

ROOT OF CONFLICT

Root of Conflict Podcast

Episode: Evaluating Peacebuilding Interventions

featuring

Ada Sonnenfeld, Former Evaluation Specialist, International Initiative for Impact Evaluation (3ie)
Current Senior Expert, Monitoring, Evaluation, and Learning, Integrity

interviewed by

Reema Saleh, Pearson Fellow
Mwangi Thuita, Pearson Fellow

October 26, 2021

(recorded in January 2021)

Full Transcript

Reema Saleh: Hi, this is Reema and you're listening to the University of Chicago Public Policy Podcast. You're listening to Root of Conflict, a podcast about violent conflict around the world and the people, societies and policy issues it affects. In this series, you'll hear from experts and practitioners who conduct research, implement programs and use data analysis to address some of the most pressing challenges facing our world. Root of Conflict is produced by UC3P, in collaboration with the Pearson Institute for the Study and Resolution of Global Conflict, a research institute housed within the Harris School of Public Policy at the University of Chicago.

How do researchers assess the impact of peace building interventions? And what can we learn from examining the existing literature as a whole? My name is Reema and, in this episode, Mwangi and I speak with Ada Sonnenfeld, a former evaluation specialist with the International Initiative for Impact Evaluation.

Ada Sonnenfeld: So, I have a technical background in impact evaluations and other types of program evaluations for international development, with a focus on evaluations and evidence in fragile contexts, particularly peace building and governance.

Reema Saleh: She talks about her work, managing systematic reviews and evidence gap map projects, which can help policymakers make more informed decisions about how to use evidence to make sense of what we know and learn from what has been done before. We discuss her recent review where she and her colleagues synthesize evidence on programs that promote intergroup social cohesion in fragile context. So, first off, what is an impact evaluation?

Ada Sonnenfeld: So, an impact evaluation is an evaluation of a project, program, policy, that tries to establish not only what changed, but what part of that change can be attributed to the policy, program or project. And so, this might be done through statistical means where you can say, using either randomization or quasi-experimental designs, use econometrics, to identify what of that change you can say with some reasonable level of certainty was due to what you did or what you're evaluating, rather than all of the other factors at play.

Mwangi Thuita: Why would someone want to do an impact evaluation? Why are they important?

ROOT OF CONFLICT

Ada Sonnenfeld: Impact evaluations help us understand what impact we are having. So, you would want to do this if, for example, you're a government, and you're trying to understand whether your policy to reduce inequality is having an effect on inequality. Or whether your policy to keep more children in school is actually keeping more children in school. Especially for government policies, these tend to be very expensive. And so, you want to make sure that the money that you're spending is having the expected results. Impact evaluations are expensive, so there are many types of programs that may not be conducive for impact evaluation, where it may not be the most relevant type of evaluation. But in general, you would want to do this to be as sure as you can be, that your impact is what you think it is.

Mwangi Thuita: Some people describe the increased popularity of impact evaluations as part of the measurement revolution and development. Aid and development organizations, they now expect impact evaluations for a lot of projects they fund. Does this expectation of evaluation affect the program design? Does it improve things?

Ada Sonnenfeld: There's a lot of things within that question. So, there are definitely more impact evaluations that are happening. 3ie has a repository of impact evaluations that at this point has over 4,000 international development programs and that is rapidly growing. When we look at the number of impact evaluations published per year, particularly around 2009, you see a big uptick in evaluations from lower- and middle-income countries that are published. So, that's great, because that means that we're growing the rigorous evidence base. Whether that means that programs are being designed differently, well, you can either say, "Are programs being designed in order to be conducive to evaluation?" And you also have another question on whether or not they are using the findings from those evaluations in order to improve design. I don't think we can answer either of those questions with any degree of certainty. We work really hard to try and get impact evaluations read and used by relevant stakeholders from implementers, policymakers, other academics working on the topic. But it's hard to track that.

Mwangi Thuita: What definition of social cohesion do you use for the systematic review? I know you said it varies, but what do you use for your review?

Ada Sonnenfeld: Social cohesion has been defined by many people in many different ways. And we adapted a definition from some work that was done by Chan et al. in 2006. And then we added to that some insights from Paletta and Cullen from 2000 and Kim et al. from 2020, which was some recent work that Mercy Corps was doing with the World Bank on social cohesion. So, fundamentally social cohesion is about the state of relationships between people, institutions, government, within a society. And you can think about social cohesion as a universe in many ways. You have the vertical relationships between the state and the society, between government and its citizens. And then you have horizontal relationships about people in institutions within civil society. For the purposes of our review, we focused just on the horizontal element of social cohesion while recognizing that for building sustainable peace, vertical social cohesion is also extremely important.

Within these sorts of horizontal and vertical spheres, you also have different types of ties between individuals and groups, and those might be bridging about intergroup, or across group ties and also bonding, within groups. Again, for this review, we're focused on bridging intergroup social cohesion. So, trying to understand how you can affect the relationships between social groups.

Finally, there are five different dimensions of social cohesion that we identified from those three main sources within the literature. And that is trust, a sense of belonging, a willingness to help, and a willingness to participate and an acceptance of diversity. And that last one, acceptance of diversity, is the one that is probably the most controversial within social cohesion discourse. There are lots of authors who have argued that it is a potential

ROOT OF CONFLICT

effect of a socially cohesive society rather than a necessary component of it. So, we decided to take a bit of a theoretical stand and say that, especially when you're thinking about fragile contexts, an acceptance of diversity actually does have to be a component of your conceptualization of social cohesion, because otherwise you could think of an authoritarian state that only allowed for a certain type of citizen to live their life freely, as a socially cohesive place.

And I think if people from different groups don't all feel a sense of belonging, then you don't have social cohesion. And it doesn't matter, even in the most homogenous state in the world, there is still diversity there. And whether that's people with disabilities or LGBTQ people or whomever, there are lots of different ways in which people are diverse, and nobody has only one identity. And so, you have to be able to have some level of acceptance for different identities within a community in order for something to be cohesive.

Mwangi Thuita: What about fragility?

Ada Sonnenfeld: The definition of fragility that we used for the review was a very nuanced one. So, because within a systematic review, we have an explicit ex-ante. So, before we start the review, we say, "This is what we're going to include in this study." And anything that meets these criteria we'll include. So, we wanted to focus on fragile contexts, and in order to operationalize a definition of fragility that would allow us to screen all of the potential records against consistent criteria, we focused on saying that either it would be in context in which the fragile states index had given the country a score of 90 or above, or it would be in all in lower- and middle-income countries, we're focused only there.

Or it would be a situation in which tensions between two groups were identified as being the driving rationale for the intervention. So, this allowed us to look also, for example, for studies that might have targeted the relationships between two different gangs in Central America, or we included studies from Bosnia and Herzegovina, and Bosnia and Herzegovina are no longer classified as a fragile state, but the tensions between the two groups are still very present in society there. And there's still a lot of work trying to address the aftermath of the nineties. So, we had a definition of fragility that tried to recognize that fragility is not constant, either over time or within a country. And by saying that the focus of the study had to be tensions between groups that either were recently or were seen as at a risk of becoming violent, was the way that we tried to find relevant contexts.

Reema Saleh: What's unique about doing an impact evaluation in a conflict or a post-conflict setting?

Ada Sonnenfeld: Any kind of work that is happening in a conflict or post-conflict setting is going to be a little bit tricky because you have to be cognizant of the fact that your actions are going to interact with the context in a way that might have an impact on the conflict or the tensions or the potential. So, whereas all interventions should have some basic level of making sure that they do no harm, that bar becomes increasingly difficult to reach in a conflict-affected area, because the potential for doing harm becomes increased, because even something that you might think looks like a good intervention at first glance, such as giving vouchers to refugees in an area might have unintended consequences that create harm for those people. If, for example, you don't provide any support to vulnerable members of the host community.

Mwangi Thuita: You and your team at 3ie did a systematic review of impact evaluation literature, which covered 37 papers, I believe, and 31 unique interventions or intervention arms. So, could you tell us briefly about the systematic review, what motivated it and what were some of your main findings?

ROOT OF CONFLICT

Ada Sonnenfeld: So, that systematic review grew out of an evidence gap map that we did of interventions that aim to build people and societies in fragile context. And so, the evidence gap map had identified a cluster of programs measuring the effects on social cohesion, but it took a wide range of different approaches in order to have that kind of impact. So, the social cohesion review identified five different groups of interventions ranging from radio dramas and media for peace to classroom, school-based peace education to intergroup contact through sports, to very complex, comprehensive, large scale programs that combined a number of different approaches. Overall, the review identified a pattern of small effects on social cohesion. But that's not really very surprising, because as we've said, interventions in fragile context tend to interact with the conditions for peace, but there's a lot of other factors that go into determining the relationship between two different social groups.

And so, we don't think it's that surprising that a social cohesion intervention alone doesn't have a very large effect on the relationships between groups and fragile contexts. However, we think it's really exciting that we're able to identify a pattern of small, positive effects that you could identify and see that well actually, these programs do have a place in the peacebuilding toolkit. They just are not going to solve all of the problems, which I think makes sense. That's a very headline finding. Within the review, like I said, we identified five different groups of interventions and we looked at the impact within each of those intervention groups. And we found, for example, that radio dramas tend to have on average, a positive impact on trust, and another group of interventions related to comprehensive multi-component programs that included elements of peace education, where they would hold workshops with community members.

And then from that workshop, they would then set up opportunities for people from the different groups to interact with each other, such as through negotiation committees or early warning systems. And then they would add to that an element of economic support. So, a way for people to work together by identifying a program, a small intervention that they could do in their community that would benefit both groups. And those kinds of comprehensive programs, we found an average and positive impact on trust and a willingness to participate. Amongst the school-based peace education interventions, the ones working with children, we identified positive impacts of the programs when they measured effects on the children who participated. There was one study that measured the effects on parents who did not participate, and we didn't find any effects there. And what we think that means is likely that such a school-based peace education program, working with children, might have a lot of capacity to influence how the children and the youth or the teenagers see each other, but that might not be sufficient for changing the way that the adults see each other, and you likely need to engage them directly.

Reema Saleh: So, one of your findings is that standalone interventions may not be enough to build resilient, social cohesion in fragile contexts without complementary interventions. So, what kind of complementary interventions do you have in mind? Is it realistic to expect major changes to group relationships without them?

Ada Sonnenfeld: I think it's not realistic to expect major changes to group relationships only through social cohesion interventions. I think they have a clear role to play, but fundamentally, the drivers of conflict in any situation are rarely identified as purely being about deep-seated intergroup prejudices. Prejudices are held by everybody, everywhere, but they don't tend to turn into violence except when there are other triggers at play. And so, I think having interventions that address those systemic drivers of conflict is very important. What those might look like will vary a lot on the particular context you're looking at.

So, there might be one context where there are major economic inequalities in that striving groups. Or there are even just perceived inequalities between how different groups are treated by the government and that might be driving tensions. In other situations, it may be tensions over the way that land is used. One type of community may want to use it in one way, another might want to use it in a different way.

ROOT OF CONFLICT

So, there's often something else that's driving conflict. And that's why it's important to be very cognizant of the local context in which you're working and understand how your intervention may interact with those situations, but also to be realistic then about what you may or may not be able to change. When it says, "We need these complementary interventions addressing structural drivers of conflict," that is not something that it's likely that any one actor can influence. And that's where you need the peace building community more as a whole, in a given context to say, "These are the different drivers that we can identify." Is there a way that we can say, "Okay, this funder might focus on this element?" This funder might focus on that one and try and build a program and just coordinate in terms of how you're doing the various approaches, and more coordination might help.

Mwangi Thuita: Your findings also suggest that social cohesion programs that identify bottlenecks to intergroup social cohesion and carry out conflict assessments tend to have a larger and more positive effect. So, this kind of seems obvious, but it's also important to understand whether or not a particular context actually needs an intervention and the lack of relevance or appropriateness to the context can be at least part of the reason for seeing no impact. Did you find that most impact evaluations do comprehensive assessments like these and if not, why?

Ada Sonnenfeld: We're always working within systematic reviews from a place of imperfect information. So, what we do in order to identify the papers is we do a really extensive search of academic databases, websites from different actors, such as the World Bank and relevant implementers and donors. And we try to find all of the impact evaluations that we can that meet our criteria. And then we do an additional search for every study that we find that meets our criteria to identify other documents written about that program, to help us get as much information as we can about what they did. But we often can't find that information. And so, while within our study, we found only one or two impact evaluations that were clearly based off of conflict analyses and based on context assessments, that doesn't mean that none of the others did that. It just meant that we weren't able to find those studies and they didn't mention having done them.

So, just with that caveat in mind, I did think nonetheless, that it was surprising that very few of them mentioned having been based on conflict analyses, but I don't know if that's just because it wasn't reported or if it actually didn't happen. To your point as well, in terms of why they may not do that or why they may, I think it is surprising, but I also think it's not uncommon. It's not unique to social cohesion interventions or peace-building interventions. Other work that I've done for other types of interventions in fragile and in non-fragile context has also identified a similar finding around how often the bottleneck seems to have been misidentified. And that might relate to the fact that the intervention was wrong, in the sense that they were trying to implement something that wasn't needed in that context. But another potential source of that is that what they were measuring might not have been quite right.

That's where things get really complicated, because what social cohesion means is very context dependent. And so, you can take the example of Nigeria, where we had four different studies that took place in Nigeria, and two of them measured farmer and pastoralist communities, one targeted Christians and Muslims, and the third targeted people from different ethnic groups in the country. And so, those are three different types of social cleavages that different interventions were targeting just within a single country. So, what the social cleavages that you're targeting and how your intervention changes perceptions across that cleavage, is going to be very context dependent, and then how you measure it will also change. So, what it means, for example, to measure acceptance of diversity, a lot of people looked at whether or not people had friends from the other group, but they measured that in different ways. And they didn't always measure that in ways that were necessarily relevant.

ROOT OF CONFLICT

And that's really tricky because maybe one way to deal with that is to say, "Oh, well, let's all measure the same thing." but then, what if what you're measuring doesn't make sense for that particular context? So, you end up in this situation where there's an issue with bottleneck identification, but it's hard for us to say whether that's because they didn't do good baseline assessments of what the conflict dynamics were and what the needs were or whether it's because they weren't measuring things quite correctly. All I can tell you is that we couldn't find evidence of conflict assessments having been done. And we think they probably would be useful.

Mwangi Thuita: Given that social cohesion is often very contextual, so if you need a theory of social breakdown in each case that you're looking at, which involves contextual information – does that have implications for how generalizable the findings of impact evaluations are across the board?

Ada Sonnenfeld: Absolutely. I think one of the conversations that we've been having with Ray and with a lot of other actors who are looking at social cohesion as a way of working towards sustainable peace in fragile contexts, is that there's a need for a framework that is general enough that everybody can say, "Yeah, this is what we mean when we're talking about it," but that the indicators can be hyper contextualized. And so, that you know where on your framework your indicator feeds in, but the indicator itself is based off of the local context. And that might help us move to a place where we can say, "Okay, this change in this context represented a big leap in the relationships between the two groups."

Whereas in this context, all they measured was something that actually was quite a small step and that can help having a sense of where something maps onto a common framework would help us interpret the findings across contexts and help us better figure out how to use the findings from one impact evaluation in another context. Because that might say, "Okay, this evaluation, this intervention in this context actually had a really big impact on trust." And maybe that helps us see why and how we can take that to another place.

The realist in me likes to always say that interventions themselves can't be replicated, but mechanisms can be transferred. And what we mean by that is the design will always have to be contextualized of your intervention. But the reactions that your design is trying to trigger in the people it targets, you can try and learn from that. So, if you can get people to work together collaboratively, that's a mechanism that you might be able to replicate, even if the way that you get them to work together, and the context, the setup might be very contextualized.

Reema Saleh: Do you think that evaluations are useful for testing assumptions about how development interventions affect change?

Ada Sonnenfeld: Yes. [Laughter] I do think they are useful for testing assumptions. I think they're very useful. Specifically impact evaluations can give us a lot of information about that, but it depends on how the impact evaluation is designed. I think increasingly we see impact evaluations being theory-based and using a theory of change. And that's incredibly important because that means that they have thought about. "All right, in order to get to social cohesion, these are the steps that need to happen." And then you can see if, for example, you find a positive impact on an early-stage outcome, but not on a later stage, you can see where your theory of change might break down and then you can test your assumptions to try and see why that might be. Some really clever impact evaluations have done specific tests of different mechanisms to try and see what was driving change. And those are really interesting.

But you still have a lot of programs that just measure those sorts of high-level impact outcomes. And then you don't really know what goes on in the middle. And that's what we often call the black box of a randomized controlled trial, for example, is a classic one where you don't really know why you're seeing the results you're seeing. And so, what we would say is that that's why mixed methods are so important. You have your statistical

ROOT OF CONFLICT

methods to answer one part of your question, but that alone is likely not going to be enough without some kind of process evaluation, qualitative information, trying to see why you're seeing those results to help you interpret them correctly.

Mwangi Thuita: Would you say that these impact evaluations do a good job of measuring intermediate effects?

Ada Sonnenfeld: I would say that very few of the studies in our systematic review measured intermediate outcomes and effects on intermediate outcomes. So, whether they are capable of doing a good job, yes, they are very much so, but it's not often that they do. And that's I think really where I would like to see the field moving towards is "Okay, we're getting to the point where we recognize that having a theory of change is really important for any type of evaluation, impact evaluations and other types, because it can really help structure what kinds of questions you ask." But what I often see is that evaluations, even where they have a theory of change at the beginning, will not revisit that theory of change after they have their findings to say what do these findings actually mean for my theory of change? Do they validate it? Do they challenge it?

Do they suggest actually it should be refined in this way? And maybe this is what the theory of change should look like. So, I often feel like that last step of closing the loop. And "All right, here's our initial theory of change." This is what we thought was going to be happening. We measured outcomes against X, Y, and Z steps. So, intermediate steps and final impact outcomes. This is what we found and they'll often leave it at that. But that can sometimes be difficult if they then don't tie that back, because it can be really hard to interpret why you might see positive effects on some indicators and null effects or mixed effects on others. And so, you really need the researchers who are working with the program team. They're the best place to then say, "Okay, what does this mean for the theory of change?" And that will also help us when it comes to understanding how the findings from that study might inform future studies as well.

Mwangi Thuita: And one thing that I think was intentional in your review is you don't include interventions that aim to build sustainable peace by providing economic support for things like job training. So, like cash transfers also. Could these be some of the complementary interventions that you were talking about earlier?

Ada Sonnenfeld: Yes and no. I mean, the reason that we didn't include those was not because we don't think they are a relevant approach to building social cohesion, but rather it's because the evidence gap map that I mentioned earlier actually identified a large number of ongoing studies of cash transfers that are trying to measure outcomes on social cohesion. And so, that would have meant that, for us to have synthesized that literature now would be a bit premature because there are so many ongoing studies, those findings would change within the next two or three years. And so, that was the rationale behind excluding those from our study. I think it will be really interesting to synthesize that literature in about two or three years, not right now. I mean you could right now, but it's likely to change.

Whether or not those address underlying drivers of contentions between the communities, I think is a slightly different question that that synthesis will probably have to answer. Cash transfers can be really important in humanitarian aid context and in addressing short-term needs. Whether they are the structural changes that you need in order to shift the situation for those communities in the long run is a question that's still open.

Reema Saleh: I was curious why there were a lot of countries that never had impact evaluations.

Ada Sonnenfeld: Why that might be?

Reema Saleh: Yeah. I was curious kind of why it's kind of uneven.

ROOT OF CONFLICT

Ada Sonnenfeld: It's very uneven. I mean, the evidence gap map is maybe a better source of that than the systematic review, but you can see there that some of the analysis we did, there's not an obvious correlation between, for example, how fragile a country is or how much ODA it receives, how much official development assistance it receives, and how many impact evaluations there are. You have quite a large number comparatively of impact evaluations from Afghanistan and DRC than plenty of other places like Syria that receive a huge amount of ODA. And Yemen. So, I don't know why there haven't been evaluations in those places. That's not a question that our research was able to ask. You would have to do a lot of stakeholder research and asking all of the different donors and all of the different universities why they don't research those areas.

But what we try to do is just say, hey, there's some really important geographic gaps where we don't have rigorous evidence. And maybe hopefully people will read the evidence gap map and see that and say, actually it would be really beneficial, not just to our own programming, but to the global evidence-base to build evidence from those less well-studied contexts.

Mwangi Thuita: Yeah. Well, thanks. Thanks Ada so much. Thanks for all your time.

Ada Sonnenfeld: You're very welcome.

Reema Saleh: Thank you for listening to this episode of Root of Conflict, featuring Ada Sonnenfeld. This episode was produced and edited by Aishwarya Kumar and Reema Saleh. Check our show notes to access the full report discussed in this episode. Special thanks to UC3P and the Pearson Institute for their continued support of this series. For more information on the Pearson Institute's research and events, visit thepearsoninstitute.org and follow them on Twitter.