Beyond synthesis: Augmenting systematic review procedures with practical principles to optimise impact and uptake in educational policy and practice

Abstract

Whilst systematic reviews, meta-analyses and other forms of synthesis are often constructed as sitting proudly atop the hierarchy of research evidence, their limited impact on educational policy and practice has been criticised. In this article, we analyse why systematic reviews do not benefit users of evidence more consistently and suggest how review teams can optimise the impact of their work. We introduce the *Beyond Synthesis Impact Chain (BSIC)*, an integrated framework of practical strategies for enhancing the impact of systematic reviews. Focusing upon examples from health professions education, we propose that review teams can optimise the impact of their work by employing strategies that 1) focus on practical problems and mindful planning in collaboration with users; 2) ensure reviews are relevant and syntheses reflexively account for users' needs; and 3) couch reports in terms that resonate with users' needs and increase access through targeted and strategic dissemination. We argue that combining practical principles with robust and transparent procedures can purposefully account for impact, and foster the uptake of review evidence in educational policy and practice. For systematic review teams, this paper offers strategies for enhancing the practical utility and potential impact of systematic reviews and other forms of synthesis.

Keywords: systematic review, impact, knowledge synthesis, evidence-based practice

Introduction

The rhetoric of evidence-based practice is ubiquitous. Since the late 1990s, practitioners and policymakers across a range of fields have been expected to ensure that their practices and policies are underpinned with rigorous research that robustly demonstrates 'what works' (Wells, 2007). As a result mechanisms and procedures for locating, assessing the quality of, and synthesising evidence were, and continue to be, constructed. Systematically synthesised evidence, in the form of the systematic review, is often favoured by policymakers in education and the social sciences (Rubin and Bellamy, 2012; Solesbury, 2001). The transparent and auditable procedures of systematic reviews, along with their rigorous assessments of methodological quality, are lauded as providing conclusions that far exceed the validity of individual studies alone. In short, the message of the rhetoric is 'to base your practices and policies on evidence, locate or commission a systematic review that responds to your specific practical problem.'

Systematic reviews were pioneered internationally in the medical arena by the Cochrane Collaboration, but have crossed over into social policy with the establishment of the Campbell Collaboration, and in the United Kingdom the ESRC UK Centre for Evidence-Based Policy and Practice (Solesbury, 2001), the Evidence for Policy and Practice Information (EPPI) Co-ordinating Centre, and the government's creation of the What Works Network to support public services (Cabinet Office, 2014). In the field of health professions education, the Best Evidence Medical and Health Professions Education (BEME) Collaboration was founded to develop the educational evidence-base through the production and dissemination of systematic reviews (Harden et al. 1999). Whilst such reviews are now increasingly used to identify 'what works', the existence of a systematic review on a particular topic only seldom results in the implementation of evidence-based policy or practice in that area. This is despite an emergent discourse in scholarly circles that emphasises the demonstration and monitoring of impact as a proxy metric of research quality (Bastow et al, 2014).

In this paper we aim to uncover the reasons for this juxtaposition. We take systematic reviews to include all studies that review the literature to analyse the evidence base, usually by synthesis, according to specified 'systematic' procedures. We do not limit this definition to meta-analyses, but to all forms of evidence review that attempt to synthesise evidence in response to a particular practice or policy issue. We argue that there are taken-for-granted

assumptions associated with the current practice of producing evidence syntheses that can undermine their usefulness in the educational arenas for which they are intended. We then set out some practical principles that the producers of evidence syntheses can use to inform their decision-making. We conceptualise these principles as the *Beyond Synthesis Impact Chain*, a framework aimed at producing conclusions that are useful and meaningful to practitioners and policymakers. The paper is authored by educational practitioners who are also involved in producing systematic reviews that supposedly make a contribution to the practices in which they are engaged. We each identify ourselves as both a practitioner and as a researcher in the arena of health professions education. This dual-identity has sharpened our focus on the disconnect between the production of synthesised evidence and its use as evidence-in-practice. In contradiction of the espoused methods, we have not undertaken a systematic reviewing process here: there was no systematic search strategy, no scanning, screening or pooling in the construction of this paper. What we offer is a reflective analysis of the sort that practitioners do when trying to understand their work and exploring how to improve their practice (Schön, 1983; Kolb, 1984). We make no apology for this. Our aim is to present principles that support the *praxis* of systematic reviewing based upon our experiential learning as both users and producers of health professions' educational evidence.

What does impact look like?

The evolution of impact as a marker of research quality has been tumultuous. The numerous criticisms of impact defined in terms of bibliometrics such as citation rates (e.g. David, 2008) have led assessors of research quality to redefine impact as 'the demonstrable contribution that…research makes to society and the economy' (Research Councils UK, 2014). The UK's Research Excellence Framework (REF) has pioneered societal impact case studies as an assessment measure of research quality, requiring universities to provide tangible evidence of non-academic impact that contributed a significant weighting (20%) to a submission's evaluation. REF (2011) defined impact as 'the effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia' with impact evaluated against the REF's (2012) criteria of *significance* (the intensity of the influence or effect) and *reach* (the stretch or breadth of influence on relevant constituencies). The validity of impact assessments is made more ambiguous though when we take account of the social complexities and proclivities of assessment panel members as Samuel and Derrick's (2015) analysis of REF evaluators' identifies.

Whilst submissions of impact case studies to the REF2014 documented a variety of societal benefits, the considerable challenge of making the case for impactful health professions education research were implied in the submission numbers. Of the 6,975 impact case studies submitted, 368 related to education (5%), but only four of these related to health professions education specifically. Despite a significant number of reviews published in health professions education during the assessment period, this finding suggests either that these reviews possessed limited impact, or that submitting institutions felt their impact did not meet the necessary standard for REF submission. Systematic reviews of health professions education were rare, though one impact case study made the case for their institution's systematic reviews' instrumental, conceptual and capacity-building influence (REF, 2015). We believe that producers of educational systematic reviews genuinely intend for their findings to support users to implement the most effective educational programmes, strategies and pedagogies for the users' specific context. However, we argue that whilst the focus of activity has been on generating the report of the synthesis as an end-product or output, monitoring its impact has inadvertently been neglected with the onus of implementation being placed firmly on the shoulders of practitioners and policymakers. Assuming that similar mechanisms for assessing impact will be applied in the future, there will be consequences for producers of synthesised evidence. The definitions of impact imply that it is a *demonstrable* phenomenon, 'an auditable or otherwise recordable occasion of influence' (Bastow et al., 2014, p.297). But monitoring the significance and reach of systematic reviews is fraught with challenges, especially if we conceptualise them as discrete end-products. Davies et al. (2005) identify strategies for tracing back the development of a specific policy or practice to its origins in the evidence base; or identifying the study's utility and tracing forward its use as it emerges in non-academic contexts to allow its impact to be articulated. Both approaches are problematic.

Martin (2011) cautions that the costs of monitoring impact could outweigh the benefits for public accountability. She suggests that capturing impact is labour-intensive and difficult, and even if it could be captured, attributing specific societal developments to specific studies remains tenuous. Sayer's (2015) sophisticated debate following publication of REF costs similarly questions the benefit of monitoring. However, the impact agenda is here to stay and, future research assessment exercises notwithstanding, we propose that building principles that *account for impact* into the planning and conduct of systematic reviews can

enhance their influence. By impact we mean the contribution that the systematic review makes to practice and policy. We also speculate on how contributions can be attributed to the review and how these might be captured. Whilst this will inevitably add to the labour of the process, we suggest that it will serve the triple purpose of more meaningfully *contributing* to educational practices, *demonstrating* impact, and consequently *justifying* future investment in evidence syntheses.

Why don't systematic reviews automatically impact on practice?

Producers of evidence (such as researchers and reviewers) and users of evidence (such as policymakers and practitioners) tend to agree that evidence-based practice is important to their work. Most producers want their research to have an impact that contributes to society, and most users want to show that policies and practices are supported by robust evidence. Yet both producers and users bemoan the insufficient impact that research evidence exercises on practice-based activities, often implicating each other's shortcomings when limited transference is observed (Hammersley 2005). Table 1 articulates the cases that users and producers of evidence have made against each other. Giluk and Rynes (2012) identify lack of relevance and resistance to change as factors that create distrust in research and inertia in the uptake of evidence-based practices. These are important issues for users of evidence. The practical implications of modifying behaviours, implementing alternative approaches or reforming organisations can be unsettling – both socially and economically – so users require evidence that is relevant, compelling, persuasive and accessible to even begin to consider modifying their approach.

INSERT TABLE 1 ABOUT HERE

Decision-making by policy-makers and practitioners is also influenced by a range of situational, temporal, economic and pragmatic factors that research evidence cannot possibly hope to fully respond to. Policy-makers can be more concerned with politically desirable – but poorly justified "quick-fixes" (Hattie, 2015), whilst practitioners employ judgement and

experience – practical wisdom – only occasionally implementing research evidence despite political and cultural pressure to do so (Freeman, 2001; Gabbay and Le May 2011). What the arguments in Table 1 also indicate is that the underlying assumptions about evidence-based practice require qualification. The fundamental principle of evidence-based practice – as a unidimensional process where producers produce evidence and users use it – assumes that research and practice are both essentially made of the same stuff, that they are both techno-rational activities (Webb, 2001). However, there are fundamental differences between the stuff of research and the stuff of practice. The former is controlled, paced and abstracted; the latter is messy, complex, immediate and contextualised. Much of the literature articulates this difference as a 'gap' or 'divide' (e.g. Bero et al, 1998; Grimshaw et al, 2012; Empson, 2013) that needs to be bridged, implicitly embedding the unidimensional assumption.

Systematic reviews are considered as one means of bridging this divide, drawing together and synthesising the evidence responding to a particular issue (Hammersley, 2013). However, the act of synthesising evidence creates a double-bind – conclusions, whilst arguably more robust when compared with individual studies, are also more abstracted. This, ironically, takes them further from the contexts within which they were originally derived, making it more difficult for users to see how they might implement recommendations in their own practices. The conduct of systematic reviews, and the contested nature of what counts as a systematic review, also relies on a number of practices that can be unhelpful for users of evidence. Firstly, some forms of systematic review rely upon strict inclusion criteria that privilege certain research methodologies over others, justified by the assertion that only the most robust research designs can offer valid findings. Screening, sifting and selecting acceptably rigorous studies often leads to a disappointing number of studies making the cut. Maclure (2005) criticises this process where sometimes thousands of potentially relevant studies are dismissed on the grounds of design, leaving a mere handful of primary studies for in-depth review. As a result, conclusions that 'more research is required' or 'insufficient evidence exists' are regular clichés appearing in the final sentences of many reviews. Petticrew (2003) suggests that this is because the more robust the research design, the less obvious the effect size. Maclure (2005: p.402) is more scathing. She argues that the process of sifting to tiny yields for in-depth review allows conclusions to be drawn that lay the blame at the feet of educational researchers for the poor quality of their reporting and 'the parlous state of their field'. Some forms of systematic reviewing also assume that studies are related to one another only in a cumulative sense – that all are studying the same phenomenon in

similar ways such that their findings can be aggregated (Davies, 2000; Hammersley, 2013). Maclure (2005: p. 399) argues that conceptualising evidence as 'nuggets of knowledge' to be extracted and aggregated misconstrues research knowledge as 'static, transparent and compliant with disciplinary boundaries' (2005, p.394). Silencing ambiguities, nuances and contexts may also serve to undermine the capacity of systematic reviews to influence practice. Whilst Maclure's criticisms neglect that some practical questions may require experimental designs with strict parameters for inclusion, they illuminate that often robust evidence risks being overlooked when it comes to informing policy and practice. Newer forms of synthesis, such as Pawson's (2006) realist synthesis, have attempted to account for context by re-conceptualising evidence as a patchwork of *mechanisms* leading to variable *outcomes* in diverse *contexts*; redefining the 'what works' question as 'what works, for whom, in which circumstances' and providing potential for more meaningfully bridging the gap between producers and users of evidence synthesis.

Illuminating the taken-for-granted assumptions of evidence-based practice and evidence synthesis help us recognise the hugely mediated relationships that exist between evidence produced and evidence used, and account for much of the variability observed in uptake and impact. Whilst we do not entirely repudiate the notion of bridging the theory-practice gap, we suggest that reconceptualising this gap as a *dialectic tension* (Bartunek and Rynes, 2014) best serves us in thinking about how evidence syntheses can be most usefully constructed for practice and policy. Bartunek and Rynes (2014) argue that tensions between academics and practitioners arise from differing logics, differing time horizons, differing communication practices and differing priorities vis-à-vis rigour and relevance. These tensions, we argue, provide the conditions for expansive learning (Engeström, 2001, Engeström and Sannino, 2010) where the tacit assumptions of competing perspectives may, through interaction and partnership, produce new and potentially impactful knowledge on the praxis of systematic reviewing. We have reflected upon our experiences as producers and users of evidence syntheses to determine how evidence syntheses might more usefully contribute to policymaking and practice. Like Mallet et al (2012) we argue that systematic reviews must be seen as a means to an end, not an end in themselves. We propose that producers of educational systematic reviews must look *beyond synthesis* when planning their review, experimenting with practical strategies that can optimise their impact. We seek to strengthen the integrity of systematic procedures by combining them with the flexibility and reflexivity of practical wisdom.

The Beyond Synthesis Impact Chain

So far in this paper we have explored definitions of impact and how evaluating impact is problematic. We have considered why systematic reviews have not consistently made meaningful contributions to practice and policy and, in debating the assumptions and tensions that characterise evidence-based practice, we have identified the following issues that undermine their impact:

1. Definitions of impact are immersed in the discourse of evidence-based practice which frequently assumes that research and practice are similar techno-rational activities;
2. The monitoring and assessment of impact are laborious and fraught with difficulty, particularly in making claims of direct attribution;
3. A unidimensional flow from evidence to practice is assumed that neglects the dialectic tensions between them;
4. The responsibility to apply evidence and implement evidence-based practice has been primarily cast as the responsibility of users rather than as a shared endeavour between producers and users;
5. Systematic review methods favour procedural techniques aimed at maximising transparency and rigour, and may inadvertently reduce relevance to users;
6. Reviewers can implicate primary research as poorly conducted or poorly reported, when it may be the review methodology that does not adequately fit the literature in a particular field;
7. Completion of the systematic review is often perceived to be the end of the process with the published report seen as the ultimate output.

In describing the *Beyond Synthesis Impact Chain* (BSIC) we hope to address some of these issues by articulating principles that explicitly bring impact optimisation strategies into the planning and conduct of systematic reviews. The premise of the BSIC, as shown in Figure 1, is that an impactful review requires (1) appropriately timed and high quality exchange between users and producers and (2) that the chain is seen as an interdependent structure rather than as distinct components. We consider each link in the chain in the subsequent sections of this paper, providing examples of good practice, illustrating how the links are interdependent and suggesting recommendations for users and producers of reviews. The elements of the chain are then integrated at the end of the paper. Taking the criticisms

levelled at research production in Table 1 as our starting point, the BSIC explains how the systematic review process can be augmented with practical wisdom to optimise demonstrable impact.

INSERT FIGURE 1 ABOUT HERE

'Practical' Problem Formulation and 'Mindful' Planning

As with all approaches to inquiry, the selection of the problem to be addressed focuses all subsequent activities in the inquiry. To optimise the impact of systematic reviews in we need to articulate their aims in relation to both the perspectives of users and the nature of practice. Schwandt (2014, p.232) identifies that problems faced by practitioners demand an action, taking the form of '**what** should I do now, in this situation, facing these circumstances?' If this problem can be addressed then practitioners might subsequently ask '….and **how** best might I do it?' These are inherently *practical* problems. In education, as in other forms of practice, decisions and actions are taken based on assumptions about desirable outcomes. Educational practices are necessarily teleological – premised on a purpose or aim. Biesta (2010, p.501) observes that 'there is no evidence to generate or collect if we do not…decide what the aim or purpose of the practice is. Evidence… needs to be "filtered" through decisions about what is educationally desirable'. In the rhetoric of educational systematic reviews, desirability is usually framed as the *effectiveness of an intervention* to meet predetermined educational outcomes. This implicates any decisions or actions as highly contingent on users' value judgements about what outcomes are desirable.

Schwandt (2014) distinguishes practical problems that demand an action from other problems that require an explanation. These problems might take the form '**why** does this approach in this situation lead to more desirable outcomes than others?' Explanatory problems, he argues, are not practical in nature. Despite this, Stevens et al.'s (2009) comparison of what funders fund and what practitioners need in children's services observed that funders favoured explanatory 'why' questions whilst practitioners favoured practical effectiveness

'what/how' questions. Practical problems focus on the pragmatics of efficacy, whilst explanatory problems can explain how context, relationships and interactions shape outcomes. Both can ultimately make a contribution. Honest discussions between users and producers that focus on utility can help to frame the most suitable problems to be addressed.

The importance of ensuring that systematic reviews address issues that are relevant to stakeholders is evident in the Cochrane Collaboration's principle of striving for relevance (Cochrane Collaboration, 2015). The emergent practice of engaging users in the systematic review process (Konnerup and Sowden, 2008) has also demonstrated attributable benefits in terms of topic refinement, review relevance and subsequent uptake (Cottrell et al., 2014). We advocate that strategies of mindful planning that engage users in helping to formulate practical problems are incorporated into the process of constructing systematic reviews in health professions education. The full inclusion of evidence users in the review team can create the dialectic tension required to reframe perspectives, ensuring that problems are scaffolded around contemporary practical dilemmas. When formulating the problem or issue that the review will address and planning the review, we propose that reviewers consider the strategies in Figure 2 to enhance the practical utility and impact of the review for end users.

INSERT FIGURE 2 ABOUT HERE

To exemplify the link between practical problem selection and mindful planning we shall compare two reviews that synthesised the evidence on internet-based education for health professionals. In the first review, Cook et al.'s (2008) objective was *to summarise* whether internet-based instruction is better at achieving desirable outcomes compared with no intervention or non-internet intervention. The selected outcomes of learner reaction/satisfaction; knowledge, skills and attitude; behaviours in practice, and effects on patients were based on Kirkpatrick's (1996; 2007) four level model, commonly used to categorise what is desirable in evaluations of health professions education. Despite heterogeneity across selected studies they conducted a meta-analysis, pooling effect sizes using a random effects model. This pooling averaged the effects across individual studies

and masked much of the variation between them. Cook et al.'s (2008) meta-analysis concluded that large effects are observed when internet-based instruction is compared to no intervention. When compared with non-internet-based instruction, the meta-analysis demonstrated no difference in effect sizes across the selected outcomes, but that variations between studies meant that in some circumstances internet-based instruction was more effective and in some circumstances less effective than non-internet-based instruction. So what does this review offer to the well-meaning health professions educational practitioner who is exploring alternate approaches to delivering their curriculum? They can infer that internet-based instruction achieves better outcomes than not teaching learners at all (which is perhaps unsurprising, as Cook's (2012) subsequent paper recognises), and that compared to other forms of teaching it might be a little better or a little worse. These inferences may be interesting but are not particularly helpful in addressing the 'what' and 'how' practical problems or the 'why' explanatory problems highlighted above. This is not to suggest that Cook et al.'s (2008) meta-analysis possesses no value. Its conclusion that there is limited value in future research comparing between internet-based instruction and no intervention controls, for example, makes a significant contribution to the research arena and the quality and transparency of the meta-analysis provides clarity for others wishing to undertake similar syntheses. However its value to responding to practical problems in health professions education remains restricted – a limitation also noted by the review team.

The second review by Wong et al. (2010) used realist synthesis to supplement Cook et al.'s meta-analysis, albeit focusing on medical education specifically. Their objectives followed Pawson's (2006) realist line: *to explain what* sort of internet-based education *works, for whom and in which circumstances*. In addition they aimed to produce pragmatic guidance for course designers and developers to optimise their courses and guidance for learners to evaluate the suitability of courses. These objectives taken at face-value seem to more clearly respond to the explanatory 'why' problem and the practical 'what' and 'how' problems. Wong et al. (2010) identified potential explanatory theories for achievement of Kirkpatrick's outcomes that were then tested against the sourced studies. In so doing, contextual conditions that optimise engagement and interactivity in internet courses were reported, including learners' needs and the course's context. The authors admitted that their conclusions fall short of their original intentions as they could not test all their candidate theories. For the health professions educator the findings of the Wong et al (2010) realist synthesis provide more scope for making decisions and, to some extent, taking practical action. The framing of

their objectives certainly resonate with the practical problems faced by users of evidence, and progress the inferences that can be extracted from Cook et al.'s (2008) findings.

Mindful planning of reviews requires producers to engage with users to consider the practical needs of those whom the anticipated conclusions and recommendations are for, and focus their review accordingly. When a review is commissioned, value is intrinsically attached to the 'answer'. Reviews are more likely to be commissioned where the anticipated value exceeds the cost of production. There is a risk that the problem focuses on the agenda of the commissioner, possibly ignoring the needs of other evidence users. Producers therefore need to be cognisant of identifying and engaging with a wide range of users, even with commissioned studies. This is not to say that reviews emanating from producers never possess demonstrable impact. For example, a meta-analysis of the effect of ethnicity on academic performance of medical students and doctors by Woolf et al. (2011) led the General Medical Council to analyse and report ethnicity data from the National Trainee Survey. The review does not indicate how educators can address this issue, but recommends future research to explain the causes of the performance gap and test interventions for improvement.

Relevant Reviewing and Reflexive Synthesis

Systematic reviews and forms of synthesis rely on robust and transparent procedures. We do not wish to re-introduce these here and refer interested readers to Kastner et al.'s (2012) excellent protocol for a scoping review of systematic knowledge synthesis methods which classifies twenty-five different forms of synthesis, and makes reference to the methodological and procedural texts associated with each form of knowledge synthesis. Rather, we consider the practices of reviewing and synthesising, and explore how these practices can be augmented with the principles of relevancy and reflexivity to enhance the impact of the review. Careful selection of the review methodology that most relevantly responds to the practical issue identified in the review question(s) to produce meaningful conclusions is the first fundamental decision that needs to be addressed. Once selected, its procedures determine what evidence is included and how (or if) that evidence will be synthesised.

The pre-synthesis procedures of review methodologies consist of searching the evidence, selecting pertinent studies against defined inclusion criteria and extracting data from these. Figure 3 shows how the evidence that *should* be included in a particular review is whittled

down to the evidence that *is* included. The square brackets explain how potential evidence that *should* be included may be 'lost' at each stage of the review process. It is also possible that evidence that does not meet the inclusion criteria is incorrectly included in the review. Not all losses are the fault of the review team. For example, while the level of publication bias can be estimated using a funnel plot (see Egger et al., 1997), a review team cannot include studies that have never been written up. However, the potential effect of publication bias can and should be modelled (Turner et al., 2009). Similarly, the review team are restricted if the original studies are of insufficient 'quality' to meet the inclusion criteria, although the subjectivity of reviewer-imposed standards needs to be recognised and challenged, since no single agreed set of quality standards exists (e.g. Armijo-Olivo et al., 2012). Some losses can be ameliorated by employing the strategies considered in Figure 4.

INSERT FIGURE 3 ABOUT HERE

INSERT FIGURE 4 ABOUT HERE

A review team may specify strict inclusion criteria for pragmatic reasons, such as limited resources to undertake the review, but this can come at the expense of the review being meaningful and relevant to users. Paradoxically, systematic design decisions that delimit the evidence included could have been the result of the review team responding to stakeholders' demands for rapid results. Systematic reviews undertaken 'voluntarily', not beholden to external commissioners, may provide a more complete evidence base, but may fail to optimise impact (Woodman et al., 2012), especially if users have not been engaged in the process.

The discussion so far has suggested that a plethora of evidence exists that may be included in a review. This is not always the case; reviews can even be 'empty'. This could occur because of the way in which evidence was sought, screened and quality-assessed (e.g. use of too narrow inclusion criteria) or because the original studies simply do not exist. In the latter case, a review can still be impactful in the long term if it leads to the required research being

undertaken. Here, a scoping review is useful to estimate the probability of an empty review, which should be discussed with users and potential modifications to the process of identifying evidence agreed. Such work will therefore initiate 'expectation management' for stakeholders, by highlighting at an early stage that an empty review could be possible.

Whilst it important to ensure all the relevant evidence is captured, review teams must also reflect upon selecting an appropriate method of synthesis. Identifying the approach most likely to facilitate useful conclusions for the practical problems being investigated is a core requirement in optimising impact. Reviewers need to be mindful of not importing their own 'pet' methodological preferences, and must make reflexive judgements about the suitability of alternative forms of synthesis. Here again, a scoping review can be useful: informing final search strategies, providing an insight into the nature of the data that are likely to emerge and helping to narrow-down a long-list of potential synthesis methods. Arksey and O'Malley (2005) identify how scoping reviews that incorporate consultation with key informants can shift focus and provide more useful results.

INSERT FIGURE 5 ABOUT HERE

Figure 5 illustrates how alternative synthesis methods may lead to contrastingly-framed conclusions and recommendations, even when the same or similar selected studies are synthesised. This can be further exemplified by contrasting two reviews that address a similar issue: the use of practice simulation as a teaching and learning strategy in health professions education. Issenberg et al.'s (2005) **qualitative synthesis** concluded that 'the research evidence is clear that high-fidelity medical simulations facilitate learning among trainees when used *under the right conditions*' (p. 24, emphasis added). These conditions included provision of feedback, integration into the curriculum and repetitive practice. For the last of these conditions, forty-three papers were found that identified repetitive practice as a key feature of effective simulation. Nevertheless, repetitive practice was subsequently the subject of a **quantitative meta-analysis** undertaken by the same research team (McGaghie et al., 2011). The authors stated that the meta-analysis responded to a 2007 call by educational

leaders from US medical schools regarding the need for research into the efficacy of simulation for enhancing performance – implying that the original review had not provided sufficiently convincing evidence of effectiveness. The latter review concluded that 'the meta-analysis outcomes favouring simulation-based medical education with deliberate practice [over traditional methods] are powerful, consistent and without exception' (McGaghie et al., 2011, p.709). What is difficult, however, is attributing the subsequent expansion of simulation in medical curricula (Pelletier, 2015) to such research evidence (i.e. assessing its impact relative to other motivators to use simulation). Furthermore, as is traditional in research, the authors note the need for further research, particularly to examine both cost-effectiveness and the organisational impacts of adopting simulation-based medical education with deliberate practice.

In this section we have examined the need to make review method decisions likely to produce findings that are most relevant to practice. We have also demonstrated that the process of synthesis – turning data from individual studies into synthesised results – needs to not only be procedurally transparent, but also reflexively transparent by identifying and reporting how their choice of synthesis may have been influenced by the beliefs, backgrounds and preferences of those involved. This can help review teams to justify their choice. Next we scrutinise methods for reporting and dissemination, aiming to optimise impact beyond synthesis.

'Meaningful' Reporting and 'Accessible' Dissemination

As approaches to reviewing have diversified so publication standards have been developed to optimise consistency amongst reports. These standards include PRISMA for meta-analyses and reviews of experimental studies (Liberati et al., 2009), RAMASES for realist and meta-narrative reviews (Wong et al., 2013a; 2013b) and ENTREQ for qualitative synthesis (Tong et al., 2012). These standards are differentiated by design. In health professions education, where multiple forms of systematic review have been conducted under the auspices of the BEME Collaboration, Gordon and Gibbs (2014) published the STORIES statement to support report writing that 'offers most to readers'. Each of these standards offers excellent guidance to reviewers with each providing examples and explanations of included items, and we would recommend that reviewers use the appropriate standard to support their writing.

The fundamental work of these reporting guidelines is to standardise the report and maximise transparency of the review process, enabling judgements to be made about their methodological quality. With the exception of ENTREQ, each standard provides relatively little insight into the sorts of content that might provide practical utility. Whilst the STORIES statement usefully encourages findings to be presented in light of the review objectives, the utility of these will be conditional upon how practically-oriented these objectives are. Gordon and Gibbs (2014) do though recommend that the report contains *descriptions of what effective education looks like*, providing fundamental insights for users. Discussing the implications of review findings is also amongst the items listed, but how reviewers might do this is less clearly specified. PRISMA asks reviewers to consider the findings' relevance to stakeholders; RAMASES indicates that recommendations for policy and practice can be offered *if appropriate*; whilst STORIES and PRISMA suggest highlighting impact and implications, respectively, for future research. ENTREQ though invites reviewers to 'present rich, compelling and *useful* results that go beyond a summary of the primary studies' (Tong et al., 2012, p.4, emphasis added). This statement is the only indication provided in any of the standards that the report should be useful.

When constructing the report, reviewers may interpret publication standards rigidly and risk omitting or disguising potentially useful content amongst technical procedures thereby missing opportunities (Gordon et al., 2014). Readers may find useful recommendations, but for the educational practitioner there may be limited detail to go on in terms of how they might modify or improve their practices. Here we suggest that systematic reviewers can document their conclusions meaningfully without deviating from the timbre of publication standards. When constructing the report we propose that reviewers consider the strategies in Figure 6 to enhance meaningfulness and utility. These strategies are not intended to be laborious, but to form a recognised part of the report-writing process.

INSERT FIGURE 6 ABOUT HERE

Examples of systematic reviews that have reported meaningful recommendations in the field of health professions education include a realist review of longitudinal practice placements (Thistlethwaite et al., 2013), a thematic review of role modelling amongst doctors as clinical educators (Passi et al., 2013) and a meta-ethnography of clinical skills learning in general practice (Park et al. 2015). These reports provide insightful recommendations for current practice and future research in their respective areas and included users amongst the review team. The review by Thistlethwaite et al. (2013) in particular aligns its report to the strategies proposed in Figure 6. Their report offers suggestions for how longitudinal placements can be implemented effectively and includes case studies of where these have delivered desirable outcomes.

The completion of the report is commonly seen as the final step in the review process, offering a product to be consumed by the users of evidence. Whilst we propose that strategies discussed so far in this article may aid uptake, we have also implied that considerable thought and planning needs to be given to post-review activities to optimise impact. Learning from the challenges experienced in the related field of clinical practice and integrating guidelines on implementation and impact, we identify strategies that reviewers can plan and implement to disseminate their work by making it accessible and targeting their energies towards those who might find their reviews most useful. Research Councils UK's (2014) *Pathway to Impact* guidelines recommends the development of specific strategies for engagement and dissemination. Such strategies include organising practitioner workshops, producing a lay summary of the review and, if possible, actionable implementation guidance to be distributed to key stakeholders, open-access publication, and exploration of the infrastructural conditions necessary to implement recommendations. Analysing infrastructural conditions and contexts can help reviewers appreciate the complexity of practice, identify the characteristics, priorities and beliefs of users and institutions and monitor the barriers and enablers to evidence use (Boerner et al. 2015). In health professions education the barriers and enablers to uptake for various potential users (including, curriculum designers, managers, clinical teachers, policy-makers, commissioners, regulatory and professional bodies, educational researchers, students, service users and the wider public) could be mapped by review teams and accounted for when planning dissemination activities. Such mapping, as exemplified by Grimshaw et al. (2012), may allow a targeted approach when exchanging the review's messages with different users.

As we have discussed, users may resist adapting their practices even in the presence of convincing evidence. Communicating the review's messages in terms of its *significance and benefits* to users, together with clear *guidance on application and implementation*, may prove more persuasive. Figure 7 provides a specific example of the dissemination impact strategy targeted at educators of a review currently being undertaken. Other less targeted strategies that can otherwise still increase accessibility to review findings include depositing reviews with Dissemination Centres and databases and publicising the review on social media. In the health professions education arena, sharing information through Twitter and other social media has increased in popularity and scope. According to Micieli et al. (2015) the *#MedEd* hashtag is considered the gold standard for immediate medical education news, so publicising reviews through this medium may lead to unanticipated, but beneficial demonstrable impacts.

INSERT FIGURE 7 ABOUT HERE

Other considerations that review teams need to bear in mind include monitoring the literature subsequent to the review so that future updates can be scheduled, and exploring how the impact of their work can be evaluated. Reconceptualising review reports from stand-alone products to 'spot-checks' on best available evidence, protocols for updating the review – either in light of emergent evidence or by considering alternate synthesis methods – can be drafted shortly after completion of the review. Kastner et al.'s (2012) scoping review protocol provides an excellent example of thinking ahead in terms of identifying potential impact. In considering mechanisms and strategies for evaluating impact, review teams will need to consider the significance and reach of their work, and anticipate the timescales over which impact might be demonstrated. Van Eerd et al. (2011) and Milat et al. (2015) review strategies and tools for assessing implementation and impact of knowledge transfer and exchange practices and review teams. They recognise that it takes time and skill to build evaluation into practice, and that limited instruments are available for capturing impact. We encourage review teams to evaluate the demonstrable impact of their work, as resources

allow, and propose that ongoing knowledge exchange activities may both facilitate demonstrable impact and contribute to its measurement.

Integration across the *Beyond Synthesis Impact Chain*

Grimshaw et al. (2012) identify mechanisms of knowledge translation as 'push' activities (the efforts of researchers to make the results of their research available to policy makers and users); 'pull' activities (the efforts of policy makers and others to access research evidence for decision-making) and 'exchange' activities (the building of relationships between policy makers and researchers that facilitate knowledge transfer on an ongoing basis). The *Beyond Synthesis Impact Chain* advocates the last of these, employing 'engaged scholarship' (van de Ven, 2007; McCormack, 2011) to optimise co-production of systematic reviews between users and producers. We have argued that exchanges between users and producers create the dialectic tension necessary to transform perspectives, address the challenges of evidence-based policy and practice, and have the potential to make systematic reviews considerably more impactful. The *Beyond Synthesis Impact Chain* articulates the practical strategies that review teams can use to strengthen review processes through interaction, engagement and partnership with users. These processes are not intended to be presented sequentially, as mutually exclusive elements, but as an integrated whole. We propose that successfully demonstrating optimal impact is highly contingent upon selecting problems that are inherently *practical*; planning review processes *mindful* of the needs and priorities of users; reviewing for most practice-*relevant* sources; being *reflexive* about the choice of synthesis method; ensuring that reported findings and recommendations are *meaningful* to users; and identifying mechanisms for making these as *accessible* as possible through targeted dissemination. The chain is also factorial. Like van der Vleuten's (1996) assessment utility formula, if any of the links in the chain are under-developed then the whole chain is compromised and demonstrable impact diminished. By considering these strategies at the pre-planning stage and articulating decisions made in the review protocol, examples of full review *praxis* that flexibly account for impact can be examined, scrutinised and reviewed against the desirable outcome of 'impact'.

Conclusions

This article makes a critical and methodological contribution to the 'what works' discourse, arguing that demonstrable impact can be optimised by consolidating the rigorous procedures of systematic review methodologies with practically-oriented principles. Whilst we have drawn upon examples from the health professions education literature, we contend that these principles are potentially transferable to the undertaking of systematic reviews across educational disciplines and the wider social sciences. In building principles that account for practical impact into the planning and conduct of systematic reviews we have engaged in the 'reflective observation' and 'abstract conceptualisation' phases of Kolb's (1984) experiential learning cycle. We plan to follow that cycle through to 'active experimentation' with the strategies of the *Beyond Synthesis Impact Chain* in our current and future engagement in systematic reviews, and encourage others to do the same. To our knowledge there exist no empirical studies that examine the clustering of principles described in the *Beyond Synthesis Impact Chain*. As a result, our proposals are to some extent speculative, though we have attempted to make clear reference to exemplar sources whose practices align with these principles.

We concede that the principles and strategies described here possess resource implications; however we would remind funders, commissioners and review teams that a systematic review that lacks impact is a fundamental waste of resource altogether. Relatively small time and financial investments may well bring a significant yield in terms of demonstrable impact and added-value. Some critics may also consider that augmenting rigorous review procedures with practically-oriented principles muddies the waters, and may insist that doing so enhances neither research nor practice. In terms of accounting for impact, such criticisms – like our advocacy – await the evidence of their convictions. We also acknowledge that the study of impact assessment is an embryonic field, driven in part at least by a neoliberal desire for institutional accountability. Milat et al.'s (2015) review of impact assessment models identifies citation analysis, interviews with principal investigators, peer assessment, case studies and document analysis as the primary mechanisms of impact assessment. Only four of the thirty-one studies they reviewed made any attempt to gain the perspectives of end users. Whilst citations and biblometric data provide proxy, and possibly dubious, indicators of impact (Pölkki et al. 2012), they remain influential amongst the academy, as exemplified by a recent analysis of top-cited articles in medical education (Azer, 2015). It is likely that multidimensional, practice-oriented accounts of impact will play a significant role in assessing the quality of systematic reviews in the future. We therefore contend that

integrating the principles and strategies of the *Beyond Synthesis Impact Chain* into the conduct of systematic reviews can optimise demonstrable impact and improve outcomes in policy and in practice and make the contribution to s.

References

Arksey H and O'Malley L (2005) Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*; 8, 1: 19-32.

Armijo-Olivo S, Stiles CR, Hagen NA, Biondo PD and Cummings GG (2012) Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *Journal of Evaluation in Clinical Practice*; 18, 1: 12-18.

Azer SA (2015) The top-cited articles in medical education: a bibliometric analysis. *Academic Medicine* [published online ahead of print]; 90, 8. doi: 0.1097/ACM.0000000000000780

Bartunek JM and Rynes SL (2014) Academics and Practitioners are alike and unalike: the paradoxes of academic-practitioner relationships. *Journal of Management*; 40, 5: 1181-1201.

Bastow S, Dunleavy P and Tinkler J (2014) *The impact of the social sciences: how academics and their research make a difference*. London: Sage.

Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD and Thomson MA (1998) Closing the gap between research and practice: An overview of systematic reviews of interventions to promote the implementation of research findings. *British Medical Journal*; 317, 7156: 465-468.

Boerner KE, Coulombe JA and Corkum P (2015) Barriers and Facilitators of Evidence-Based Practice in Pediatric Behavioral Sleep Care: Qualitative Analysis of the Perspectives of Health Professionals. *Behavioral Sleep Medicine*; 13, 1: 36-51

Cabinet Office (2014) *What Works? Evidence for decision makers*. London: Cabinet Office. Online at: https://www.gov.uk/government/publications/what-works-evidence-for-decision-makers [accessed 3 June 2015]

Cochrane Collaboration (2015) *Our vision, mission and principles*. Available online at http://www.cochrane.org/about-us/our-vision-mission-and-principles [accessed 23 June 2015]

Cook DA (2012) If you teach them, they will learn: why medical education needs comparative effectiveness research. *Advances in Health Sciences Education*; 17: 305-310.

Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ and Montori VM (2008) Internet-based learning in the health professions: a meta-analysis. *Journal of the American Medical Association*; 300: 1181-1196.

Cottrell E, Whitlock E, Kato E, Uhl S, Belinson S, Chang C, Hoomans T, Meltzer D, Noorani H, Robinson K, Schoelles K, Motu'apuaka M, Anderson J, Paynter R, Guise JM (2014) *Defining the Benefits of Stakeholder Engagement in Systematic Reviews*. AHRQ Publication No. 14-EHC006-EF.: Agency for Healthcare Research and Quality. Available online at www.effectivehealthcare.ahrq.gov/reports/final.cfm [accessed 23 June 2015]

David ME (2008) Research Quality Assessment and the Metrication of the Social Sciences. *European Political Science*; 7: 52-63.

Davies P (2000) The relevance of systematic reviews to educational policy and practice. *Oxford Review of Education*; 26, 3: 365-378.

Davies H, Nutley S and Walter I (2005) *Approaches to assessing non-academic impact*. Report of an ESRC Symposium, Swindon, ESRC.

Egger M, Smith GD, Schneider M and Minder C (1997) Bias in meta-analysis detected by a simple graphical test. *British Medical Journal*; 315, 7109: 629-634.

Engeström Y (2001) Expansive learning at work: Toward an activity-theoretical reconceptualization. *Journal of Education and Work*; 14, 1: 133-156.

Engeström Y and Sannino A (2010) Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review*; 5, 1: 1-24.

Evans J and Benefield P (2001) Systematic reviews of educational research: does the medical model fit? *British Education Research Journal*; 27, 5: 527-541.

Freeman AC and Sweeney K (2001) Why general practitioners do not implement evidence. *British Medical Journal*; 323, 7321: 1100-1102.

Gabbay and Le May A (2011) *Practice-based evidence for healthcare: Clinical Mindlines*. Abingdon: Routledge.

Giluk TL and Rynes SL (2012) Research findings practitioners resist: Lessons for management academics from evidence-based medicine. In D Rousseau (ed.) *The Oxford Handbook of Evidence-Based Managemen*t (pp.130-164). Oxford: Oxford University Press.

Gordon M, Carneiro AV, Patricio M and Gibbs T (2014) Missed opportunities in health care evidence synthesis. *Medical Education*; 48: 644–645.

Gordon M and Gibbs T (2014) STORIES statement: Publication standards for healthcare education evidence synthesis. *BMC Medicine*; 12, 143.

Grimshaw JM, Eccles MP, Lavis JN, Hill SJ and Squires (2012) Knowledge translation of research findings. *Implementation Science*; 7: 50.

Hammersley M (2005) The myth of research-based practice: The critical case of educational inquiry. *International Journal of Social Research Methodology*; 8, 4: 317-330.

Hammersley M (2013) *The myth of research-based policy and practice*. London: Sage.

Harden RM, Grant J, Buckley G and Hart IR (1999) BEME Guide No 1: Best Evidence Medical Education. *Medical Teacher*, 21, 6: 553-562.

Hattie J (2015) *What doesn't work in education: The politics of distraction*. London: Pearson.

Issenberg SB, Mcgaghie WC, Petrusa ER, Gordon DL and Scalese RJ (2005) Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher*; 27, 1: 10-28.

Kastner M, Tricco AC, Soobiah C, Lillie E, Perrier L, Horsley T, Welch V, Cogo E, Antony J and Straus SE (2012) What is the most appropriate knowledge synthesis method to conduct a review? Protocol for a scoping review. *BMC Medical Research Methodology*; 12, 114.

Kirkpatrick D (1996) Revisiting Kirkpatrick's four-level model. *Training and Development*; 50, 1: 54-59.

Kirkpatrick D (2007) *The Four Levels of Evaluation: measurement and evaluation*. Alexandria: American Society for Training and Development.

Kolb D (1984) *Experiential learning: experience as the source of learning and development*. New Jersey: Prentice Hall.

Konnerup M and Sowden A (2008) *User involvement in the systematic review process*. Campbell Collaboration Policy Brief from Campbell Collaboration Users Group. Available online at http://www.campbellcollaboration.org/artman2/uploads/1/Involvement_in_review_process.pdf [accessed 23 June 2015]

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J and Moher D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *British Medical Journal*; 339:b2700.

Maclure M (2005) Clarity bordering on stupidity: where's the quality in systematic review? *Journal of Education Policy*; 20, 4: 393-416.

Mallett R, Hagen-Zanker J, Slater R and Duvendack M (2012) The benefits and challenges of using systematic reviews in international development Research. *Journal of Development Effectiveness*; 4,3: 445-455.

Martin BR (2011) The Research Excellence Framework and the 'impact agenda': are we creating a Frankenstein monster? *Research Evaluation*; 20, 3: 247-254.

McCormack B (2011) Engaged scholarship and research impact: integrating the doing and using of research in practice. *Journal of Research in Nursing*; 16, 2: 111-127.

McGaghie WC, Issenberg SB, Cohen ER, Barsuk JH and Wayne DB (2011) Does Simulation-based Medical Education with Deliberate Practice Yield Better Results than Traditional Clinical Education? A Meta-Analytic Comparative Review of the Evidence. *Academic Medicine*; 86, 6: 706–711.

Micieli AMMI; Frank JR and Jalali, A (2015) A Medical Educator's Guide to #MedEd. *Academic Medicine* [published online ahead of print]; 90, 8.
doi: 10.1097/ACM.0000000000000776

Milat AJ, Bauman AE and Redman S (2015) A narrative review of research impact assessment models and methods. *Health Research Policy and Systems*; 13: 18.

Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA and PRISMA-P Group (2015) Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement, *Systematic Reviews*; 4, 1.doi:10.1186/2046-4053-4-1

Passi V, Johnson S, Peile E, Wright S, Hafferty F and Johnson N (2013) Doctor role modelling in medical education: BEME Guide No.27. *Medical Teacher*; 35, 9: e1422-e1436.

Pawson R (2006) *Evidence-based policy: a realist perspective*. London: Sage.

Pelletier SG (2015) *Use of Simulation Expands in Medical Education*. Association of American Medical Colleges Reporter – February 2015. Available online at: https://www.aamc.org/newsroom/reporter/february2015/426092/simulation.html [accessed 23 June 2015]

Petticrew M (2003) Why certain systematic reviews reach uncertain conclusion. *British Medical Journal*; 326: 756-758.

Pölkki T, Kanste O, Kääriäinen M, Elo S and Kyngäs H (2012) The methodological quality of systematic reviews published in high-impact nursing journals: a review of the literature *Journal of Clinical Nursing*; 23, 315–332.

Research Councils UK (2014) *Pathways to Impact*. Available online at: http://www.rcuk.ac.uk/innovation/impacts/ [accessed 19 June 2015]

Research Excellence Framework (2011) *Assessment framework and guidance on submissions*. Available online at http://www.ref.ac.uk/media/ref/content/pub/assessmentframeworkandguidanceonsubmissions/GOS%20including%20addendum.pdf [accessed 12 July 2013]

Research Excellence Framework (2012) *Panel criteria and working methods*. Available online at http://www.ref.ac.uk/media/ref/content/pub/panelcriteriaandworkingmethods/01_12.pdf [accessed 12 July 2013]

Research Excellence Framework (2015) *Systematic reviewing: building capacity for better informed policy-making*. Impact case study from Institute of Education. Available online at http://impact.ref.ac.uk/casestudies2/refservice.svc/GetCaseStudyPDF/44325 [accessed 19 June 2015]

Rubin A and Bellamy J (2012) *Practitioner's guide to using research for evidence-based practice*. New Jersey: Wiley.

Samuel GN and Derrick GE (2015) Societal impact evaluation: Exploring evaluator perceptions of the characterization of impact under the REF2014. *Research Evaluation.* Advance Access Online at: http://rev.oxfordjournals.org/content/early/2015/04/16/reseval.rvv007.full.pdf+html [accessed 3 June 2015]

Sayer D (2015) *Why the £246m tab for the REF is over three times as much as the RAE (it's not just because of impact)*. Maximising the impact of academic research – the impact blog. LSE. Online at: http://blogs.lse.ac.uk/impactofsocialsciences/2015/08/03/why-did-the-2014-ref-cost-three-times-as-much-as-the-2008-rae-hint-its-not-just-because-of-impact/#author [accessed 9 October 2015]

Schön D (1983) *The Reflective Practitioner: how professionals think in action*. London: Temple Smith.

Schwandt TA (2014) On the mutually informing relationship between practice and theory in evaluation. *American Journal of Evaluation*; 35, 2: 231-236.

Solesbury W (2001) *Evidence-based policy: whence it came and where it's going*. ESRC UK Centre for Evidence-Based Policy and Practice. Working Paper 1.

Stevens M, Liabo K, Witherspoon S and Roberts H (2009) What do practitioners want from research, what do funders fund and what needs to be done to know more about what works in the new world of children's services? *Evidence and Policy*; 5, 3: 281-294.

Thistlethwaite JE, Bartle E, Chong AAL, Dick M-L, King D, Mahoney S, Papinczak T and Tucker G (2013) A review of longitudinal community and hospital placements in medical education: BEME Guide No.26. *Medical Teacher*; 35, 8: e1340-e1364.

Tong A, Flemming K, McInnes E, Oliver S and Craig J (2012) Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. *BMC Medical Research Methodology*; 12, 181.

van der Vleuten C (1996) The assessment of professional competence: developments, research and practical implications. *Advances in Health Science Education*; 1: 41 – 67.

van de Ven A (2007) *Engaged Scholarship: A Guide for Organizational and Social Research*. Oxford: Oxford University Press.

van Eerd D, Cole D, Keown K, Irvin E, Kramer D, Gibson J, Kohn M, Mahood Q, Slack T, Amick BC, Phipps D, Garcia J, Morassaei S (2011) *Report on Knowledge Transfer and Exchange Practices: A systematic review of the quality and types of instruments used to assess KTE implementation and impact*. Toronto: Institute for Work and Health.

Webb SA (2001) Some considerations on the validity of evidence-based practice in social work. *British Journal of Social Work*; 31: 57-79.

Wells P (2007) New Labour and evidence-based policy-making: 1997-2007. *People, Place and Policy Online*; 1, 1: 22-29. Online at: http://extra.shu.ac.uk/ppp-online/issue_1_220507/documents/new_labour_evidence_base_1997-2007.pdf [accessed on 31 May 2015]

Wong G, Greenhalgh T and Pawson R (2010) Internet-based medical education: a realist review of what works, for whom and in what circumstances. *BMC Medical Education*, 10, 12.

Wong G, Greenhalgh T, Westhorp G, Buckingham J and Pawson R (2013a) RAMESES publication standards: realist syntheses. *BMC Medicine*; 11:21.

Wong G, Greenhalgh T, Westhorp G, Buckingham J and Pawson R (2013b) RAMASES publication standards: meta-narrative reviews. *BMC Medicine*; 11, 20.

Woodman J, Thomas J and Dickson K (2012) How explicable are differences between reviews that appear to address a similar research question? A review of reviews of physical activity interventions. *Systematic Reviews;* 1, 37

Woolf K, Potts HWW and McManus IC (2011) Ethnicity and academic performance in UK trained doctors and medical students: systematic review and meta-analysis. *British Medical Journal*; 342, 7803: d901.

Table 1: How producers and users implicate each other for the failures of evidence-based practice (see also Hammersley 2005; Giluk and Rynes 2012)

Users' criticisms of research:	**Producers' criticisms of practice:**
It does not focus upon specific practical problems in my day-to-day work **(Irrelevant and impractical)**	*It is set in its ways, unwilling to be challenged and resistant to new perspectives* **(Resistant)**
I*t generates conflicting and confusing evidence that provides no obvious recommendation for the context of my practice* **(Meaningless and unpersuasive)**	*It lacks the skills and capacity to understand and utilise research findings* **(Unscholarly)**
It generates conclusions that are at times over elaborate; qualified by limitations; jargon-ridden; or poorly disseminated; rendering them inaccessible. **(Inaccessible)**	*It chooses to not seek out research that might challenge embedded practices, and favours research that validates current practice* **(Risk-averse)**

Figure 1: The Beyond Synthesis Impact Chain

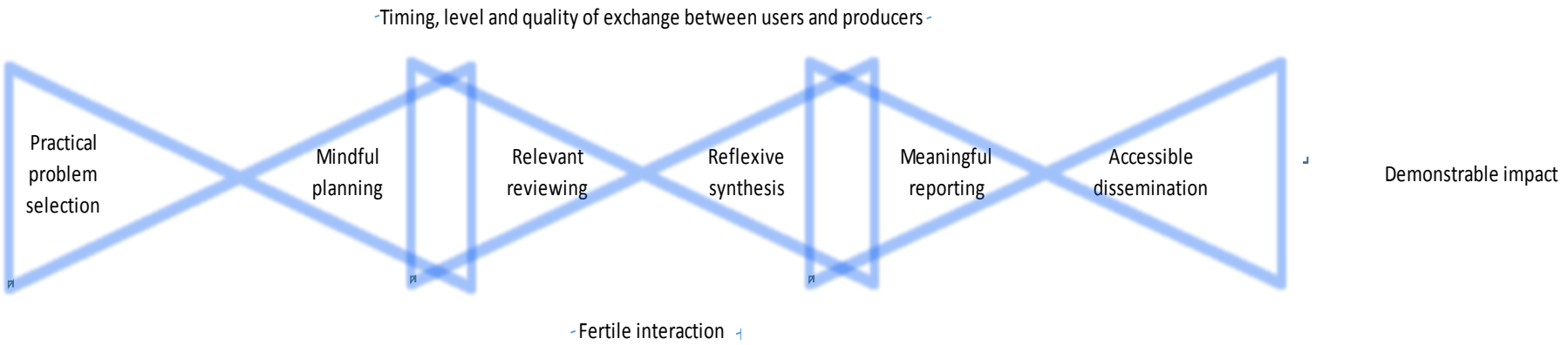


Figure 2: Strategies to consider in the formulating of practical problems and the mindful planning of impactful reviews

- *Ensure review aims/questions are practically-focused.*
  Drawing on the practical expertise of users, if possible, identify what changes in practice and/or policy are most desirable. Frame review aims/questions in such a way as to respond to these. If the aim is to explore and summarise the effectiveness of a particular intervention, then ensure that supplementary questions allow the conditions for successful outcomes to be articulated. Contextualise the issue within the practice and policy landscape. Make and report reflexive judgements about how beliefs, backgrounds, and preferences of those involved may have influenced the framing of the question.

- *Plan how users can be practically engaged in the full review process.*
  Once the problem has been articulated explore how engagement in the process can be achieved. In some cases, engagement activities may already exist or local conditions may be facilitative to full user involvement; in others a partnership between the review team and a user steering group may be more appropriate. If possible access user networks to request reciprocal networking. Construct the review protocol collaboratively, mindful of the ways in which the review could be potentially impactful.

- *Anticipate the form of recommendations that will most benefit users.*
  Identify how practice contexts, infrastructure and institutional structure may impede or facilitate uptake of recommendations, and structure the review to account for these factors. Accept that impact and uptake will vary and may not be immediately obvious and be tolerant of users' needs. Explore with users how recommendations should be articulated.

- *Consider dissemination processes at the planning stage.*
  Write strategies for engagement and exchange into the review protocol and ensure these are followed through. Work in partnership with user organisations to organise dissemination events such as workshops and seminars. Ensure that a short lay summary of the review is openly accessible and circulated to potential users. Consider open-access publication of the review.

Figure 3: Possible sources of evidence loss during systematic review

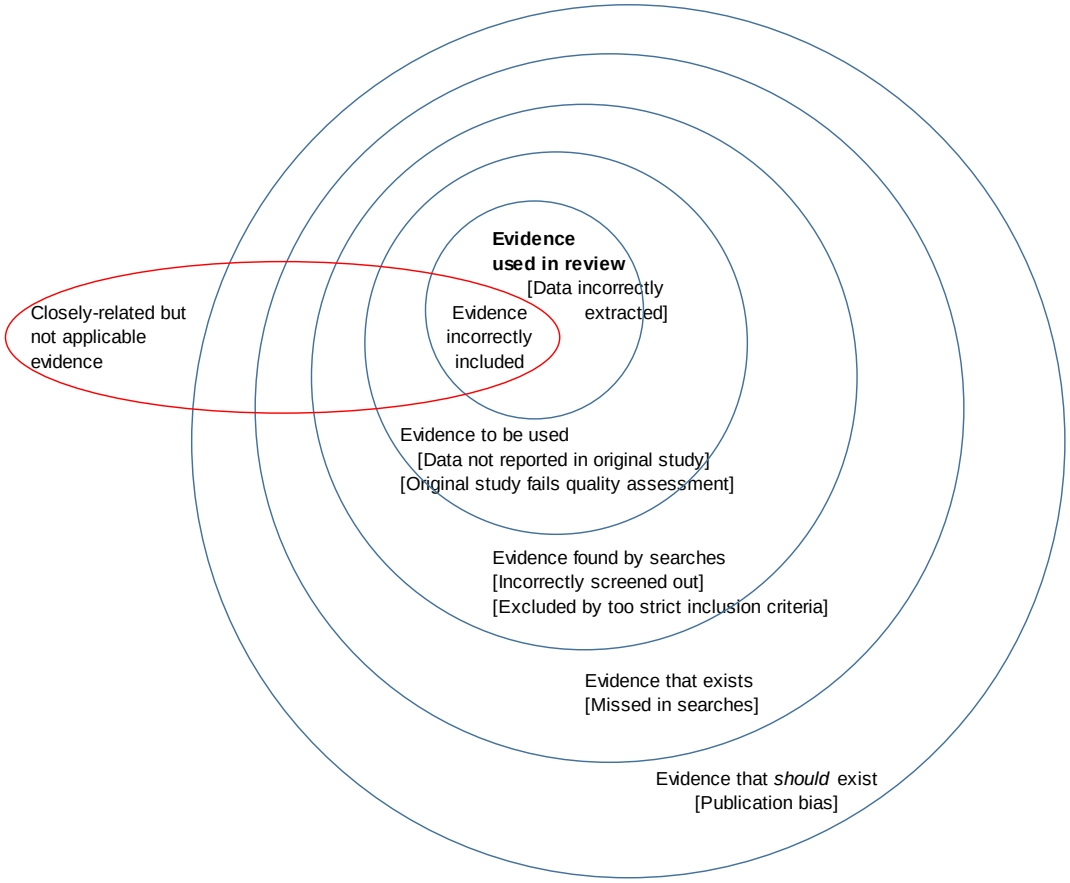


Figure 4: Augmenting the review process to reduce evidence loss

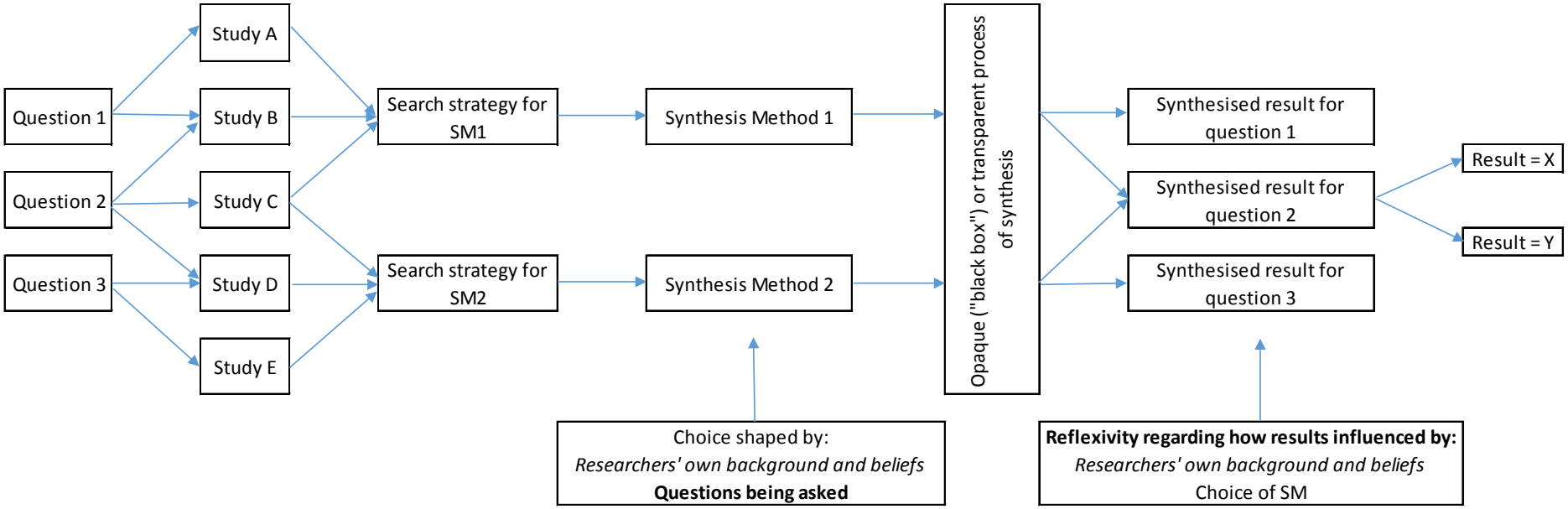
- *Sources missed in searches*
  Extend databases searched and search for grey literature; including hand-searching and citation search of included studies; work with information specialist to ensure search strategy is appropriate and undertake a scoping review to check 'known' evidence sources are identified.

- *Sources lost by strictness of inclusion criteria*:
  Excluding studies based on their design limits generalisability so ensure stakeholders are consulted when determining inclusion/exclusion criteria; determine final criteria following scoping review but before full review, but be mindful of the effect of these changes on potential impact.

- *Sources inadvertently screened out or data incorrectly extracted*:
  Undertake double-screening and data extraction; discuss screening and extraction methods comprehensively across the team to ensure consensus amongst reviewers.

- *Data not reported*:
  Attempt to contact study authors where possible. Whilst they are not obliged to respond, they may welcome interest in their work – especially if it is to be included in a systematic review.

Figure 5: The influence of alternate synthesis methods on results and potential impact of a systematic review



The implications of selecting different synthesis methods (SM) in determining what findings and recommendations are produced in answering the same question:

- SM1 allows questions 1 and 2 to be addressed; SM2 allows questions 2 and 3 to be addressed.
- SM1 identifies studies B & C to address question 2; SM2 identifies studies C & D to address question 2.
- SM1 produces result X for question 2; SM2 produces result Y for question 2. These results may lead to differing recommendations for practice

Not shown, are the additional methodological contingencies that even if both synthesis methods identified the same studies, the differing processes of synthesis could lead to different emphases in the findings and therefore recommendations made.

Figure 6: Strategies to enhance meaningfulness when reporting the review

- *Demonstrate how recommendations have been or could be implemented.*
  Use examples from reviewed studies to describe how the issue has been implemented in other settings. Direct readers to specific studies that provide rich descriptions of implementation.

- *Seek contributions from anticipated users.*
  Engage with groups of practitioners and policymakers to ensure that conclusions also provide advice, tools or guidance on how recommendations can be implemented. If this is not possible, relate the review's findings to the literature on implementation or guidance that complements the review findings.

- *Propose specific future research.*
  Resist using "limited evidence exists" or "more research is required" as the primary finding of the review. If the inclusion criteria are for your selected review design restricted the studies selected for detailed review to a small quota, then identify whether a future review using an alternate design may offer more useful review findings. Identify areas where future studies would enhance the evidence base and propose specific primary and secondary studies and research questions that could meet the needs of the substantive field.

- *Resist implicating or blaming the literature.*
  Avoid pejorative labelling of the existing evidence base. It is likely that researchers did not have your future review in mind when conducting their research. Make recommendations that indicate how researchers and journal editors can report published studies in a way that suits specific review methodologies.

- *Seek critical feedback from anticipated users.*
  Engage with groups of practitioners and policymakers to gain critical feedback on the report. Listen to feedback and adjust the report to ensure its findings are practically-oriented and accessible. If possible, tailor recommendations that are context-specific and relevant. Be mindful of not burying potentially useful findings, instead clearly signpost them.

Figure 7: Impact strategy for a systematic review of interprofessional education (IPE) targeted at educators

IPE educators and curriculum leads in Faculties of Health and Social Care

Educators seek guidance on the most effective model of IPE to employ in their undergraduate curriculum. The review will identify the most effective models of IPE currently available.

A knowledge exchange outreach event will be held one month after the completion of the report. Delegates for the event will be recruited through organisations committed to IPE, such as the Centre for the Advancement of Interprofessional Education (CAIPE) who have agreed to host the event, and established networks of Higher Education Institutions and Clinical Practice Placement Providers.

The event will target educators and curriculum leads involved in pre-qualification health professions programmes. It will raise awareness of the most effective models of IPE and discuss some of the barriers and facilitators to implementing these models in the contexts of educators' host institutions. The event will contextualise recommendations within current debates and recent legislation related to integrated care and proposals for change in the training and workforce development. Delegates will interactively explore how effective models of IPE can be implemented to meet these developments.

Implementation case studies developed through this knowledge exchange will subsequently be reported and disseminated as appendices to the review. The review and case studies will be further disseminated through targeted mailing to curriculum leads, publicity on social media, submission to the National Institute of Health Research (NIHR) Dissemination Centre and Database of Abstracts of Reviews of Effects (DARE), and open-access publication online.