RAMESES

Realist and Meta-narrative Evidence Synthesis: Evolving Standards

Delphi Panel Briefing Document: developing publication standards for realist and meta-narrative reviews

Trish Greenhalgh, Geoff Wong, Gill Westhorp, Ray Pawson

CONTENTS

What we would like you to do, how and when

The task is to produce consensus publication standards for two sorts of systematic review: realist and meta-narrative. You have agreed to be a member of our Delphi panel. A Delphi panel is a way of working towards consensus on a topic or question. It consists of a number of rounds. In a preliminary round, you will be asked to suggest topics which you would like to see covered (or statements you would like to see included). In each subsequent round (usually two more), you will be asked to do a task which involves *scoring* a draft set of statements. There will be a deadline for this, because we can't analyse the responses until everyone has replied.

After each *scoring* round, you will be sent your own scores *and* the average score for everyone in the group. If you find you are an 'outlier', you have two choices: amend your score (after reflecting on the statement and why you scored it as you did) – or stand your ground and argue your case to the group (they won't know how you scored the statement). Even if you scored a statement similarly to the group average, you may be swayed to change your score by arguments put subsequently.

Each statement is scored on two dimensions: [a] relevance (should we include this topic / theme at all?) and [b] content (should we word it like this?). High scores for relevance *and* content mean the statement will be included 'as is'. High scores for relevance but low scores for content means we need to word the statement differently (we'll ask for suggestions). Low scores for relevance mean the statement gets dropped. But when some panel members score a statement high and others score it low, we need a discussion. For references on the validity and methodology of the Delphi process, please ask us.

Here's what we'd like you to do:
- Pull out now if you've changed your mind (so you don't count as a 'withdrawal')
- For ROUND 1, please read this background paper (and, if you've got time, the full study protocol and the other documents we have provided)
- Respond within one month to Geoff *only* by hitting the reply button with your suggestions.
- Wait while we analyse all the responses and build the draft statements
- Respond to the ROUND 2 email (expected mid-October 2011) within one month by looking at the statements and entering your scores for each (we'll give you a link to an online questionnaire)
- Wait again while we analyse the data and send you back your scores
- Join in an email discussion on how we might amend the statements
- Repeat the last three steps for ROUND 3 (expected late November 2011)

This Delphi panel is part of the wider RAMESES project, which has six work packages: [a] produce publication standards for realist and meta-narrative reviews; [b] refine and extend existing methodological guidance; [c] develop, pilot and run training modules; [d] run a JISCmail discussion list (www.jiscmail.ac.uk/RAMESES); [e] support teams undertaking reviews; and [f] contribute to the academic literature (e.g. on the methodology of doing [a] to [e]). The RAMESES study protocol is appended.

Authorship policy

We want to acknowledge the input of everyone who contributes to RAMESES. We propose two levels of authorship:

a. People who contribute materially and significantly to conceptualising the study, undertaking the research, analysing the data or writing up will be named as co-authors alongside us on publications. The format of the author list will be "Smith A, Jones B, Bloggs D on behalf of the RAMESES group".
b. Members of the Delphi panel who do not fulfil the above criteria will be listed directly below the authors in the following format: "The RAMESES group comprised: Aardvark H, Bloggs D …etc to Zindel B".

Please let us know if you are looking for a formal authorship role or if at any stage you believe you deserve to join the author list. We will also be alert to input from Delphi panel members above and beyond what is expected of an ordinary participant. It is quite possible that the RAMESES statement will have a large number of authors and we are comfortable with that.

Whatever your level of input to this project, you won't get paid unless you were costed on the grant application. Nevertheless your input is greatly valued.

Briefing on meta-narrative reviews

Background

Meta-narrative review is a new method of systematic review, designed for topics which have been *differently conceptualised and studied* by different groups of researchers. Here's an example. Many groups have studied the building of dams in India. Some have conceptualised this dam-building as engineering; others as colonialism; others as a threat (or promise) to the local eco-system; others as inspiration for literature and drama, and so on. If we were to summarise this topic area in a way that was faithful to what each different group set out to do, we would have to start by asking how each of them approached the topic, what aspect of 'dams in India' they chose to study and how. In order to understand the many approaches, we would have to consciously and reflexively *step out of* our own world-view, learn some new vocabulary and methods, and try to view the topic of 'dams in India' through multiple different sets of eyes. When we had begun to understand the different perspectives, we could summarise them in an over-arching narrative, highlighting what the different research teams might learn from one another's approaches.

(NOTE: some reviewers might be interested only in summarising the findings of randomised controlled trials of 'dam present' versus 'dam absent' on a predefined outcome, and if that was the focus of the review, a Cochrane review with statistical meta-analysis would be the gold standard approach. The meta-narrative approach is only intended for those reviews where the underlying research goal is to identify and explore the *diversity* of research approaches to a topic.)

Methodological issues in meta-narrative review

The methodology of meta-narrative review was developed by Trish Greenhalgh and her team in 2004 when reviewing the literature on diffusion of service-level innovations in healthcare.[1] A methods paper was published in Social Science and Medicine in early 2005.[2] The inspiration for this method was Kuhn's 1962 book *The Structure of Scientific Revolutions*, which argued that science progresses in paradigms (i.e. particular ways of viewing the world, including assumptions about how the world works) and that one scientific paradigm gives way to another as scientific progress renders yesterday's assumptions and practices obsolete.[3] Newton's theories and methods, for example, became less and less able to answer the emerging questions of particle physics, leading Einstein to develop his theory of relativity. Meta-narrative review looks historically at how particular research traditions have unfolded over time and shaped the kind of questions being asked and the methods used to answer them. A research tradition is a series of linked studies, each building on what has gone before and taking place within a coherent paradigm (that is, within a shared set of assumptions and preferred methodological approach shared by a group of scientists).

While researching the background for the RAMESES project, we came across *meta-triangulation review*, another synthesis method described by Marianne Lewis and Andrew Grimes in the Academy of Management Journal in 1999.[4] We had been unaware of this approach when we published our original work on meta-narrative review but have subsequently communicated with Prof Lewis, who has offered her input to the RAMESES study (subject to other commitments). The (many) similarities and (few) differences

between these two approaches are shown in Appendix 1. When collecting meta-narrative reviews for the RAMESES study, we looked at examples of meta-triangulation reviews but decided to exclude these because meta-triangulation review seeks to understand and analyse topics at a the level of theoretical differences between paradigms, whereas meta-narrative review is more interested in the research tradition as its unit of analysis (a working definition of a research tradition is "what researchers get up to *within* a paradigm"). However, the methodology of meta-triangulation review offers some transferable insights which will help us refine the quality criteria for a meta-narrative review (Appendix 1).

Summary of published examples of meta-narrative reviews

With the help of a specialist informaticist/librarian (Jeanette Buckingham), we identified a sample of 9 published papers which were described as meta-narrative reviews. These were examined independently by Geoff and Trish. As expected, the 9 reviews covered a range of complex topic areas which had been differently studied by different groups of researchers (e.g. electronic patient records, environmental health, fundamentals of nursing care, knowledge translation and exchange). Most were published after 2009, and we know of several more reviews which are ongoing or in press. We considered that five of our sample of 9 were "true" meta-narrative reviews, defined by three working criteria: [a] the authors clearly understood the need to consider the topic from multiple paradigmatic perspectives and used the research tradition (or something comparable to it) as their unit of analysis; [b] the authors made efforts to step out of their own world-view and recognise and value alternative world-views; [c] the synthesis included a comparison of how the topic area was approached from at least two contrasting perspectives. Another of the 9 reviews appeared to "almost" meet these criteria. Three papers described as meta-narrative reviews did not meet even these fairly loose criteria.

Preliminary thoughts on publication standards for meta-narrative reviews

Our analysis of these published reviews, along with our discussions with review teams who are currently undertaking meta-narrative reviews, have surfaced the following issues and implications for the RAMESES project. These are preliminary – we hope the Delphi panel members will add to and/or challenge them.

1. TERMINOLOGY. Key terms were used inconsistently by review teams (partly because we had omitted to define some of them in our original methods papers).

   => *We need a glossary and set of definitions.*

2. PHILOSOPHICAL BASIS. The philosophical assumptions of meta-narrative review (e.g. Kuhn's notion that science progresses in paradigms, and a paradigm is a particular conceptual lens which shapes what counts as knowledge) appear to be widely misunderstood. Misunderstanding or undervaluing the importance of the philosophical basis of meta-narrative reviews and its implications appeared to lead to mis-application of the method.

   => *We need to find ways of making the philosophy accessible and its implications clear.*

3. CLASSIFICATION. Some review teams did not appear to understand the fundamental difference between the meta-narrative method and an old-fashioned narrative review. The term "meta-narrative" seemed to be used as a synonym for a form of thematic analysis or on one which offered an under-theorised "thematic analysis" (by which was meant that the findings section listed the themes raised by the empirical papers).

   => *We need to include very clear criteria for classifying a review as "meta-narrative" and an alert that the term is sometimes misused.*

4. UNIT OF ANALYSIS. Some review teams were confused about "the research tradition" as the unit of analysis.

   => *We need to clarify what a research tradition is <u>and</u> the importance of inter-disciplinary working.*

5. TITLE. Some but not all meta-narrative reviews were described as such in the title.

   => *We need to encourage authors to do this.*

6. RESEARCH QUESTION. Some review teams either asked no research question or asked a question which did not seem amenable to being answered using meta-narrative review. Successful reviews had sought to make sense of a topic by appropriately adapting of one or more of the 'generic' questions that underlie the meta-narrative method (How has the topic been conceptualised in each separate research tradition? What are the key concepts, theories, assumptions? What are the preferred study designs and ways of knowing? What are the main empirical findings? What can we learn from the range of different approaches?)

   => *We need to clarify what a research question (and sub-questions) would look like in a meta-narrative review. We also need to highlight the kinds of questions which are UNsuited to this kind of review.*

7. METHODS. Some review teams appear to have cut and pasted the methods section from a published meta-narrative review virtually verbatim into their own paper, thus *claiming* to have followed all the recommended steps even when it was clear that they had not. This suggests that some journal editors and peer reviewers are unable to judge whether the method is being followed or not. Some review teams described a "modified" meta-narrative approach but did not say how and why they modified it.

   => *We need to include techniques for confirming that the methods were actually followed and an alert to the cut-and-paste ruse. We need to include the instruction that if teams modify the method, they have to say <u>how and why</u> they modified it.*

8. SEARCHING. A number of reviews did not undertake any iterative searching. Rather, they used a one-off, predefined search strategy (as is standard in many Cochrane reviews). But because meta-narrative review is aimed at making sense of the

literature, it may only become clear which data are needed as a review progresses. Hence searching needs to be not only iterative, but also purposive and flexible (e.g. in terms of inclusion criteria).

=> *We need to encourage review teams to begin with a broad, "browsing" search and progressively refine this in the light of emerging data. If iterative searching is not undertaken, such a decision should be justified.*

9. FOCUSING THE REVIEW. Several teams reported difficulties in making the review manageable within the time and resource available. This is an inherent problem since meta-narrative review is designed to make sense of large and contested bodies of literature. Where reported, containment and focusing was achieved through discussion within the review team and with reference to interested parties (e.g. service users, experts in the field).

=> *In the publication standards, we need to expect a statement of how the review was shaped and contained. In the methodological advice we need to suggest approaches to this.*

10. APPRAISING PRIMARY STUDIES. Review teams appraised studies in different ways. Some used a flexible approach, using judgement to include (or exclude) and appraise studies in an iterative manner as their data extraction and synthesis unfolded. Others preferred to develop a formal list of inclusion and exclusion criteria and used this systematically (and somewhat inflexibly) to rule papers in or out of their dataset *before* reading them in detail. We strongly favour the former method, which aligns with the interpretive basis of meta-narrative review. Meta-narrative review assesses studies using the quality standards accepted within a paradigm. The purpose of this is to aid the sense-making process as it allows (for example) a review team to say "in this tradition, the X, Y and Z are considered to be high quality studies".

=> *We need to encourage an interpretive and iterative approach to assessing primary studies for inclusion.*

11. FINDINGS. Some review teams did not provide sufficient detail to support the inferences in their findings section.

=> *We need to include clear guidance on how we expect review teams to present and justify their findings in a way that allows others to judge their coherence and plausibility.*

12. CONCLUSIONS. Some but not all teams provided a clear line of reasoning linking findings to conclusions and recommendations.

=> *We need to require conclusions should be 'traceable' back to detailed presentation of the mapping and analysis of the research traditions and their underpinning paradigms.*

Briefing on realist reviews

Background

The realist research question is often summarised as "What works for whom under what circumstances, how and why)?" Realist inquiry considers the interaction between context, mechanism and outcome. In a realist world, intervention X is *not* thought of as having effect size Y with confidence interval Z. Rather, intervention X (e.g. a programme introduced by policymakers who seek to create a particular outcome) alters context (for example by making new resources available), which then triggers mechanism(s) which produce both intended and unintended outcomes. X may work very will in one context but poorly or not at all in another context.

Realist inquiry seeks to unpack the context-mechanism-outcome relationship, thereby *explaining* examples of success, failure, and various eventualities between. Theoretical explanations of this kind are referred to as "middle-range theories" (i.e. ones which *"...involve abstraction... but [are] close enough to observed data to be incorporated in propositions that permit empirical testing".*[5]

The basis of realist inquiry is a realist philosophy, whose key tenets are as follows (feel free to challenge these – this is just to get us going):

1. There is a [social] reality which can't be measured directly (because it is processed through our brains, language, culture and so on) but can be known indirectly. Realism thus sits, broadly speaking, between positivism ('there is a real world which we can apprehend directly through observation') and constructivism ('given that all we can know has been interpreted through human senses and the human brain, we cannot know for sure what the nature of reality is').
2. Social programmes (including complex interventions) may change the social context (for example by introducing legislation) or may change the resources or opportunities available to participants and, in that sense, change the *context* for those participants.
3. To understand the relationship between context and outcome, realism uses the concept of *mechanisms*, defined as "...underlying entities, processes, or [social] structures which operate in particular contexts to generate outcomes of interest." [6]

The realist approach has informed empirical studies (realist evaluation),[7] and offers the potential for insights (e.g. in relation to complex interventions and the implementation of research findings) that go beyond the narrowly experimental paradigm of the randomised controlled trial.[8]

Methodological issues in realist review

"Realist synthesis" was first described by Ray Pawson in 2002,[9] updated in an ESRC-commissioned monograph in 2004[10], published as a book in 2006[11] and summarised in a short methods paper in the Journal of Health Services Research and Policy in 2005.[12]

A realist review (or realist synthesis) applies realist philosophy to the synthesis of findings from primary studies that have a bearing on a single research question. Reviews begin by eliciting from the literature the main ideas that have gone into the making of a class of interventions (the programme theory). This programme theory can be thought of as setting out how and why a class of intervention is though to 'work' to generate the outcome(s) of interest. The pertinence and effectiveness of each constituent idea is then tested using the available evidence (qualitative, quantitative, comparative, administrative, etc.) that has gathered in the primary literature on that family of programmes. In this testing, the ideas within a programme theory are re-cast and conceptualised in realist terms and for each idea reviewers have to seek out the contextual (C) influences that have triggered the relevant mechanism(s) (M) to generate the outcome(s) (O) of interest. Synthesis consists of comparing 'how the programme was supposed to operate' to the 'empirical evidence on the actuality' – all along CMO lines. Analytic purchase comes from the ability to describe and understand the many contingencies that need to be put in place (or avoided) to improve the likelihood of such interventions generating their intended outcomes – in other words, explaining how an intervention might change the context or provide resources in such a way as to most likely trigger the right mechanism(s) to produce the desired outcome.

Summary of published examples of realist reviews

We identified a sample of 35 published papers which were described as realist reviews. These were examined in detail by Geoff, and aspects of his analysis checked by Trish, Ray and Gill. They were published between 2004 and 2011 and covered a broad range of topics (e.g. health, education, human resources). We classified 8 of these 35 as high-quality realist reviews, five as having many but not all features of a high-quality realist reviews and 22 lacking many substantial aspects of a realist analysis. Our classification of these reviews was based on our judgment of whether a realist analysis (the application of realist logic and concepts in a review and synthesis) had been undertaken.

Preliminary thoughts on publication standards for realist reviews

Our analysis of these published reviews, along with our discussions with review teams who are currently undertaking realist reviews, have surfaced the following preliminary issues and implications for the RAMESES project. Many of these are similar to the problems found in meta-narrative reviews.

1. TERMINOLOGY. Key terms were used inconsistently by review teams (partly because even when we defined them clearly in our own publications, other conflicting definitions exist in the literature or teams chose to redefine the concepts themselves).

   => *We need a glossary and set of definitions (and we need to make sure we don't just privilege what we ourselves have written before).*

2. CLASSIFICATION. Currently the number of reviews which we have judged to have significant limitations from a realist perspective outnumber those which we have judged to be robust realist reviews.

=> We need an accessible way of determining the quality of realist analysis within realist reviews.

3. PHILOSOPHICAL BASIS. In our judgement, the commonest flaw in our sample of published realist reviews was lack of appreciation of the philosophical basis of realism and the implications of this for the review methodology. These reviews used the term "realist" to mean (variously) "qualitative", "narrative", "non-Cochrane" and/or did not explain its methods in detail.

   => We need to explain the philosophy and its implications for the methodology. We need to devise ways in which assessors can determine whether review teams understood the implications of realist philosophy and its application.

4. THE REVIEW TEAM. High-quality realist reviews tended to have been undertaken by a team of reviewers with relevant methodological expertise. This appeared to have allowed them to "bounce ideas off" each other to focus the review and apply the realist logic of analysis appropriately and effectively.

   => We need to highlight the opportunities offered by working in a review team.

5. RESEARCH QUESTION. Some review teams either asked no research question or asked a question which did not seem amenable to being answered using realist review (e.g. one that could not be mapped to the generic question "what works, for whom, in what circumstances, to what extent, how and why?").

   => We need to clarify what a research question (and sub-questions) would look like in a realist review. We also need to highlight the kinds of questions which are UNsuited to this kind of review.

6. METHODS. There was a mismatch between what review teams said they had done and what the findings section suggested had actually been done. Sometimes, a review would explain that realist review had been chosen as the preferred method for one of the following reasons; heterogeneous data, a wish to synthesise quantitative and qualitative data, and/or to address the "what works for whom and in what circumstances" question. An explanation would then be provided as to what the realist review method is. This section (if provided) often suggested that a realist analysis had been undertaken, but the findings would consist of a thematic and/or narrative synthesis. In some reviews which made claims to be "realist", realist concepts were not mentioned at all or incorrectly conceptualised (e.g. mechanism was confused with intervention). Some review teams described a "modified" realist approach but did not say how or why they modified it.

   =>We need to include techniques for confirming that realist methods were actually followed and an alert to the cut-and-paste ruse. We need to include the instruction that if teams modify the method, they have to say how they modified it and why.

7. SEARCHING AND INCLUSION CRITERIA. A number of reviews did not undertake any iterative searching. Whilst this may not always be necessary, it is highly likely that as the process of theory-building and theory-testing progresses, additional searching will be needed after the initial papers have been identified. Some of the realist reviews in our sample searched for and included only randomised controlled trials (RCTs), and these found that they had too little detail in their included studies to build and test theory.

   => *We need to encourage review teams to search iteratively, purposively and continuously throughout the review, and refine searches in the light of emerging data. We should discourage including only RCTs in a realist review, since the data needed to enable reviewers to make coherent and plausible theoretical inferences can come from a wide variety of sources.*

8. FOCUSING THE REVIEW. Several teams reported difficulties in making the review manageable within the time and resource available. This is not surprising since a realist review on any topic is potentially endless as more and more refined explanations are sought to explain increasingly diverse aspects of a review. Where reported, as with meta-narrative review, containment and focusing were achieved through discussion within the review team and with reference to interested parties (e.g. decision makers, experts in the field).

   => *In the reporting standards, we need to expect a statement of how and why the review was shaped and contained. In the methodological advice we need to suggest approaches to this.*

9. THEORISING. In reviews which had some but not all characteristics of a robust realist review, what was missing was the use of theory to try to provide an overarching coherent and plausible explanation of the observed patterns of outcomes. An important aspect of this theory development is that it is iteratively tested against the reported data in the included studies.

   => *We need to include an expectation for this type of theorising in the publication standards (and explain how to do it in the methodological guidance and training materials).*

10. APPRAISING PRIMARY STUDIES. Review teams appraised studies in different ways. Some used a flexible approach, implicitly or explicitly following Pawson's judgement-dependent criteria of "relevance" and "rigour", and appraised studies in parallel with their data extraction and synthesis. Others preferred to apply a formal critical appraisal checklist and used questions on this checklist as a tool for excluding studies before undertaking the detailed synthesis work. We strongly favour the former method, which aligns with the explanatory basis of realist synthesis. Quality appraisal may need to be iterative, because as the process of theory-building and/or refinement unfolds, a different section of an included study may yield relevant data.

=> We need to encourage approaches that assess the relevance and rigour of primary studies for inclusion. Studies should be included in a review if they are able to make a contribution to theory building or testing (relevance). We should caution against excluding studies based on an overall assessment of their 'quality', since small segments of included studies may contain useful insights for theory-building even when other aspects of them are flawed. We should encourage a parallel rather than sequential approach to appraising papers and synthesising the insights from them.

11. FINDINGS. Some review teams did not provide sufficient detail to support the inferences in their findings section.

    => We need to include clear guidance on how we expect review teams to present and justify their findings in a way that allows others to judge their coherence and plausibility.

12. RECOMMENDATIONS. Few studies contained sufficient detail on contextual influences. The explanations in realist reviews are highly dependent on contextual influences. It follows that recommendations must be contingent (for example only under certain contexts will a particular mechanism be triggered to generate the desired outcome) rather than a list of "dos and don'ts".

    => We need to stipulate the format of recommendations in a relist review (e.g. "In pursing programme theory A, attend to the following contingencies, context and implementation features B, C. D, E, ...N.").

Appendix: Meta-triangulation vs. meta-narrative review

	META-TRIANGULATION REVIEW[4]	META-NARRATIVE REVIEW[2]
Purpose	To build theory. "Studying multifaceted phenomena characterized by expansive and contested research domains"	To build a rich, multifaceted picture of a complex topic, especially when a summary is needed for policy decisions
Philosophical basis	Constructivist (Kuhn's philosophy of science)	Constructivist (Kuhn's philosophy of science)
Intended audience	Academics	Policymakers
Type of insights	Analytic	Predominantly descriptive but recognises potential for analytic, theory-building insights
Examples of topics reviewed	Theoretical topics at high level of abstraction e.g. power, strategy	Policy and/or practice-relevant topics e.g. electronic records, knowledge translation
Empirical data	Included only as an aid to theorising	Included as substantive component of review
Unit of analysis	**Paradigm:** "the assumptions, practices and agreements among a scholarly community"	**Research tradition:** the historical unfolding of research on a particular theme by a group of scientists, which occurs within a paradigm
Key stages	**GROUNDWORK** Define phenomenon of interest Focus paradigmatic lenses **SEARCH** Collect data interpretable from multiple lenses **MAPPING PARADIGMS** Plan paradigm itinerary (ordered use of different paradigmatic lenses) Code data Write paradigm accounts **THEORY BUILDING** Explore metaconjectures Attain meta-paradigm perspective Reflect critically on the process	**GROUNDWORK** Assemble multidisciplinary team Outline research question Agree outputs with funder **SEARCH** Browse literature to identify research traditions Search within each tradition to identify seminal conceptual and theoretical papers Search systematically for empirical papers **MAPPING RESEARCH TRADITIONS** Describe paradigmatic basis for each tradition Highlight the 'storyline' of each tradition (key issues and discoveries as they unfolded) Appraise and summarise primary studies **SUMMARY / SYNTHESIS** Summarise each research tradition separately, highlighting similarities and differences View discrepancies as higher-order data; explain as contestation between paradigms **RECOMMENDATIONS** Consider implications for sponsor / audience
Principles and approaches to assure quality of the review	**Reflexivity:** Theorist should be fully aware of own assumptions **Systematic cross-paradigm synthesis techniques:** e.g. paradigm bridging (seeking commonalities), paradigm bracketing (highlighting differences), interplay (exploring tensions); meta-theorizing (exploring patterns that span conflicting understandings)	**Pragmatism:** What to include is not self evident **Pluralism:** Include multiple perspectives and ask what we can learn from each **Historicity:** Trace research traditions over time **Contestation:** Use "conflicting findings" in a positive way to generate new insights **Peer review:** Present emerging findings periodically to a critical external audience

Reference List

(1) Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organisations: systematic literature review and recommendations for future research. Milbank Q 2004; 82:581-629.

(2) Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O, Peacock R. Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. Soc Sci Med 2005; 61(2):417-430.

(3) Kuhn TS. The structure of scientific revolutions. Chicago: University of Chicago Press; 1962.

(4) Lewis MW, Grimes AJ. Meta-triangulation: Building theory from multiple paradigms. Academy of Management Review 1999; 24(4):672-690.

(5) Merton.R. On Theoretical Sociology. Five Essays, Old and New. New York: Free Press; 1967.

(6) Astbury B, Leeuw F. Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. American Journal of Evaluation 2010; 31:363-381.

(7) Pawson R, Tilley N. Realistic evaluation. London: Sage; 1997.

(8) Berwick DM. The science of improvement. JAMA 2008; 299(10):1182-1184.

(9) Pawson R. Evidence-based policy: the promise of 'realist synthesis'. Evaluation 2002; 8(3):340-358.

(10) Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist synthesis - an introduction. ESRC Working Paper Series. London: ESRC; 2004.

(11) Pawson R. Evidence-based policy: A realist perspective. London: Sage; 2006.

(12) Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review--a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy 2005; 10 Suppl 1:21-34.