning process</td>
</tr>
<tr>
<td>Hjortso</td>
<td>Enhancing public participation in NRM using soft OR</td>
<td>SODA</td>
<td>Questionnaire and group discussion 10 people</td>
<td>Soda can improve practice and wider stakeholder integration</td>
</tr>
<tr>
<td>Joldersma and Roelofs</td>
<td>Impact of soft OR on problem structuring</td>
<td>SODA, OMT</td>
<td>Quasi experiment data collected by observation and survey</td>
<td>No positive effects of PSM were found. 2ndary analysis showed PSM had effect in information exchange measurement</td>
</tr>
<tr>
<td>Sorenson, Vidal, Engstrom</td>
<td>Using soft OR in a small company</td>
<td>SCA SWOT</td>
<td>Single case-study</td>
<td>Conclude that Soft OR important factor for developing IT strategy in small companies</td>
</tr>
<tr>
<td>Franco, Cushman, Rosenhead</td>
<td>Project review and learning in the construction industry</td>
<td>SCA</td>
<td>CVF questionnaire (70% response rate) and group deliberation on effectiveness</td>
<td>PSM should be embedded in continuing business practice</td>
</tr>
<tr>
<td>Bryant and Darwin</td>
<td>Exploring inter-organisational relationships in the health service</td>
<td>Drama theory and role play</td>
<td>Case-study through process observation and questionnaire</td>
<td>Ideas applicable in other different domains</td>
</tr>
<tr>
<td>Connell</td>
<td>Evaluating osft OR</td>
<td>SSM</td>
<td>Case-study, retrospective reflections</td>
<td>Why successful structuring leads to unsuccessful implementation…</td>
</tr>
</tbody>
</table>
Theories of practice

What do we do when describing the meaning of our findings?

• In terms of use of evidence- sometimes its hard to tell if an explanation is from the study or otherwise. If going for demanding theory the evidence would need to be strong and can be very difficult to achieve in a single intervention. Loose or weak theory makes it difficult to accept conclusions

e.g. Hjortso claimed that PSMs can enhance citizen participation- evidence was not strong..later he suggested that PSM can improve stakeholder perception of being involved

• In terms of structure of argument – if too ambitious, conclusions may undermine credibility of the evaluation. If too weak ..gets ignored… not much bite.. Then hard to see how you get from here to there...

• Eg Sorenson et al made loose claims and conclusions did not fit evidence . Felt I could ignore the paper.. Boring..
• What works and why… muddling of theoretical bases… vague and cryptic concepts.. making too modest claims

• Eg....Bryant … participants gain more cognitive appreciation of the situation in which they are immersed … the greatest benefits will occur when it leads into and helps inform the creation of clear agenda for change

• Sorenson: Soft OR was useful to x in that x learn about complexity through ease of use , transparency and facilitation (although organisation failed to make decision)

• Phal and Friend: A process using NGT and SCA together is successfully in engaging the full group.. A learning process

• Hjortso  Demanding theory ..PSM enhance citizen participation.. Later he wrote that SODA can improve perception of being involved
An approach

• 1. Why should a more pragmatic approach be adopted? Firstly, it is impossible to incorporate all situations, expectations and interpretations into one evaluation process. Secondly, we will always fail to do justice to this diversity, thus an evaluation will at no time lead to any final true assessment. A pragmatic approach recognises that we are a self-interpreting, meaning making beings, with the task of interpreting the value of our activity and actions as always contingent, complex, contested and never finished.

• 2. even if we know that a PSM intervention worked, we are often less certain about why it worked. Thus, it is important to identify the contextual and intervening factors that hinder or facilitate an intervention’s processes and outcomes in order to reach a better understanding of why an intervention worked, for whom and in what circumstances. Often the outcomes depend on both context and mechanisms, and an intervention that works in one situation may not work in another.
An approach

3. We should be looking towards middle-range theories; neither relying on new theories based on local wisdom from each investigation, nor formal philosophical positions (i.e., theories with universal validity (just society!))

4. MRTs should be moderate in focus and be open to change...
Case

Sure Start in Tower Hamlets

- The area is characterised as
- as mixed residential with many low-income families living in poor-quality housing.
- where the environment is poor and there is little in the way of community facilities.
- The area experienced high levels of Asian immigration during the 1980s, particularly Bengalis.
- This group has a wide range of unmet needs and there is poor availability of appropriate services
- Programme governed by Partnership Board (mainly service providers) and programme director
- Initially set up as a set of services.. Board and PD wanted more CD approach
- Parent’s forum established at a later date
Case

Board wanted to identify and explore policy options which reflected local parents’ priorities and informed by local views.. Dominant professional perspective

Brief:

• Working with board to develop a set of priorities and options for parent involvement
• Working with parents forum and local community to identify local priorities and build capacity to engage with Sure Start board
• Having a final workshop with board and parents group to agree a series of community development principles and activities
Case

- Use PSMs to reach an agreement on the way forward
- Handle the diverse and complex arrangement and conditions surrounding the development of policy options
- The need for planning and preparation and care was taken in the selection of venues (for the parents it was on the estate)
- The work with the Sure Start Board involved two workshops using a combination Strategic Choice Approach (SCA) and Nominal Group Technique (NGT) over a period of 3 months.
- The use of ‘oval’ post-its and SCA assisted the sharing of focus.
- The workshops produced a number of areas which were discussed in detail. The main issue was whether to continue to support mainstream services versus community-based activities. Also, to explore what types of community based services would support the objectives of the Sure Start.
- The group worked in a participative way and ideas emerged which the group worked on, agreed and accommodated.
- There were different constructions of what the priorities were and different versions of the same views.
Case

- The workshop with the parents (later the parents’ forum) involved the use of PRA and elements of SCA.
- PRA uses a range of visually oriented methods for eliciting and collating views and representing the information in a transparent and accessible way.
- In the work with the parents, role play and Venn diagrams were used and the options developed were evaluated using pair-wise comparisons.
Case

Final workshop

• For the final workshop, representations (conceptual models) from each area of work were fed-back to each of the groups. This was to ensure an overview of progress was given and the groups would be given an opportunity to clarify issues or respond to concerns.

• It was also agreed by both groups that the final workshop would focus on participation and partnership. The design, therefore for the final workshop had to keep this in mind.

• The final workshop used a combination of SCA and some facilitation processes for large group interventions.

• The workshop began with a short plenary that set out the aims which then led into the session.

• The workshop adopted a technique called ‘carousel’ to support joint working and sharing of information among a large group of participants.

• A plenary - after the original sub-groups had had an opportunity to review the views relating to their zone - was held to discuss differences and shared views and to begin the task of reaching a convergence. The plenary explored commitment package and closed with the use of dots for voting.
Evaluation

- Use of PSMs. Firstly, to derive a description of what happened and to understand how and why certain things happened or did not happen. This aim goes beyond describing what occurred and attempts to explain causes.
- The seconds aim was to carry out further reflections of the findings in order to provide some insights (in the form of ‘middle range theories’) which could provide guidance about other similar interventions.
- The evaluation would need to be a pragmatic use of methods of date collection, have an idea of the theoretical base in keeping with the philosophies/theories behind the PSM interventions and acceptable to those who may use the evaluation findings; ie, the practitioners, clients and sponsors
Evaluation

• Data
• Pre and post intervention interviews were conducted with the s/h participants and review of the artefacts produced (i.e, flip chart notes, diagrams and maps created with ‘oval’ post-its etc.) were reviewed. Observations of mood, what got suppressed, what resistances, tensions and how these changed through out the process. These were the key sources of data.
• Important for author to be part of data generation ensure unique knowledge was fully utilised in the enquiry in to the use of PSMs in the intervention
• Data organised using Atlas/ti
Evaluation

- *Some theories from lit and records*
- *The reasons for the success or otherwise of the PSM interventions are to be found in the context in which they operate and in the mechanisms and processes employed or not employed in the pursuit of identifying policy options for the Sure Start board.*
- PSMs can deal with complexity, multi-perspectives, help agree a problem focus and agreement on course of actions
- PSMs are successful in single orgs (many s.h with many views) and can be successful in multi-org situation
- Stakeholders thought that PSM would help reduce conflict, build capacity
- Parents thought process would develop new skills and capacity... build confidence
- Mechanisms are transparent approaches, modelling and facilitation
Evaluation

• It was clear that the focus on transparency, modelling and facilitation could account for what was going on.
• The board and the parents’ group found that the use of visual methods was important to the success of the intervention.
• The board thought that the use of mapping allowed perspectives to be shared and that complex issues could be represented in an accessible way.
• They also found that the maps helped the group to have constructive dialogue, which they felt would not have happened in a normal board meeting. The PSMs methods were effective in terms of visualisation.
Evaluation

- The maps did ensure engagement with the issues but it also surfaced tensions and resistance. There was a clear understanding that via the intervention collaboration between different groups was possible, even if it was not taking place.
- During the intervention people felt confident to develop their ideas about networking and collaboration
Evaluation

• Parent’s previous experience with SS was via 1-to-1 interaction. They were not used to working as a group. Weak networks and mistrust of outsiders.

• Intervention emphasised alternative social processes and social learning and a focus on relationship building.

• The final workshop which attempted to focus the group on commitment to particular outcomes led to a surprise. When something was seen as a good outcome, people did not want to lose it. They organised to keep it going, sometimes in a cooperative way (does this say something about empowerment?)
Evaluation

Mechanisms that had worked were:

• Basic visualisations helped share perspectives. People were happy to apply it in whatever situation they need.

• Favouring facilitation.. working in groups.. helped to develop ‘mediation’ capacities in places where individualism and conflict are endemic.

• Being pulled along.. Something feels good then other people (parents in this case) not directly involved want to be a part of it (final workshop).
Evaluation

• Problem-structuring is the most interesting part of the intervention as it is a manifestation of the strength of voices and power. Thus the techniques that the facilitator uses need to deal directly with power (power distribution was asymmetrical.. Are those who are use to control willing to cede power to others? Does the process only lead to a legitimacy of the agencies decisions rt genuine involvement?

• Previous studies appear to prioritise the tool over relationships.. The present study indicates that the tool(s) should be subordinate to relationships

• Facilitation seems to be about brokerage.. Need to reflect on this in terms of the relationships formed.. Who is included and excluded

• What works varies between an agreed fact to a social process around core set of technologies. Heterogeneous audience with different requirements generally demand different accounts
Conclusion

- PSMs are complex technologies and thus begs question whether desirable or feasible to evaluate
- Earlier debates have focus on approaches to evaluation and conclusions drawn range from being vague to demanding
- Effect has been to divert attention from what works, why, where and for whom?
- Sympathise with interpretivist...Suggest pragmatic use of methods ...but as Eden suggested that evaluations should consider the context in which the interventions are enacted and the different interests and views of stakeholders..consider how means and mechanisms will work to produce outcomes
- Focus on what works means dealing with theories of practice... original intentions ... Purpose of evaluation to test the theories ... purpose of generalisation/conclusions re towards middle range theories which are moderate in view and open to change
- Future - study of intervention (PSMs as technology Latour...)
- Facilitation...focus on brokerage ... maybe focusing less on content and more on bespoke (fit for purpose) relational process
- More auto-ethnography.
- Comparators are difficult. What about the Counterfactual ... what if we did nothing....?