**KTDRR and Campbell Collaboration Research Evidence Training:**

**Evidence Gap maps and Other Innovative Reporting Strategies**

*Presenter:*

*Howard White, PhD*

A webcast aired on May 15, 2019, sponsored by AIR’s Center on

Knowledge Translation for Disability and Rehabilitation Research (KTDRR)

<https://ktdrr.org/training/webcasts/webcast68/index.html>

Transcript for audio/video file on YouTube:

<https://youtu.be/gpFjTx5zBqI>

JOANN STARKS: Hello, and welcome to today's webcast brought to you by the Center on Knowledge Translation for Disability and Rehabilitation Research, or KTDRR at American Institutes for Research, or A-I-R, in coordination with the Campbell Collaboration. The Center on KTDRR is funded by the National Institute on Disability, Independent Living, and Rehabilitation Research, known as NIDILRR, in the US Department of Health and Human Services, Administration for Community Living. The Campbell Collaboration is an international organization that promotes positive change through the production and use of systematic reviews and other evidence synthesis for evidence-based policy and practice. The Center on KTDRR partners with Campbell's Disability Coordinating Group, or DCG, to help increase the number of Campbell reviews in the disability field.

I'm Joann Starks from the Austin, Texas office of AIR, and I will be the moderator today. I also want to thank my colleagues Shoshana Rabinovsky and Ariana Hammersmith who are helping with the logistics. The KTDRR Center and the Campbell Collaboration are working together to offer a five-part training course that focuses on high quality methods for synthesis of evidence, including the procedures and methods for conducting systematic reviews, as well as software, tools, and strategies for analyzing and reporting data. We kicked off the series in February with Dr. Terri Pigott providing an overview of systematic review and research synthesis, followed in March by Dr. Pigott presenting basic steps and procedures for a Campbell systematic review. Last month, the third session on management analysis tools for reviews introduced four different software packages to assist teams in keeping track of studies and data for their systematic reviews.

Today we are privileged to have Dr. Howard White on hand to discuss strategies other than systematic reviews for reporting data, primarily evidence gap maps or EGMs. Howard is leading efforts with the Campbell Collaboration to develop EGMs with and for different agencies. Today, he will talk about evidence gap maps and provide a number of examples on the use and usefulness of this strategy.

Howard is Campbell's chief executive officer. He was the founding executive director of the International Initiative for Impact Evaluation, or 3ie, and before that led the impact evaluation program of the World Bank's independent evaluation group. He started his career as an academic researcher at the Institute of Social Studies in the Hague and the Institute of Development Studies, University of Sussex. Howard, are you ready to begin?

HOWARD WHITE: Yes, thank you Joann. So thank you, Joann, for the introduction. And as Joann mentioned, I'm going to be talking primarily about evidence and gap maps and what we call other innovative reporting strategies at the end of the presentation, which are various other ways in which evidence can be presented. So first of all, we will talk about what happens in gap maps and then provide a short overview of the history of mapping. And the bulk of the presentation will focus on different types of maps which are addressing different types of research evidence.

So to begin, what is a map? An evidence and gap map is a systematic presentation of all available relevant evidence for a particular sector or subsector. And the way that maps are presented these days-- and we're going to be seeing many examples during this presentation-- are a tabular or matrix format in which different interventions are in the rows of the matrix and different outcomes are in the columns of the matrix. And the cells of the matrix present the available evidence for that particular combination of interventions and outcomes. As I said, we'll see several examples of that in the second half of the presentation.

So, maps live in the evidence synthesis space shown in this particular diagram. What we're mostly familiar with are systematic reviews up in the top left hand corner of this diagram. Systematic reviews of the sort that were presented by Terri Pigott in earlier lectures in the series, present the summary of what the evidence says based on a summary review of what we call the primary studies. Those are the relevant studies to address the research question being assessed.

There are of other types of evidence synthesis products we can produce. The next along the chain is review of reviews, where rather than reviewing primary studies the authors review existing reviews. So you could pick an area in which there are many reviews already. In an international development context that would apply, for example, to conditional cash transfers. There are many reviews looking at different outcomes of those interventions in a domestic context. There are many reviews of child protection policies. And so you might look at doing a review of reviews there. The unit of analysis in those reviews of reviews are the systematic reviews rather than the primary studies.

Next along the chain, evidence and gap maps that we're going to be talking about today, and evidence and gap maps present what evidence is available in systematic reviews-- that's SR, systematic reviews-- and primary studies. Next along the line is a mega map, which we're going to meet later on, which only shows other reviews, reviews of other maps. And finally, a map of maps we'll also be meeting later on. It only shows other maps.

And as the diagram shows, there's a tradeoff between what we call the scope of the evidence synthesis product and the content, or the depth of content. So as you go across the horizontal axis, the scope of the evidence synthesis product is getting broader. So for example, the only map of maps we have to date helps the whole of international development. It's a very broad scope. The mega map helps all of child welfare in low- and middle-income countries. Again, a very broad scope.

Whereas a systematic review might look at the topic, for example, the role of assisted design devices for people with disabilities in improving their learning outcomes. It's a much more specific, focused question, and as the scope gets narrower in that way, so the depth of analysis, the content, gets deeper. So systematic reviews can really dig down into the primary studies and unpack the causal chain, the theory of change, and live intervention and explore heterogeneity of impact across the studies and so on.

The review of reviews relies on the information reported in reviews, not the primary studies, would generally have a broader scope of possibly covering a broad range of interventions or outcomes or both, but not be able to dig so deeply into the evidence because it relies only on the evidence reported in reviews, not the primary studies themselves. An important break point comes in this diagram as evidence and gap maps. Because evidence and gap maps, EGMs, the mega map, and maps of maps (which are types of maps), only tell you what evidence is available, they do not tell you what the evidence says. And that is a very important point to note, at least insofar as maps that can produce support in the production of the maps are a way of displaying the available evidence, not of summarizing what the evidence says. And there is important theoretical reasons why we believe strongly in that approach, that again, I will refer to as we go on.

So whilst maps are very popular when you speak to commissioners of research, funders of research, funders of programs -- I mention maps, they get very excited and want to have one. You have to really be very clear to them that this will only tell you whatever is available, they do not summarize what that evidence says. As I will go on to say later on, it's a basis for then exploring what the evidence says. It helps you make decisions about where to synthesize the evidence, but the map itself does not present that synthesis of what the evidence says.

In addition to the overall framework of the map, the rows and columns presenting a matrix, there are additional filters, what we call filters, which might be country, might be region, might be study design, might be population subgroups like youth, women, people with disabilities, people with humanitarian contexts. There's all different common filters that would apply to a map that you use. You can then select the filter and filter the evidence to only show that evidence corresponding to that particular filter or subpopulation, such as a region, say, only evidence from North America. If the map has any synthesis of global or cross-border regions, and you wanted the evidence from a particular country or particular region, it will, of course, allow you to do so.

Again, emphasizing that the maps only show evidence there, not what the evidence says, and finally, the evidence might be global or particular regions, so we do have in Campbell both global maps, but also maps that color particular regions, specifically, low- and middle-income countries. The map might show primary studies and reviews, it might show just one of those. Most of them would show a good bit of both, and most particularly, the maps can cover different types of evidence.

To date, most of the maps we have, just like most of the reviews we have, are effectiveness reviews, so there are studies that look at the effects of those interventions, what effect those interventions have on the outcomes of interest. I mentioned, for example, the role of assistive devices in people's abilities-- to impact on their attending education and the learning outcomes they achieve. That’s an effectiveness intervention, but you could also carry out maps and other synthesis for any research question.

It's simply a method to be applied to any research question, and we're going to see in this presentation, the thing I chose to focus on is the way in which we've been innovating in the types of maps we've been doing to cover a broader range of research questions, and therefore, different types of evidence to effectiveness studies. We're going to see examples of both effectiveness maps, but also maps covering other types of research design.

Before moving on to that, it should be very clear that evidence mapping is nothing new. In our review of looking at the history of evidence mapping, we have found the first evidence mapping using that terminology going back to the early 2000s, but that was a descriptive overview of the evidence. Similarly, in the same year, EPPI-Centre, a systematic review group based in London, started using the term systematic maps, systematic mapping, to map the evidence that's available. But at this point, they were presenting maps not in pictorial form, but a descriptive overview of the evidence with graphs showing the breakdown of values both by country, by outcome, and so on.

They weren't doing a pictorial representation of the maps in the matrix form, which began in 2010 when the International Initiative for Impact Evaluation, 3ie, first started doing maps. The first map 3ie worked on was on health and nutrition impact of agricultural interventions, but since that time, 3ie's now been associated over 20 maps. And many organizations working in international development have worked with 3ie, organizations like USAID, International Rescue Committee-- both in the US, obviously-- but Sightsavers in the UK, an NGO working on people with sight impairments-- as all have gone on and done evidence mapping, again, working with 3ie. That's really been an important turning point, is the work of 3ie promoting evidence mapping. And Campbell has built, in particular, on that approach.

As you see, other organizations have been adopting mapping as we go through time, and if we go to the next slide, we'll see the continuation of mapping. I think quite important to point out, here, is Epistemonikos, which is an online database for health studies, health impact evaluations, and systematic reviews. So they are effective studies, and it does have a mapping utility to map the evidence of particular sectors or questions that you choose to specify with Epistemonikos.

And then importantly--the last one for this presentation--in 2017, so two years ago, now, Campbell developed the guidelines and standards, and so on, for producing evidence and gap maps according to Campbell criteria. And you'll see also mentioned on this slide, as I mentioned already, the work of Sightsavers working on interventions for people with visual impairments in developing countries, and International Rescue Committee, which works mainly in humanitarian contexts, and also their [INAUDIBLE] reference tools including evidence maps, working with 3ie.

So you see, now, in this slide, as I already mentioned, there's been a large increase in interest in mapping in recent years, and this has been international [INAUDIBLE] development, [INAUDIBLE] in quite large part by 3ie's own work, or the work that's been supported by 3ie. This is a graph taken from the map of maps, I've mentioned already. The map of maps shows 79 ongoing or completed maps international development space. This is a study I'm taking two years ago, now, so it's a bit dated. We know of at least another 20 maps have started since then.

And from this graph, you see very clearly the rapid rise in mapping from 2013 through the three years up to 2016. 2017 is not a drop off like that, the study was undertaken in mid-2017, so 2017 maps weren't yet available to be identified to go into this graph. Once we update this study later this year, we would fully expect to see new increases in this graph continued from 17, and indeed, into 2018.

So a large increase in evidence maps in recent years. This is just in international development, but you would look also in other areas, like in social welfare, like child protection in developed countries. Then I'm aware there's many maps being commissioned in those sorts of places, as well. There's been a big increase in the number of maps being commissioned in recent years, and more and more follow the sort of visual presentation I'm going to be showing you in the second half of this talk.

The important point about maps and why they're part of this lecture series is evidence and gap maps are a systematic product following the same principles of evidence synthesis as does a systematic review. Following the procedures that have been explained in the previous talks in this lecture series. And that's why I'm not going to go into detail in explaining how you do the map, because the idea of having PICOS, or having a clear search strategy, systematically related research strategy, reporting the search strategy, systematic coding of studies, systematically reporting of what you find-- all of that's the same for a map as it is for a review. Where there are differences, I will mention those now, but the similarities are much greater than the differences. Evidence and gap maps are a systematic product in the same way systematic reviews are.

So you specify the scope of the map having a PICOS, the population intervention, comparison groups, outcomes, and study designs, just as you would for a systematic review. But most importantly-- and I'll come back to this in a moment-- you define the scope of the map with the framework. The framework defines the intervention and outcomes for the matrix, and defines the filters. These help define your search strategy and your coding form. So it's very important that in the protocol stage, or even the title registration stage you go through in Campbell, that you specify clearly what the framework for the map is.

You have to have a pre-specified protocol for the map, saying what you are going to do, how you're going to do your analysis. This is standard for a Campbell product, and it's necessary for maps. So you have to have a systematic search strategy that says where you're going to search, and how you're going to search.

So a search strategy has two components. Where you're going to search-- which databases, which websites, with journals you're going to search-- and how you're going to search, which are the search strings or the key search terms you're going to use. So just as for a systematic review, the protocol has to specify the search strategy, similarly for the evidence and gap map, the protocol has to specify the search strategy. Once you've done the search, you've got your search results, then you need to have your screening using clear, transparent, ex-ante inclusion-exclusion criteria.

Note that because the scope of the map will be generally broader than the scope of a review, you're going to get many more hits. So in a review, you tend to get 2,000 to 3,000 hits, you might get up to 10,000, wouldn't be unusual for a review. Whereas, for a map, getting 20,000 or more hits is not at all unusual. We're working on a map right now where we started with 70,000 hits. Because the scope is so broad, you'll find a lot more potentially relevant studies that have to be screened.

And so that's one difference. The procedure is the same, but the amount of work involved in the screening is going to be more because you have to go through a lot more studies. Of course, a lot would be done on title and abstract screenings. That goes very quickly, but you do need to go through that procedure. And for a Campbell map, like a Campbell review, has to be double screened-- at least at the start. If you can get consensus of your screeners and coders, at around 85%, and you can report that, then you can do single screening with a random check or a random sample by a supervisor. And that's the procedure we're allowing now, and on maps that have a very large number of studies to be screened.

As I mentioned, screening needs to have very clear inclusion and exclusion criteria, just like for a review, and then you need to code the studies. This is another difference. Because maps only show evidence that's there, not what the evidence says, the coding required for a map is far less per study than the coding required for a review. When you're coding for a review-- particularly an effectiveness review-- you have to code all of the effect sizes, and the separate statistics-- sample size, statistical control groups, standardizing of the outcomes, and so on-- and you have to then code also sub-sample analysis, sub-sample characteristics, and so on. You don't need any of that when you're coding for a map. You just need the basic demographic information, you need the intervention and outcomes, and you need to add any filters you're applying, like study design, country for study, and so on. Basic information you'd expect to collect. What is normally left is simple to code.

So with a coding form for a map-- and it depends on the map-- but it's not unreasonable to expect to be able to code three to four studies an hour. So you could actually manage maybe 20 a day. Whereas, for a review, you often expect it to take two or three days to code a single study, depending on how deep the review is going. But there's a big difference in terms of coding time for study-- and a map could have more studies. The review might only have 10 or 12 included studies, and a map can have 200. So your total coding time might be similar or even more for a map, but the time required per study is far less. Because, when it comes to the last bullet point-- systematically report all eligible studies-- that means putting them into the map and reporting your tables or other evidence, but you don't have to actually synthesize the evidence.

Also, the stage of analysis is also much quicker for a map, because you're doing univariant, bivariant tabulations of what evidence is there, by study design, by country, and so on. You're not doing detailed analysis of content, qualitative synthesis of the values and percentages, and so on. The synthesis component, you could do it in a half-day, or a day in an evidence and gap map-- which you certainly can't do in a systematic review.

So, typically, while a Campbell review might take 12 to 18 months or longer, a Campbell map can be done in 6 to 12 months. The work on a map can be done in less than six months. Going through the Campbell review process usually takes a bit longer, but the time required for the actual production, and it's going to depend on the number of studies, is typically three to six months.

A key component of that is the framework. The framework is the key component of an evidence and gap map. And to repeat, the framework is intervention and outcome categories, and then any additional filter you want to code. But it's getting the intervention and outcome categories right that is certainly key. And that needs to have stakeholder engagement, some sort of a consultative process with stakeholders to get something they will recognize and buy into. If you put something out there where they say, well, I don't recognize where is this intervention, where's that intervention, I can't find it, I don't understand these categories, then you're going to undermine the credibility and use of the product.

So we'll adopt a couple of approaches. One is that we look at existing strategies that are out there in your particular sector, and where there is an existing framework, then we use that. So we have done an evidence and gap map on interventions for people with disabilities in low and middle income countries, and for that, we used the WHO Community-Based Rehabilitation matrix for both intervention categories and the outcome categories. And we did that because the CBR is commonly recognized amongst governments and international agencies working with people with disabilities. And so, by using that framework, we knew we'd get acceptance of what we had done, rather than inventing our own, or take one from a particular agency, we took what had already universal acceptance.

When you haven't got that-- if we're doing a map of access to justice, where there is no such framework and no such consensus-- then you have to try and look at strategies to identify what that framework should look like. And go through stakeholder consultation with intended users to get agreements around that framework. When there's an existing framework we can use, designing the framework takes a little time-- it can take a week, maximum. But where you need to go through consultative process, it can take two or three months. It depends on how long the consultation is, whether you use physical meetings to do it, or whatever.

But again, as is with reviews, Campbell maps are strongly advised to have an advisory board, and that advisory board should comprise both academic experts in the area, but also practitioners and policymakers in the area. And the advisory board are amongst the groups you're using to help advise you on the framework and you need to get them to agree to the framework. There are some caveats, there, I will come back to when we see some actual maps.

Let's move on to examples. The first examples we're going to see are not the Campbell maps. We will start with a 3ie example. But there is a wide variety of mapping formats out there, and it's interesting to take a look at some of those, because you may feel your organization would like something different to what we have, and so it's good for you to see the range of choices which are out there.

So the first example, here, is a 3ie map. And so this is as described earlier, and this example comes from adolescent sexual reproductive health. The interventions are in the rows, and the intervention categories, and then subcategories. And typically, you'd want to have around 20 to 25 subcategories. So five to six, maybe even four intervention categories, and then, within that four to six subcategories for each row, and similarly, five or six outcome categories-- that's in the columns. You see here-- five. And then within those, four to six outcome subcategories. So your dimensions of your matrix should be around 20 to 26 rows, 20 to 26 columns. Any more than that, it gets harder to navigate one single page. And in particular, though, what we find is that once you get more disaggregated than that, it becomes harder to code, as well, because you haven't got such clarity of distinction between the different categories once you subdivide the space too much.

The principles, I'll repeat again later. Designing a framework, you can try and go for a small number of large holes rather than a lot of small holes. But don't try and get too fine with this aggregated matrix. It's going to be less usable in the end. So people driving to it thinking it will be more useful. It won't. It will be less useful.

Now, going back to look at the content of the map, as you look across the top row, which is provider training and youth-friendly service adjustments, and we go across to the health service column, where the first of those is access in utilizing health services. And the top cell, there, in that row has a big red dot and a big grey dot. In the 3ie maps, the grey dots correspond to impact evaluations. And the larger the circle, the more impact evaluations there are. The colored dots correspond to systematic reviews. And they're color-coded by what we call "strength"-- confidence in the evidence. So low, which is red, medium, which is orange, and green, which is high. We also include in the map ongoing reviews, which are coded blue.

If this were a live version of the map, you can hover over a particular circle and get a list of all the studies there. So if there are 16 studies in that circle, it will list all 16 studies-- so for a primary study, it will tell you the author, the year, and the country, and often, there on the map, the study title. And you can, from that list, click on any one of those studies, and click through to the database record for that study-- which summarizes the study and gives a link to the URL. So these maps not only display the evidence in an easy-to-understand, navigable way, but they also actually take you through to find the evidence, find the underlying studies. That's an important function of maps, is putting the studies back into the public domain. And I’ve mentioned this a few times. It makes the evidence discoverable and accessible, and therefore usable.

Because many people would not be able to discover these studies, because they're in databases they don't know how to use or don't go to, and they can't access them, so they can't find how to get them. The access is there, both because of the database summary, which goes beyond the abstract, but also because of the link to the URL. Now, the URL might, then, be behind a paywall, and there's nothing we can do about that. Maybe it just has a description, maybe it doesn't. But the user, depending on where they are, may be able to access the actual research office, as well. If not, maybe you'll get a summary in the database record. So maps play an important role in surfacing evidence that's not [INAUDIBLE], and putting that evidence out there so it can be more discovered, more accessible. And we've seen that in some examples I'll come back to in the second half of the presentation as being a primary function of maps.

Here's the Sightsavers map. It looks a bit similar to the 3ie map, but actually has some very important differences. So all the Sightsavers maps are related to visual disabilities, visual impairments, and this one is for cataracts. And as we go across the top of the row, where the column headings are, you'll see that they have a number of outcomes around burden of disease, biomedical outcomes, service delivery outcomes, and health systems outcomes. And that each circle in this map is a review. This map only includes reviews-- it doesn't include primary studies. And then they're color-coded the exact same way I described before-- red is low confidence, meaning we're not really sure we can trust the review findings, so shortcomings in the review, medium confidence, and high confidence. And you'll see the majority of studies are coded red. That's not unusual. To make this assessment-- it's called critical appraisal. It's something we also do in systematic reviews. We critically appraise the include studies.

Sightsavers used a SURE check method. You can Google this, look online. You'll find quite a large number of checklists for critical appraisal. And actually, the group that produce [SURE](https://www.cardiff.ac.uk/specialist-unit-for-review-evidence)-- there at Cardiff, in the UK-- they have a website that gives a list of critical appraisal tools for different types of research designs, including reviews, but also different types of primary studies. And so, Sightsavers used the SURE checklist to critically appraise the included reviews, and decide whether they were high confidence or low confidence.

The other thing they do in this map is to say whether the evidence supports the intervention-- if there's an effect, or not. So you see, down the left-hand column, it's says "strong," inconclusive” and "weak." So that's where the intervention review concludes there's weak evidence of any effects, inconclusive evidence of the effects, and strong evidence of the effects. Now, in this case, because they're basing it on reviews, I don't object so much to that approach. When we come to the next map I'm going to show you, I do object more strongly. But we're not that keen on that approach, for reasons I will explain in a moment. And you'll see. I'll lead to that. But I do intend to have a look at the Sightsavers map, and to explore it. Any of these maps are all online. You can access them and explore them.

Here is the Evidence-based policing matrix. This is one of the older maps that's around. And this map shows primary studies, not reviews. So every single dot is a primary study. And this has multiple dimensions. There's three dimensions you can see, and the fourth dimension is given by the color coding. That fourth-dimensional color-coding-- and shape, because they have also this triangle-- is to do with the effect. So the upside-down red triangle means the intervention has an adverse effect, whereas the black circle is a positive effect, and white circle null effect, and grey is what they call "mixed findings" from the single study. So that would be positive effect on some outcomes, or some measures, and not other measures. Because in this map, there's only one outcome, and that outcome is the impact on crime rates.

There are three other dimensions and I'll go through very briefly. On the horizontal axis, the x-axis, you have what they call the scope or target going from whether the intervention targeted individuals, groups, what they call "micro-places," which is like a block, a neighborhood, which is a larger space, or then a jurisdiction, which is a larger geographical area still, or nation state. Then they have specificity of prevention mechanism-- so whether it's a focused prevention mechanism focused on that particular group, or whether it's a more general one which tends to affect that group-- and then level of proactivity, which whether it's something the police go up and do, or if it's something they didn't place, and is a reactive intervention to things that happen. And so the interventions are coded by those different categories.

Now, why we object to this-- if you look at the studies down the bottom left-hand corner, for example, you'll find they all correspond to the same intervention. What this design of the map encourages you to do is what we call "vote counting," which Terri (Pigott) may have spoken about in her lecture. Vote counting says, OK, well you see, there are roughly 10 studies there. Half of them—OK, there are 11 studies there. Five of them find no effect at all. Three have mixed findings. Two say positive impact. One says negative impact. So overall, you know, it looks a bit mixed, the evidence. We're not really sure, but maybe it works. On balance, some say it works, some say it doesn't.

That's not the correct way to provide evidence. The correct way to provide evidence is through meta-analysis. Maps like this can lead you to tend to report study effects. This is a map of report study effects, but maps that report study effects that encourage the user to engage in vote counting and at Campbell, we're very, very, very, very opposed to vote counting. We believe in meta-analysis. We think that showing in the map what the finding of the individual studies is, is actually counterproductive. It can lead people down into bad habits of vote counting, which we're strongly opposed to.

So in the Campbell map, like the 3ie map, our coding of the studies represents coded study findings but doesn't tell you what the studies say. But I'm about to argue-- it's not everyone's view-- but the following map I'm going to show you, which is a 5D bubble plot, which are produced by Veteran's Affairs in the United States-- there again show you what the evidence says. We don't mind so much, because it comes from reviews. We mind less when the reported evidence comes from reviews, like with Sightsavers, than single studies-- but still, we'd rather leave that to doing the evidence synthesis yourself.

On the horizontal axis, there's evidence of no effect on the far left, unclear evidence in the middle, and evidence of positive effect on the right-hand side. The bubbles have various sizes, and the key is in the top left-hand corner, there. The larger the bubble, the more reviews there are in that bubble. These bubbles are color-coded into sets of interventions. So you see that there is this blue bubble-- the light-blue bubble is a hypertension. So you see, in the absence of positive effect column on the far right, the top one of those is a blue bubble which says "hypertension" on it-- there's another "hypertension" bubble in the middle section. And so the color-coding corresponds to an intervention class, and then specific intervention categories. It's like our intervention categories and subcategories. Their color coding captures the categories, individual bubbles the different subcategories.

And they have a vertical axis which corresponds to literature size. So it's the number of articles impact of evidence there are in the reviews of that particular intervention subcategory. And whilst they don't synthesize or map the individual studies, they do tell you how big is its literature, which is quite a useful thing to know. I think that's what I like about a 5D bubble plot, is that it shows you how big the literature is. So you can see that the health outcome in this particular map, there's a large number of studies, whereas for most of them, they're lying on the bottom, there, there's just zero, one study. So not a lot of evidence actually to synthesize. The reviews don't have many included studies. So there's a whole different number of ways in which these maps can be presented.

You can read about those in our discussion papers, you'll find online on the Campbell website. We go through all these different agencies and discuss their approaches, how they've done these maps, and basically, what we think of the pros and cons, strengths and weaknesses of each approach-- which we looked at to aid others and to inform the Campbell standards and approach doing these maps.

Now I'm going to go ahead to evidence mapping initiatives. These are different maps we're in doing in Campbell which show the scope of maps you can do, the sort of research questions you can approach, and some of the innovations I think we'll be undertaking in order to push the frontier of what's being done by this mapping. So we're going to look at the map of maps, the mega map, and so on. We're going to go through the examples right now.

The first one we've mentioned already is the map of maps, the Effectiveness of International Development Interventions. This is a very broad map that has an extra-extra large scope. It's all of international development, and because its scope is so large, we've only included other evidence and gap maps. We don't include reviews, we don't include primary studies, there have been thousands of them. We only include other evidence and gap maps.

And so it looks like this, this is the version published a couple of years ago. It had-- so there are actually 73 maps, of which 55 were completed and 18 ongoing. Just as in the other maps we've seen, there are different outcome intervention categories in the rows—agricultural development, climate change environment, conflict management, and so on-- and then different outcomes, which are sustainable development goals in the columns. You can hover over any one of these circles. The circle size corresponds to the number of maps and you can click on the circle to access the actual maps themselves, description of the maps, and links to URLs for the maps.

With this map, what we find is that the majority of internet maps to date are in health. Any maps you do, you'll find there are many more studies in health than there are in any other area. We identified under-represented areas. This map was produced by 3ie, but it's part of its work with what's called CEDIL. [CEDIL](https://cedilprogramme.org/) is Centre of Excellence for Development Impact and Learning, CEDIL-- you can Google it and find it-- which is a consortium involving 3ie and Campbell, amongst others. 3ie made the map of maps, and published by CEDIL. The UK Department for International Development, DFID, is financing CEDIL.

It's a useful map to identify three areas, of which the disability map was one, I've already mentioned that. Access to justice, I've already mentioned was another one, and transport was a third. We're doing maps in those three areas based on the fact that the map of maps shows these are under-mapped areas, and so we want to do a map now to look into more detail why there's no maps there. I don’t say there are reviews there, but there are primary studies there. And just a mention of interest, this particular investigation found very few studies-- fewer than a hundred studies-- on disability. It's a very small number, particularly in Africa. Most of the studies we found were in China and Iran and Turkey, a reasonable number in India, but not many studies at all for large parts of the world. So the map of maps is one innovation we've done by mapping, not reviews, not primary studies, but mapping maps.

The second innovation-- actually it was the first, because it was the first one we did-- is a mega map on child welfare. Again, I've mentioned this, it's an evidence and gap map with a large scope. The mega map had an-- sorry, the map of maps had a very broad scope. This has a large scope-- all of child welfare; health, nutrition, education, child protection, and so on. It doesn't include primary studies, because, again, there would be thousands of them-- particularly in health. It only includes other maps and systematic reviews. So it looks like any other map; here it is. There's only colored circles, because it's showing only the reviews and no primary studies, and we've done a critical appraisal of the included studies.

And so you have red are low confidence—sorry, gone back. Red is low confidence, green is high confidence, and orange is medium confidence. You see quite a lot of studies here, there are over 300 reviews. This actually is the first version we published. We updated it recently with a grey literature search and there are now around 350 reviews in the map. And what you'll see here-- I think I've got the slide. So the majority of studies, about 2/3 of the studies are in health and nearly most of the rest are in early child development and education. So a study might be counted as being more than one of these areas. You'll see actually nearly every map covers health, the health sector amongst its outcomes. But the others are-- a lot of others cover ECD and education and then areas like social protection, governance, social work and welfare, and these include child rights, child protection, trafficking, and so on. There's also, actually, some disability-related issues. They have far less coverage, far fewer reviews in those areas.

Because this is a mega map, it only contains maps and reviews. We can't say there's evidence in these areas. We can say there's no evidence synthesis in these areas. There's an absence of evidence synthesis in these areas; whether there's no evidence, we'd need to do a map to find out. The use of this map has been precisely to say, well we can see that in these areas around what we call the new areas of child rights-- around child abuse, trafficking, and so on-- there's very little synthesis. UNICEF have commissioned Campbell to do a violence against children evidence and gap map to see is it really true there are no studies in this area, and we have some preliminary findings. There were some studies, but again, it's another map where there's fewer than a hundred. When you're doing something as broad as a map, and violence against children includes both abuse and neglect, so that's fairly broad scope, covered by a fairly broad range of interventions. When you find so few studies, that's a fairly sparsely-populated map, so we are finding from that map there's some areas without much evidence. It's not true of all maps. So the mega map, again, is an innovation because we've looked at broad scope, looking just at reviews in the map, not primary studies.

The next example is the homelessness effectiveness map. So this is a more standard map, where we have looked at homelessness primarily in developed countries. We're doing this for the What Works in homelessness in the UK, the Centre for Homelessness Impact. It's looking at effectiveness studies, impact evaluations, and reviews and the innovation was the critical appraisal of all included studies. There are one or two maps that have done critical appraisal reviews, but haven't done critical appraisal of all the primary studies, because there's a lot of them. Doing critical appraisal of primary studies is just a lot of work, but we undertook to do it and so we did that.

In the map, here, you can see, though not quite so well, the red dots you can see from the key at the bottom is low quality. We use the word "quality," because it fitted but we should say, low confidence study findings of primary studies is red; medium or high quality, which are not so many, is green for primary studies. Orange is low quality reviews and a different shade of green is medium to high quality reviews. If you look at the numbers at the top, you'll see there are 238 impact evaluations, and only 22 reviews.

This is a very under-reviewed area. If we do a map in a health-related area like disability, there are as many reviews as there are primary studies. If that's going a bit too far, in areas where there's a lot of public health research, which is true of disability, you'll find a lot of reviews because health researchers do reviews, and the number of reviews will be equal to or more than the number of included studies.

Here in homelessness, there's not a tradition of doing reviews. It’s outside of the health sector and health interventions in the sector. Outside the health interventions, you don't see a lot of reviews being done. It's an under-reviewed area, which is one of the important messages on this particular map. But the innovation here was doing critical appraisal of all included studies. And so, what's been done with this map is to identify the areas where there are studies where we could do a review, and already the Centre's commissioned three Campbell reviews to use the evidence identified in the maps to do study reviews in those particular areas around increasing access to health for people who experience homelessness, for example.

Moving on to the next innovation is, again, on homelessness, but now what we've done is to map the process evaluations, not the impact evaluations. So the scope is the same, it's homelessness, but we're looking at process evaluations, implementation studies. They are very different and they come from a very different sort of background. In the homelessness map, you saw there were a total of 238 included studies. They're nearly all in North America, they're nearly all from the US and Canada. There's just 12 from the UK. Now, we've done that map for a UK organization who knew that was going to be the case, so we need to sell that map to a domestic audience. Including academic research in the UK working on homelessness, none of the studies are included in the map that don't do impact evaluations.

So one way we've addressed that is to do a map of process evaluations, nearly a quarter of which do come from the UK because we're only doing it in English language. So again, most are from the US and nearly a quarter are from the UK. The important innovation, here, really, was that these are nearly all grey literature. These process valuations are not even in Google Scholar, which is normally quite good at finding grey literature. They are certainly not on OpenGrey, or anything like that, where you would normally try and search for grey literature.

So the way we found them was by going through every single state in the US, every single county in the UK, every single state or province in Australia and Canada, going to the relevant website of the relevant agency and searching on that website and going over the publication lists of that agency on their website, and then going to the main research agencies like AIR, like Abt Associates, Mathematica, and so on, going to their websites and searching their websites, also, for studies they've been involved in. And similarly research engines in the UK, then NGOs active in this sector, and searching their websites.

And so it's very much a website-based search strategy, not a database search strategy. So when you're doing an effectiveness study, most of the evaluations will come from the database searches, just supplemented with website searches. In this case, nearly all the literature in the map-- I mean, I had 220 impact evaluations which we did-- and I would say nearly 200 of those came from the website searches.

In the map itself, the columns are no longer outcomes. The columns are barriers and facilitators. So we've developed a framework of barriers and facilitators from the implementation science framework. We ran that by, first of all, according to the research that we worked with in the social sector, but then, through that, the UK academics on homelessness, to get a coding framework. Actually, we have worked with a group in Edinburgh working on homelessness, to do the coding of these process evaluations, to look at the barriers and facilitators of interventions for people experiencing homelessness.

And so, whilst we say that maps don't tell what the evidence says, that's not really, entirely true of the process evaluations maps, because a process evaluation map does tell you what are identified as being the main barriers and facilitators to successful implementation of interventions for people experiencing homelessness.

And so the gaps, the empty cells in this map, do[es] represent there being evidence in this map, which they do in a traditional effectiveness map. This has not been identified as a barrier and facilitator. We have to get that correct. We can't say it's not a barrier and facilitator, we have to say it's not been identified as one. So it may have been one, but not been identified in the process of barriers and facilitators to look at it. We would have to be careful on what we say and we don't say.

So this map-- these two maps-- have been developments in a couple of ways. One is that we are doing the systematic reviews, and they're going to include the process evaluations and mixed-method reviews, so they include the process evaluation evidence, as well. But the second is that they're building an intervention tool, an evidence portal, which I'll show you in the next few minutes, in the concluding section of this talk, when we talk about different ways of presenting evidence. And so evidence summaries are being produced from the process evaluations around the issues identified as barriers and facilitators in the process evaluations map.

Having done the first map of process evaluations on homelessness, the next time we've done that is in the Uganda country evidence and gap map. Uganda is a country where I have quite a strong association for a number of years, and I was there last summer doing a presentation for the Uganda Evaluation Association. And just as a question whilst doing the presentation, I asked people how many randomized control trials do you think have been carried out on development interventions in Uganda? And some people said none, others were saying nine or 10. And that was the highest guess that I got.

Now, 3ie has a database of development impact evaluations. So I checked the 3ie database (I had done for the talk) and there were 92; 92 randomized control trials of development impact evaluations had been carried out. I then looked at the overall impact evaluations, and I found that over 220 impact evaluations were carried out in Uganda. So I thought, what about process evaluations? I did a quick scoping of those, and I think there was over 200 of those, as well. So since 2000, there have been around 500 evaluations of development interventions in Uganda, and the fact is that most people don't know most of those studies are there.

This is an evidence and gap map, but its scope is to include all evaluations of whatever kind-- process evaluations, impact evaluations, and formative evaluations-- carried out as development interventions in the last 18 years. And so we've put that map together, now, this is a snapshot of it. The green circles are the impact evaluations. The blue circles are process evaluations. There is, at the top left-hand cell, there's a small orange dot. That's formative evaluation. There's only seven in the map, so you're lucky we got one in this screenshot. We have not done many formative evaluations.

Doing this map has made me really emphasize the discoverability, accessibility sort of maps. Because one wrong evaluation is lesson learning; the lesson learning, functional evaluations of global public good. But no one's learning lessons from evaluations that no one ever reads or knows about. And the vast majority of these evaluations shown in this map were not known about, not being read, not being used. So by putting them up there in the public domain through this map, they're now accessible.

There's a master's program on monitoring and evaluation at the University of Technology and Management in Kampala, Uganda and they're now using the map with their students. Whenever they've had an assignment they tell them, well, go and find the evaluation in the map, and use that evaluation as the subject of this assignment. They could have done that before, but they'd have all done it on the same one or two evaluations that were commonly known about. Now, all 500 are out there in the public domain.

Again, this is one that's a map done mainly by website searching. Database searching is not the way to go about constructing a map like this. Typically with process evaluations, it's based mainly on website searching. There's a number of locally commissioned studies. Locally commissioned studies by a ministry or local NGOs are harder to find, may not be on websites. And there, we worked with the office of the prime minister of Uganda, our partner when doing the map. They contacted different ministries to collect the evaluations they had commissioned and had over the last 18 years to include in the map. So you have to use different approaches of identifying the evidence for a map of this sort. But the interest and excitement over this map has been very great, and so I would expect to see many more maps of this sort, going ahead.

That's right-- there's a zooming into the map, we can see, for example, on education literacy, or education outcomes, there are 23 studies. And you can see it's just about that for the health and nutrition cell, where there were many more studies.

So what's been done, one other thing we did-- because you can code whatever you like-- so one thing we coded was who the authors were. What we found is that the vast majority of studies actually had no Ugandan authors. You can see from this, over 2,000 studies had no Ugandan authors-- which is quite scandalous. You can pick on any item of interest when you're doing a coding. What's interesting to code? People don't normally code nationality of study author but in a map like this, that is something that was interesting. So you can code it and present it like we do here. I think this slide alone, I think this slide was the most tweeted slide when I made the presentation map back in Uganda a couple of months ago. Because it's something of great concern, of course, to local evaluators-- to see who winds up being bypassed in the production of local evaluation studies.

What's being done with this map is where there's a hodge-podge of evidence, in education and governance, there are quite a lot of studies. So we could do country-level evidence synthesis of those studies to see what we could learn from them. We can see areas where there are evidence gaps, where there simply aren't studies, to go ahead and commission more studies in those areas. But I think really, the next step for us is that many more countries are interested in having these maps, and we expect to see many more of them in the years to come.

In fact, we have, now, 18 ongoing maps. I've presented a few of them. I'll just pick out one more, that I mentioned at the start. Naturally, as mentioned there, it's the pathways between agriculture and nutrition. And actually, it's not even that map. So we have a map being done on methods, metrics, and tools between agriculture and nutrition. So they're not mapping what evidence is available-- they're mapping what metrics and tools and methods are used to analyze agricultural and nutritional linkages. This is being done by a research group who is interested in investing in developing or promoting the use of rigorous tools and metrics for looking at agriculture and nutrition in developing countries. And so they're surveying all the existing tools, to then do reviews of those tools to try and come to a conclusion on recommending which tools should they use. I'm working on this map, and I can only say, it's really, really needed. Because there is a total proliferation of people-- every research team inventing their own tool, and no coordination, no realizing economies of scale, no standardization across tools.

This actually is another part of the disability map. Even though there are these [INAUDIBLE] questions are meant to be used to assess disability. it's not being done, not being applied in a standard way so they need to converge around outcome measures in most areas. Also, they need to converge around the tools and the instruments being used, so you can actually get comparative study findings, so you can actually get comparative effectiveness. The map of methods, metrics, and tools can be very helpful in saying, clearly, we need a review of this particular tool or tools, when it's doing it's particular thing, so we can actually go dig a bit deeper into these tools, and learn about them, and how comparable they are.

I mentioned the violence against children map. They had a workshop a couple of weeks ago on violence against children, and there are a half-dozen different survey instruments to measure violence against children. So again, I said, well, we need not just a map, but we need to do a review of them. And someone present said, oh, but actually, we didn't invent our own. We looked at this tool, and [INAUDIBLE] we used that instrument. No one else in the room knew that, even though they were meant to be experts in the area. Just by mapping them, we don't see, actually, these different instruments overlap. We need to go on and do the review.

So this brings me very nicely to my last point, which is that evidence maps are an important building block in the evidence architecture. But they are that, they're a building block. Evidence maps are a means to an end and they're not the end in themselves. Evidence maps can be used to commission reviews, and we've seen that used several times, identifying it leads to commission. In particular, they can help build evidence portals, and to write guidelines and checklists.

So this pyramid you see in front of you now is what I call the knowledge brokering pyramid. And it goes through different stages of knowledge brokering. As you go up the pyramid, knowledge is being more heavily brokered or translated. It's not the idea that the top row of the pyramid is the best-- it's not an evidence pyramid in that way-- it’s the most heavily brokered.

So at the bottom level, we have data. And data is used, collected—those data are analyzed or translated into primary studies. Those primary studies are summarized in reviews, and reviews and studies go to databases, and subsections of those databases can be put into maps, or script platforms. Then you have portals, guidelines, and checklists. The important point is that maps take the user back to the original studies. Portals, guidelines, and checklists enable a decision maker to make evidence-based decisions without having to go back and look at the original research.

The best example of this is produced by the What Works [Network] in the United Kingdom. The Education Endowment Foundation has a Teaching and Learning Toolkit to inform school level decisions on interventions to adopt to improve learning outcomes. It has now identified 34 interventions, and for each of those 34 interventions, they commissioned systematic reviews. They put this tool together that shows you the name of the intervention; what it costs, in the range of one to five pound symbols; how strong the evidence is; then the impact. The impact is the amount of additional learning outcomes a child gets from being exposed to this intervention. You can sort by that, so you can see that giving a child feedback is very low cost, reasonable evidence, and it's like getting an additional eight months of learning. So including feedback in your curriculum is a very powerful way of including them in the outcome.

By contrast, repeating a year is very expensive, and actually sets a child back by four months compared to what they would have been had they not been sent back a year. It actually had adverse effects on the child. Now, this is a very good goal. The decision makers might decide what interventions are for the school. A report by the UK National Audit Office (similar to the MBO in the US) carried out in 2015, found that nearly 2/3 of schools are using this Toolkit to decide how to allocate school resources.

That means that 2/3 of schools in the UK are making evidence-based decisions based on systematic review evidence because they’ve got access to this tool, not because they really researched it all. Because they don't need to, all they need to know is here. They can click down and read more about the intervention. They can click down and read the research reports. But they don't need to, this is all they need to see there. So this is one way of supporting decision making-- here's the evidence, you decide what to do.

Another way is using-- oh, sorry. This is the Homelessness Intervention Tool. So I meant earlier What Works and the Centre on Homelessness Impact do the same thing. And this is the prototype intervention tool that they have out there. So again, it's an example of an intervention tool. Here's what the evidence says, please go ahead and decide what to do.

Another way of doing is using guidelines. The next level up on my pyramid is guidelines, evidence-based guidelines. This has been done by WHO for many years, now, and this is a quote from the WHO guidelines on producing guidelines; WHO guidelines on producing guidelines. WHO guidelines have to be based on high quality systematic reviews, so what we're doing, here, is institutionalizing this evidence from systematic reviews. That's what we'd like to see happen in all sectors based on evidence portals, like the Teaching and learning Toolkit, or evidence-based guidelines, or checklists.

Checklists are one level up on the pyramid. Guidelines say, we've looked at the evidence, recommend you do this. Checklists say, just do this. So checklist is based on the evidence. Here are the things you need to think about when you're designing a fall prevention program. So checklists are often ways to help you think about designing a program or intervention, often means much more than practice the program decisions.

As you go up the evidence brokering pyramid, you go through different levels of decision making. Portals help decide what to do, and guidelines can do both what and how, and checklists are much more about how to do things. So their use is different, and they have different degrees of agency for the user. They're all ways of presenting evidence, and presenting evidence in ways that help it being used by the user in evidence-based ways-- by the decision maker in evidence-based ways.

And they're all sitting up in this pyramid with data, rigorous studies at the bottom, reviews of those studies, and then you have database and maps and maps are the foundation for how to identify where can we do reviews, where can we produce guidance, what are sort of checklists should we produce.

The key takeaway points are that maps are an interactive tool to make evidence discoverable, accessible, and usable. They can be applied to a range of research questions. That it's not just about effectiveness. We looked at process evaluations, we looked at formative evaluations, we looked at maps, just as reviews. It's one thing to put the evidence out there. So the relevant evidence you put in a map depends on the research question you want to address with that map, the scope of the map, and the initial evaluation depends on your research question.

They are a building block in the evidence architecture. If they're very excited about them and want to have them, but always ask them, what are you going to do with this map? They're about actually going to the next stage of getting evidence used. The mapping itself is not usually sufficient to get the evidence used. And often, they're using maps for that purpose, and guiding people to evidence-based areas. But really, they're to get you to the next stage of doing reviews, and building portals, and building guidelines.

I hope that's been a useful overview of maps. Read more about it, again, in our paper that I mentioned. Visit our website, the Campbell website has more information available on it. And to mention, also, our What Works Global Summit in Mexico, this October, where there will be a lot of maps being presented. There will be workshops on producing maps, and there will be many presentations of ongoing evidence maps. So thank you and back over to Joann.

JOANN STARKS: Well, thank you very much, Howard, for taking time to prepare, and to give us this comprehensive overview and description of evidence and gap maps, and other reporting strategies such as toolkits, guidelines, and checklists. I want to thank everyone for participating this afternoon. We don't have any time for a question and answer today, so please ask any questions you may have when you take a few minutes to give us some feedback about the webcast by filling out a brief evaluation. The link is listed here in the slides. We will be sending an email with the evaluation link to everyone who registered. Of course, you can also send your questions to us, and we will get them to Howard.

Finally, I want to thank all the AIR staff, along with Howard White and other representatives from the Campbell Collaboration who helped with planning and logistics, and of course, we want to thank NIDILRR for their support to offer these webcasts and other events. We also want to invite you to attend the final webcast in this series, an overview of meta-analysis and effect size, that will take place on Wednesday, June 19th, at this same time. As a courtesy, everyone who registered for a previous session will be registered for all sessions in the series.

We will be holding some follow-up activities as part of the training course, so information will come soon on how you can participate. We also invite everyone to visit our recently refreshed and updated website at www.ktdrr.org. Thanks again, everyone, and good afternoon.