**Webisode 1 – Dimensions of difference in systematic reviews**

Presenter: David Gough (EPPI-Centre, UCL)

EPPI-Centre Evidence Tools, Products, and Projects – A series of webisodes from the Evidence for Policy and Practice Information and Co-ordinating (EPPI) Centre. Hosted by AIR’s Center on Knowledge Translation for Disability and Rehabilitation Research (KTDRR).

**YouTube video:** <https://youtu.be/uvmCxsM16C8>

JOANN STARKS: Welcome to another session from EPPI-Centre Evidence Tools, Products, and Projects. This series of brief webisodes will introduce the audience to several tools, products, and projects of the Evidence for Policy and Practice Information and Coordinating Center or EPPI-Centre. Based at University College London's Institute of Education, the EPPI-Centre focuses on the development of systematic reviews and studies the use of research evidence. I'm Joann Starks from the Center on Knowledge Translation for Disability and Rehabilitation Research, or KTDRR, at American Institutes for Research.

The Center on KTDRR is sponsoring these webisodes with support received from the National Institute on Disability, Independent Living, and Rehabilitation Research, NIDILRR, in the US Department of Health and Human Services. This session focuses on dimensions of difference in systematic reviews. Our presenter is David Gough, who is the director of the EPPI-Centre, and is also professor of evidence informed policy in practice. David's interests include the development of systematic review and the study of research use. Welcome, David. I will now hand things over to you.

DAVID GOUGH: Thank you, Joann. So, in this presentation I'm going to talk about how systematic reviews vary. What my colleagues and I call dimensions of difference in systematic reviews. So, first of all, what are systematic reviews? Well, they are reviews of research, and research is systematic critical inquiry. People do primary research out in the real world, and research reviews are a meta level of research where you bring together the findings of already completed studies.

Instead of addressing research questions by studying the world directly, reviews use the findings of existing studies to address research questions. As reviews are a form of research, they should be systematic and explicit about methods. Really, we shouldn't really need to use the word "systematic." The problem is that traditionally, reviews of literature weren't very rigorous or systematic, and so we now use the word systematic to be clear that these are research exercises. They are research processes for bringing together research findings.

Now, in primary research, there's extreme variation in the type of research questions asked and the methods used. So you would expect, and it's true, that systematic reviews would reflect this variation. If you ask a question at the primary level, then you can ask the same question at the secondary level. You can ask a systematic review question that reflects the primary research.

So, if you ask a particular type of question, then the review is likely to include primary research studies asking the same question. But in addition, the methodology used in the reviews are likely to have similar methodological approaches and concerns as the primary studies that they include. So for example, if you are asking a question about the effect of an intervention, you will be concerned about bias leading you to come to the wrong conclusion. So methodology would protect against bias. If you are concerned about understanding a process or phenomena, then you'd be concerned more about interpretation and representation.

So, the method of the review is likely to reflect not only the type of studies included, but the philosophy of research method used. Now research, people often make a distinction between quantitative and qualitative research. And it's very easy for people to understand what that means, but we think that-- my colleagues and I think that-- the distinction falls down If you look at it closely. So we prefer to make the distinction between aggregative research and configuration research. And aggregative research is where you add all things. You aggregate things. So in primary research, you try to aggregate things. And in systematic reviews of aggregative research, you aggregate the aggregations, as it were. It's all about trying to count things up to understand them.

Whereas configuring primary research and configuring reviews, those approaches to analysis are looking for patterns. So, if we look at the different research questions and methods that you might ask, I've already referred to the question of impact effectiveness, what works. And normally that is using statistics and counting up things. So that is predominantly a configurative-- sorry, not configurative-- that is predominantly an aggregative approach to research. However, there can be configurative aspects.

So when you do a systematic review on a what works question and you find the primary studies, and you look at their effect sizes and you combine them, and come up with a combined aggregation or combined effect size, that is an aggregative exercise in aggregative analysis But if you then go in and you question, look for correlations in the studies, look for patterns to develop hypotheses about what may have what sub-aspects of the data might reveal, that is a configuring exercise. You're looking for patterns.

If your question is about causal processes and mechanisms, that may again be a configurative exercise. Because the whole approach is about trying to understand how things relate to each other. But, in doing a review, a primary study or review on causal processes and mechanisms, there may be some aspects which are adding up data and some research looking for patterns. You may have a mixed methods review-- and in a later slide I will give you an example of that-- where you're combining things.

So again, it's not quantum quall. It's about different approaches to analysis, whether you're aggregating data and whether you're configuring data, and the degrees to which you're doing that. And then the final example on the slide is prevalence. And prevalence is counting things up, so that is very clearly an aggregative exercise. So we just need to be aware that the forms of analysis can be aggregative, configurative, or a balance between these.

So, if we have that in our mind, we can think about a range of ways in which reviews may vary. So, the first thing is the review question. We've all researched the question, is the rock the starting point that then drives everything that you do? And you may have open questions. So that's where you are taking more of an exploratory approach. You've not got a very closed view, you haven't got a very strong a priori-- narrowly defined, specifically defined question-- you're much more open and exploratory.

So, that is on the left hand side of this graph. And an example of that would be a primary research or a review which is trying to make conceptual claims, developing concepts, developing ideas, very configurative. On the right hand side all more closed questions. So statistical match analysis asking questions like what works is testing a very specific hypothesis. I have the hypothesis that this intervention will have a different outcome to a controlled or comparison intervention. It's hypothesis testing. So it's a very closed question.

In the left hand side, those open questions are about developing emergent concepts rather than the right-hand side which pre-specified. The open emergent approach will have methodology, but there'll be more iteration. There'll be less formalized procedures, whereas the closed pre-specified a priori question are more likely to have tightly controlled formal methodology. I must emphasize, this is not suggesting that the emergent approaches don't have methodology, methodological rules, principles, and processes. It's that the extent that they are specified and allow iteration.

And then again, that leads to different types of inferences. Theoretical inferences, or statistical inferences. And in terms of the results, and how they're used, that can also lead to differences with the more open emergent ones developing theories and concepts. It's about enlightenment. Carol Weiss devoted this idea that research is probably having more effect by changing the way we understand the world, and enlightenment effect, that the instrumental effect of facts and information providing us with more information.

So what's really important in looking at these dimensions the difference is that, in any particular bit of primary research, or a review, quite often there'll be balances between these things. Or mixtures of the things. And I've already given an example of post-hoc regression analysis to explore and develop ideas. What works. Statistical meta analysis. That's just one example of how things can be mixed.

But just to give a more concrete view, this is what's called a forest plot. So this is used in statistical meta analysis to test what were his questions. And this example-- and it's a made-up example-- shows five studies. And although studies vary in their effects life. And they also vary in how much confidence-- the confidence intervals-- that they have. So in the graph, the squares, the black squares, is showing you where the effect size is. Against the line along the bottom. Of the x-axis. There's one, two, to three.

And the size of it gives some indication of the sample size. And how big of a result it was. And then the lines are the confidence interval. So the studies with larger samples will tend to have more confidence in the results. And therefore, the confidence controls will be narrower.

But the main point about this is that the bottom, the diamond-shaped square with no filling-- not black-- shows the combination, the combined effect sizes in a statical meta analysis. So if you just followed the results of one study, you probably wouldn't get the correct result. Because you haven't taken into account the other research evidence on the same question. So this is a clear example of aggregation.

The next example is a mixed methods review, where there was a question about barriers and facilitators to fruit and vegetable intake amongst young children. So this was concerned about trying to increase the fruit and vegetables that young children and young people ate. So that was two sub-questions. And they took different approaches to systematic reviews.

So first of all, there was a broad mapping exercise. And there's another episode about what mapping is. But basically, it's describing a research field. And in this review, there was a very broad question, including very many different studies. And those studies could be used to both answer the sub-question about trials, And the sub-question about views.

Now the sub-question about trials with asking whether public health interventions that were used to try and encourage drawing young people to eat more fruits and vegetables were effective. And then, that was compared with a study on young people's views and understanding about health and eating. So the trial's systematic review, the what works systematic review, was predominantly one of aggregative synthesis. And the view study was one of mostly configuring synthesis. Trying to look at what children and young people understood about fruit and vegetables. And then trying to see patterns to create a thematic synthesis.

Third example is a more complex question. And this question was asking about whether legislation to stop people smoking cigarettes in cars where children were present was likely be effective. And so that question was unpacked. So it was seen as like a hypothesis, a theory of, why would you expect banning smoking in cars to have a positive effect?

So here-- these lists are wonderful-- are some of the sub-questions that you would answer. And all of those are often very different questions with different types of evidence. Toxicology, survey research, political science, evaluation of policing. So in this type of real-- this is an example of a realist synthesis. The first part of the review process was unpacking what the question meant, with all these different sub-questions. And then, each of those sub-questions were answered separately.

So in example two, there were just two sub-areas. But in this example, there are very many more sub-areas. And each of them will have different balances of aggregative research-- so answering the sub-question-- and configuring research in order to understand the question. But predominantly, unpacking the review question is a configurative exercise. And then going to find evidence in support of those different sub-components is predominately an aggregative exercise.

This shows how reviews are not just one thing. You have different questions. You need different methods. The questions may be simple, or they may be complex. The method may be single, or there may be multi-component, mixed methods. And this slide explains some of the further dimensions by which reviews may vary.

So we really covered the questions. Conceptual framework, studies considered, and single- or multi-component reviews. But reviews will also vary in how much they try to address. Do they address a narrow question, or a broad question? And to what detail do they examine those questions? So those multi-component reviews in the previous few slides, of course, are quite broad. But to what depth are they examining those issues?

And that will depend on the resources available. Time, staffing costs, and things like that. I've already mentioned that these can be balancing in the extended aggregation and configuration. And therefore, also, they're going to vary in the extent that they're going to be using existing theory, developing new theory, and exploring empirical data.

So the work done by review is not only determined by the question. But how it's addressed, how much resources there are. And not only can you have a systematic review, you can have a review of reviews, which is at a higher level of analysis. So the idea that review is one thing is not really true. It can be many things.

Some people talk about rapid evidence assessments, or rapid reviews. Now, as a systematic reviewer, I get anxious about whether it's systematic rapid review. But if it's rapid, then you are reducing the time and resources available. And therefore, what you can achieve is likely to be less. In a way, something has to give. Now that may be that the question is very narrow. Or it may be that some of the methodological rigor will become a little weaker and shorter, I think.

But you've got set amount of resource. Then that will affect the choices that you make. You can do something broadly, and channelly, narrowly, and in-depth. If you have lots of resources, you can do all of these things.

So finally. I want to say something about review brands. Because in literature on systematic reviews, there are lots of names describing different types of reviews. And these are very, very helpful as a shortcut label to understand the approach being undertaken by review. But I would urge caution, because within these types of reviews, these brands of reviews, there may be quite a lot of variation. Or there may be lack of specification about how the reviews vary on all of these dimensions that I've briefly discussed in this presentation.

So brand is hopeful, but it's a quick label. It's not telling you the specifics. So if you're undertaking your review, or your reading and trying to understand a review, you really need to go beyond the brand and look at all these different dimensions of difference.

So if you're interested in this, then this approach to considering reviews throughout a book, a kind of textbook on systematic reviews that my colleagues and I at the EPPI-Centre have produced. And that will explain this whole philosophy in more detail. Thank you very much.

JOANN STARKS: Thank you very much, David, for sharing the EPPI-Centre perspective on dimensions of difference in systematic reviews. We also want to thank our funding agency, NIDILRR, for supporting this and other webcast activities. Please look for the other sessions in the series on the EPPI-Centre evidence tools, products, and projects.