WhatWorks
TO PREVENT VIOLENCE

UKaid
from the British people

WHAT WORKS TO PREVENT VIOLENCE AGAINST WOMEN AND GIRLS: RESEARCH AND INNOVATION PROGRAMME

THEMATIC PAPER 1: RESEARCH

Mapping the complexities and challenges of researching violence against women and girls in developing contexts

iMC worldwide

COMMS CONSULT | Specialists in strategic communications

UNIVERSITY OF PORTSMOUTH

Thematic Paper 1: Research – Mapping the Complexities and Challenges of Researching Violence Against Women and Girls in Developing Contexts

March 2020

Authors: Bradley, T. with Crawford, S. and Lloyd-Laney, M.

About What Works: The What Works to Prevent Violence against Women and Girls programme is a flagship programme from the UK Department for International Development, which is investing an unprecedented £25 million over five years to the prevention of violence against women and girls. It supports primary prevention efforts across Africa and Asia that seek to understand and address the underlying causes of violence, and to stop it from occurring. Through three complementary components, the programme focuses on generating evidence from rigorous primary research and evaluations of existing interventions to understanding What Works to prevent violence against women and girls generally, and in fragile and conflict areas. Additionally, the programme estimates social and economic costs of violence against women and girls, developing the economic case for investing in prevention.

Conflicts of Interest: The authors confirm that there were no conflicts of interest identified for this evaluation. The team were also able to work independently and freely without interference from outside sources.

CONTENTS

ACRONYMS AND ABBREVIATIONS

DFID   Department for International Development
DRC   Democratic Republic of the Congo
GBV   Gender-Based Violence
IPV   Intimate Partner Violence
IRC   International Rescue Committee
RCT   Randomised Control Trial
REF   Research Excellence Framework
SAMRC   South Africa Medical Research Council
ToR   Terms of Reference
UNDP   United Nations Development Programme
VAWG   Violence Against Women and Girls
VSO   Voluntary Services Overseas
WW-VAWG   What Works to Prevent Violence Against Women and Girls (Programme)

1 INTRODUCTION

1.1 THEMATIC PAPER OVERVIEW

This thematic overview is one of three papers that form part of the outputs for the independent evaluation of the 'What Works to end Violence against Women and Girls' research programme. What Works is DFID's flagship VAWG knowledge generation programme (WW-VAWG), and it is divided into three components global programme, conflict and crises, costs of violence.[1]This paper is intended to draw out the important lessons that have emerged across these components in relation to research design, data collection, ethical protocols, data analysis and the identification of key findings. In particular, it seeks to better understand the legacy of the first WW-VAWG programme in research terms, and to be realistic in presenting the challenges and the key lessons learnt. Appreciating the reach and significance of the research and where it leaves us is important not just for the global movement to end VAWG but also in terms of informing where the next phase of the programme needs to focus its efforts.

The WW-VAWG programme has been pioneering in driving the formation of a new and innovative evidence base to support better resourced and more effective interventions to end VAWG. The collective body of research that it has generated makes a significant contribution to raising the awareness of VAWG as an urgent issue, drawing particular attention to the long-lasting and far-reaching implications of it for the rights and wellbeing of women and girls, as well as economic growth. Crucially, it also demonstrates that VAWG can be prevented, the policy and programming implications of which cannot be overstated.

When evaluating the combined efforts of the three components, many positive outcomes, including examples of rigorous new data sets, the development of new research capacity, and successful programme design and evaluations of VAWG interventions, can be identified. This paper brings together the successes of What Works as a collective body of research, doing so in a way that celebrates its achievements whilst also highlighting the remaining gaps that should to inform future research on VAWG.

1.2 METHODS

Building on three years of engagement and experience as the evaluators of WW-VAWG between 2017 and 2020, sixty interviews were conducted with various actors who have been involved with the research, in different ways. Interviewees included senior leads, data collectors and technical support, implementers, and a range of stakeholders, some of whom were engaged as national advisors and others who were positioned outside of the programme but were active in the end VAWG space. Additionally, a small number of interviews were conducted with researchers who work on VAWG, but who were not part of the WW-VAWG programme. The findings from these interviews have been analysed as a qualitative data set, using a thematic coding approach. The key observations that emerged through and in the data are reflected in the thematic areas highlighted below.

The paper is structured in the follow way, Section One provides an overview of the academic publications produced by the WW-VAWG programme so far. It should be noted, however, that given the inevitable time lag between writing and publishing peer reviewed research there will be many more papers to come. The summary considers the different approaches to publications taken by each component, as well as the different disciplinary challenges in achieving outputs, and reflects on co-authorship, what it means, and how it has been positively achieved. Section Two reviews the methods used, beginning with the ground-

---

[1] For more details see https://www.whatworks.co.za/

breaking use of Randomised Control Trials (RCTs) in measuring social norm interventions. Section three then goes on to review the operationalisation of each component in terms of partnerships, capacity development and ethics, and also explores the quality and innovativeness of the data collected. The conclusion begins by considering the impact of the research beyond the immediate What Works community, and ends with a summary of the key achievements and the new research questions left by phase one.

2 ACADEMIC PUBLICATIONS

2.1 WHAT WAS PRODUCED BY WW-VAWG?

At the point of compiling this paper the academic outputs for each component stood at 83 for Component 1, 3 for Component 2 and 10 for Component 3. The differences in number of outputs reflects differences in the size of each and also their publications strategies (see details below).

2.2 ASSESSING THE RESEARCH GENERATED: AN ADAPTED RESEARCH EXCELLENCE FRAMEWORK

Both at midterm and final evaluation, an adapted UK Higher Education Research Excellence Framework (REF) [2], which was developed by the evaluation team, was used to review all academic papers produced through WW. The assessment criteria followed the same principles as the REF in considering the rigour, reach, significance and impact of each piece of work.[3] The adaption of these criteria reflected the specific concerns of the programme as a whole, which were:

- To generate impact at the level of policy and practice;
- To build capacity to conduct applied research, evaluation design and data collection;
- The operationalisation of interventions; and
- The building of reciprocal North-South and South-South partnerships.

The assessment also considered the extent to which impact at these levels was demonstrated through the specific academic publications.

Given that all academic outputs have been through a rigorous academic peer review as part of the publication process, it was not the intention of the assessment to judge the quality of the data, as such, but rather to pull out the specific methodological and knowledge contributions made by each piece. The assessment conducted for the final evaluation also takes a step back from the separate components to offer an overarching summary of the diversity of methods, findings and approaches across the research programme.

2.3 WHAT WE FOUND?

In adapting the REF grading system of 1-4 stars (with 3* being internationally recognised and 4* reflecting 'world leading' research), the evaluation found that almost all publications across the three components from the beginning of the WW-VAWG programme were rated as 3*, meaning they offered significant

---

[2] See https://www.ref.ac.uk/ for more details.
[3] In addition, further categories were added that sought to assess the research uptake impact of the research generated by the programme. Please see the evaluation framework annex for more details.

originality and provided new evidence that filled knowledge gaps. The few early publications that were rated as 2* are not without an important function in the programmes as a whole as they often consisted of literature reviews or reviews of secondary data sets, which as part of the overall body of research helped to highlight the innovation and importance of the primary research that later emerged.

These initial publications came before the primary data collection had really begun and before strong working partnerships between researchers in the North and South, and between researchers and evaluators in the case of Component 1, had been forged. Once these relationships began to flourish, the legacy of WW-VAWG emerged both in terms of research quality, because of the new data collected through these local partnerships, and in the sophistication of the analysis, which benefitted from the knowledge and expertise that exists within each research team.

2.4 FOUNDATIONS OF SUCCESSES

Whilst an important tool, the REF style grading system only demonstrates the success of the research outputs. In order to better understand the reasons for this success, we undertook interviews with many key actors to unpack how and why it was possible to produce such high-quality research that was able to fill knowledge gaps. We found that the foundations of this success included, a clear publication strategy, employing mixed methods during both data collection and analysis, and building strong partnerships through co-authorship with southern-based researchers. The importance of having a clear publication strategy from the outset is a key lesson in terms of ensuring that a flow of outputs begins to emerge quickly and can be sustained. Component 1, for example, had a particularly robust publication approach in place from the beginning, which consisted of a core team of senior researchers, with a proven track record and profile of publishing on VAWG. This in turn led to very early outputs on Component 1 that drew on the data and literature already available.

The immediacy of these outputs allowed Component 1 to begin to establish an international identity as a global programme on VAWG prevention, thereby building its audience in preparation for future publications. Outputs in the first years, as already stated, mainly focused on literature reviews and secondary analyses of the World Health Organisation multi-country study on VAWG[4]. These fulfilled, and still fulfil, an important academic purpose: they laid the foundations of subsequent research by providing a clear map of knowledge gaps, which continues to support and guide the work that is still to come.

Taking a mixed method approach to research also resulted in publications being ranked 4* (i.e., 'world leading') more often. This may be partly because mixed methods in general is a relatively new approach to research, and therefore provides more opportunity for novel insights to emerge. It may also be because the questions that are best addressed by mixed methods, like those relating to WW-VAWG, are typically complex and multiple faceted, with separate and overlapping components that may each be best explored by one approach or the other[5]. As Creswell and Tashakkori (2007)[6] noted, developing meaningful and effective mixed approaches to research methods, for various paradigmatic and political reasons, including publication strategies, has been neglected in favour of more easily disseminated research. However, in terms of VAWG research, mixed methods offer a promising means of advancing

---

[4] Chan, K. L., Fulu, E., Haardörfer, R., Huque, H., Jewkes, R., McCook, S., Miedema, S., & Roselli, T. (2017). Pathways between childhood trauma, intimate partner violence, and harsh parenting: findings from the UN Multi-country Study on Men and Violence in Asia and the Pacific. *The Lancet Global Health, 5*(5), 512-522..

[5] See e.g., Tariq, S., & Woodman, J. (2013). Using mixed methods in health research. *JRSM short reports, 4*(6)

[6] Creswell, J. W., & Tashakkori, A. (2007). Developing publishable mixed methods manuscripts. *Journal of Mixed Methods Research 1*(2), 107-111.

the study of why violence happens and how to confront it. As Testa, Livingston and VanZile-Tamsen[7] noted, "*Integrating semi-structured interviews and qualitative analysis into a quantitative program of research on women's sexual victimization has resulted in valuable scientific insight and generation of novel hypotheses for testing*" (2011, p. 236). Though much of those research for WW-VAWG used multiple methods (qualitative and quantitative), those publications that *combined* and *integrated* both quantitative and qualitative data and analysis, and did so from the start, were consistently rated 4. *.This is because the qualitative approach was able to provide depth and nuance that could inform and shape the subsequent research, and could then be triangulated with survey findings to provide more robust conclusions.[891011] This assessment does not mean that a 4* rating cannot be achieved through quantitative or qualitative methods alone, but rather that in answering the key overarching question, '*What works to end VAWG?*', mixed approaches provide the necessary nuance and complexity.

Though mixed methods are clearly important, the research design and the programme as a whole has largely been dominated by two disciples; public health and economics, which, though often supplemented with qualitative research, has resulted in a general tendency toward focussing on quantitative data and methods. That said, researchers from a range of social science backgrounds were part of the research teams, and they were present in each context and component. But in the next iteration of What Works, the greater inclusion of specialists from disciplines in which qualitative and mixed methods are used more frequently would help to further deepen the analysis. This will contribute to a greater understanding of why specific interventions work and others do not, and will also help to identify which types of changes are likely to be sustainable. It should be understood that this critique is not directed at the researchers, but at the limitations of the research environment in which research of this kind is conducted.

With all of this in mind, it is worth acknowledging that publishing mixed methods research can be difficult, in part because many journals focus exclusively on *either* qualitative or quantitative studies, or are heavily biased towards one or the other. This is in part due to the lack of familiarity with mixed methods among journal editors, including in health sciences journals, as noted by Curry and Nunez-Smith (2017). Creswell (2007), and many others since[12] (in the Journal of Mixed Methods Research and elsewhere) have discussed potential strategies for successful production and publication of mixed methods research. But even if these strategies were to have been applied, there was an inherent restriction on this approach embedded within WW-VAWG, in that the DFID ToR stated that particular research questions, i.e., ones that required quantitative methods to be used, were to be answered in the first instance.

For future research, it is recommended that a greater emphasis on qualitative research that is integrated with, rather than positioned alongside, quantitative research is given greater emphasis. In particular, sociology, social anthropology, and political science as disciplines are highly sensitive to context. Social anthropology, for example, brings with it the capacity to develop a longitudinal lens that can track and

---

[7] Testa, M., Livingston, J. A., & VanZile-Tamsen, C. (2011). Advancing the study of violence against women using mixed methods: Integrating qualitative methods into a quantitative research program. *Violence against women, 17*(2), 236-250.
[8] Clark, C. J., Ferguson, G., Shrestha, B., Shrestha, P. N., Batayeh, B., Bergenfeld, I., ... & McGhee, S. (2019). Mixed methods assessment of women's risk of intimate partner violence in Nepal. BMC women's health, 19(1), 20.
[9] Gibbs, A, Jewkes, R and Sikweyiya, Y. 2017 Avoid Bad Friends": The Role of Social Contexts in Shaping the Transformation of Masculinities in a Gender Transformative and Livelihood Strengthening Intervention in South Africa, Men and Masculinities 1-20 2017
[10] Maureen Murphy, Jeffrey B. Bingenheimer, Junior Ovince, Mary Ellsberg & Manuel Contreras-Urbina (2019) The effects of conflict and displacement on violence against adolescent girls in South Sudan: the case of adolescent girls in the Protection of Civilian sites in Juba, Sexual and Reproductive Health Matters, 27:1, 181-191.
[11] Kalysha Closson, Abigail Hatcher, Yandisa Sikweyiya, Laura Washington, Smanga Mkhwanazi, Rachel Jewkes, Kristin Dunkle & Andrew Gibbs (2020) Gender role conflict and sexual health and relationship practices amongst young men living in urban informal settlements in South Africa, Culture, Health & Sexuality, 22:1, 31-47.
[12] See e.g., Curry, L., & Nunez-Smith, M. (2015). Publishing mixed methods studies in the health sciences. In L. Curry, & M. Nunez-Smith. *Mixed methods in health sciences research* (pp. 309-343). Thousand Oaks, CA: SAGE Publications.

document the sustainability of positive change or shifts in mind-sets.[13] A great number of public health researchers, including within WW-VAWG, have backgrounds and/or experience in these fields, and it is important to ensure that there is sufficient flexibility within the research approach to make full use of this valuable knowledge.

2.5 CHOICE OF PUBLICATION

The **choice of academic publisher** for the research generated by each component seems to have been determined by three factors; firstly the disciplinary background of the core research team; secondly, the intended impact of the publication (i.e., which specific academic and/or applied audience it is meant for); and thirdly, the extent to which journals are receptive to approaches that may not sit clearly within one subject area.

Component 1 is largely orientated and informed by public health discourses and this has, to some extent, influenced the choice of journals in which they have published. This includes the Lancet and British Medical Journal, both of which are largely accessible to both academics and practitioners. As such, the decision to target these publications means that the potential impact at a dual level, reaching across the academic and practitioner divide, becomes possible. Conversely, some of the more qualitative outputs have been published in wider social science journals, which highlights the breadth of the research but also the inclusion of a range of perspectives within teams.

The team working on **Component 3** have stated that they feel their outputs fall between disciplines, meaning that finding the right academic 'fit' has been challenging and has taken some time. This is perhaps reflected in their choice of more interdisciplinary and area (rather than discipline) focused journals.[14]

Component 2 took a risk in their decision to prioritise non-academic reports and briefing notes before submitting to journals. Academic journals tend to be very conscious of publishing only original data that has not already been published elsewhere in any other form. Yet despite much of the data used already being available in the public domain, when approaching peer reviewed journals the team made a strong argument that the rigour and significance of the findings were such that they should also be published in an academic form.

The lesson that can be drawn from the examples of Components 2 and 3 is that it is useful to select and contact journals in advance in order to discuss their protocols and the likely enthusiasm of the editorial board with regard to publishing their data. In other words, a process of priming journals in advance may save time later; it allows the authors to get to know the journal, and perhaps write in a more targeted way and better communicate their findings to the specific readership. Academic articles are often written in a manner that is tailored to specific publications (or types of publications) with a given audience in mind. Writers can (and do) contact journals in advance to discuss their publication criteria, what their plans are in future and what might potentially exclude articles from being published with them. It is may be useful to be particularly proactive with this, and potentially even consider it at the stage of research design. It may also be especially useful when seeking to publish mixed-methods research, for which it can sometimes be difficult to find a natural home in terms of publishing.

---

[14] Lewis, J. (2007). Analysing qualitative longitudinal research in evaluations. Social Policy and Society, 6(4), 545-556.

Co-authorship between more experienced and less experienced researchers is a key part of developing capacity, and it is also a fundamental part of supporting professional development and building the publication profiles of more junior researchers, particularly those from the Global South. The WW-VAWG programme has been highly successful in co-publishing, with some outputs being led entirely by southern-based researchers and, for the first time, by the individuals concerned. These types of outputs represent a significant academic achievement.[15][16] However, a lesson that emerged for future programmes is that there is a need for consistency in publication protocols across components. In one example, the southern country research team felt they had been promised co-authorship for all outputs but, as one team member shared:

> …in reality this did not happen and whilst we collected the data we have not been involved in analysis or writing of outputs. This feels like a huge missed opportunity not least because of the amount of local knowledge and insight we could have brought to the analysis and shaping of more theoretical narratives.

(Researcher, interviewed January 2020)

2.6 EXPANDING QUALITATIVE ANALYSIS FOR SUSTAINABLE CHANGE

In some interviews with qualitative researchers, and perhaps unsurprisingly given the lens they work through, there was a feeling that many of the outputs were very descriptive and data heavy, lacking in deeper theoretical and nuanced analysis. Differences in the orientation of an article, as in more theoretical versus data heavy, reflect the different disciplinary approaches but also the specific goal of the article and the gap the research was intended to fill. The feeling from some of those interviewed for this paper, was that the data collected could have been used to say a lot more than it did in the publications. However, it should be noted that not all conclusions or recommendations that will be drawn from the data have yet emerged. Data sets take time to be analysed, and once the research paper is written it can take even longer to be published and become publicly available, particularly in highly ranked peer-reviewed journals.

It was pointed out that specific research could have contained a deeper level of analysis that allowed for a greater number of connections to be made between intersectional dimensions. This could have drawn links between life cycles and the complex cultural, political and economic environments that have an impact on the likely success of scaling up the efforts to tackle VAWG. This is not intended as a criticism of the research, as such – it was not part of the original research mandate – but is instead an observation that a gap remains and there is an appetite to address it. In relation to Component 1, for example, an intervention may work for a specific group at a specific time in their lives and at a particular moment in a country's history: understanding why is a crucial part of assessing whether an approach will work in the longer term and be sustainable. Though this was not an explicit part of the mandate for the research, it points to an existing gap in knowledge and an interest in filling it among the stakeholders.

---

[15] For example, Chirwa, E. D., Sikweyiya, Y., Addo-Lartey, A. A., Alangea, D. O., Coker-Appiah, D., Adanu, R. M., & Jewkes, R. (2018). Prevalence and risk factors of physical or sexual intimate violence perpetration amongst men in four districts in the central region of Ghana: Baseline findings from a cluster randomised controlled trial. *PloS one, 13*(3).
[16] For example, Saeed Ali, T., Karmaliani, R., Mcfarlane, J., Khuwaja, H. M., Somani, Y., Chirwa, E. D., & Jewkes, R. (2017). Attitude towards gender roles and violence against women and girls (VAWG): baseline findings from an RCT of 1752 youths in Pakistan. Global health action, 10(1), 1342454.

Assessing the likelihood of sustainability and scale through research remain as two key questions for the second phase of WW-VAWG. As previously noted, academic outputs of each component will continue to emerge, and it may be that answers can and will begin to be provided in the next few years as phase one data sets continue to be analysed. Getting to these more complex narratives takes time, and a process of digging still deeper into what the data can tell us, sometimes with supplementary data added, will likely lead to the emergence of new and critical findings. An important lesson for phase two of WW-VAWG is not to ignore or disregard the data from phase one, but instead to use it as the starting point for new research.

2.7 SYNTHESISED PUBLICATION

The process of synthesising across such a diverse programme is challenging. Synthesised publications were planned late in the programme and the approach to producing them, for some, did not feel very inclusive. Even when a country focus is shared, it is difficult to bring together various data sets, often collected through different disciplinary lenses and driven by different ideological and political objectives. Different localised social and/or political objectives may represent unique ideological perspectives. And then, due to structural and political reasons, when being synthesised across the programme some of these become more dominant and/or visible than others. The issue with deciding to create synthesised cross-component publications so late in the programme is that this issue may have potentially been magnified. Greater integration and/or communication and planning of such publications from the beginning could have mitigated this. A lesson that can be drawn, here, is that the discussion across components perhaps needed to begin much earlier and be embedded as part of the overall coordination of the programme. When it is done well an additional rich layer of insight and evidence emerges. [17]

2.8 LESSONS FROM PUBLICATION

- A consistent cross-component protocol on co-authorship is needed that recognises that 'contributing' goes beyond the literal writing up and should also include input in design, collection of data and analysis. The deliberate inclusion of local researchers in this process should be part of this protocol and recognised as essential to support the framing of the design and the contextual analysis of data. In short; all involved in the data collection, analysis and writing should be included as equal authors, with the acknowledgement that such an inclusive approach improves and strengthens the end paper whilst also supporting the careers of all involved.
- The quick generation of briefing papers in the early stages of each piece of research should be prioritised ahead of the denser and lengthier academic outputs to enable immediate sharing and feedback and to begin building a potential audience for uptake.
- Early identification and contact with journals will help to mitigate the possible risks associated with publishing data in a non-academic form first and will also enable researchers to write strategically for the journal readership, thereby increasing the chances of acceptance.
- Plans for synthesised outputs should be embedded at the programme design phase in order to ensure that the data collected supports high quality, cross-reaching thematic outputs.

---

17 Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., & Sutton, A. (2005). Synthesising qualitative and quantitative evidence: a review of possible methods. *Journal of Health Services Research & Policy, 10*(1), 45-53.

3 REVIEW OF RCTS AND MIXED APPROACHES

3.1 WHAT WAS PRODUCED BY WW-VAWG?

Across all components, a range of methods were used. Component 2 and Component 3 adopted a mixed approach, fielding surveys and qualitative interviews, focus groups and some observational methods. Component 1, in addition to the use of Randomised Controlled Trials (RCTs), conducted in DRC, Tajikistan, VSO Nepal, HTAC, Ghana and Bangladesh pre-post, modified interrupted time series studies and quasi-experimental studies as well as a range of qualitative approaches. This section draws on interviews conducted with researchers who were involved directly in data collection in order to capture their reflections on operationalising these methods. They were asked to think about what had gone well and the challenges they faced, and what could have been done better or differently. Given the relative size of Component 1 and the reach of its portfolio, many of the interviews were with researchers involved in the RCTs. Lessons on methods and approach that have been drawn from the other components are discussed later in the sections on the operationalisation of the research methods, with a focus on partnerships, capacity development and measuring rigour and significance.

3.2 WHAT WORKS IN WHAT CONTEXT?

The interview responses from those researchers involved in Component 1's research using RCTs revealed that context is a driving factor in terms of the success of the approach. For example, in Rwanda the researchers involved were positive about the application of the RCT and cited a 98% retention of their groups. In thinking through why this may be the case, it is likely that the specific political economy of Rwanda, with strictly imposed post genocide government policies on community participation, led to a greater commitment among participants to seeing the trial through than is in other contexts. In South Africa, for example, the focus on a peri-urban setting made the operationalisation of a randomised approach more complex, not least because of the fluid nature of people's lives. As documented below, migration between neighbourhoods is common, which makes the randomisation and control element of the method difficult to implement.

3.3 RCTS AS A 'QUASI-EXPERIMENTAL' APPROACH

With context in mind, a number of interviewees shared reflections on whether RCTs were applied in a strict sense or if they were in fact adapted into more of a quasi-experimental approach, "*For us, it was a randomised controlled trial but the question of the unit of randomisation, which were not many, brought on discussion as to whether we should refer to it as a quasi-experimental study*" (RCT Researcher, interview conducted August 2019). Component 1 acknowledged that it was not always possible to conduct an RCT, and so the approach was adapted according to context. This however should not be considered as a mark of less rigour, but rather as an indication that there is a need to build and shape methods in a grounded way that is sensitive to the specific location where it is to be applied.

The issue of ethics in relation to how decisions were made provoked some challenging discussions both within research teams and at the community level. For example, randomisation requires that decisions have to be made regarding who receives the intervention and who does not. Researchers shared that these conversations had to be handled very delicately with community leaders, particularly with those who would not be part of the RCT. One researcher stated that, "*I think the approach [RCT] brings up interesting ethical issues around the challenges of doing research on gender violence prevention. More funding is needed so that when an intervention is successful it can be broadened out to the control community*" (Interview with Researcher, conducted in August 2019).

As stated above the specific country and urban/rural context had a significant impact on how an RCT could be approached. For example, as one researcher recounted:

> *One challenge in the design of our What Works trial was the clustering in a peri-urban/slum setting. The challenge was that the approach required the clusters to be separate so that you can have an intervention in some of the neighbourhoods, but not in all of the neighbourhoods. But in our peri-urban settlements we had a lot of interchange with people moving from one neighbourhood to another. So, it is a challenge to know if the cluster randomised design was clean.*

(Interview conducted in August 2019)

3.4 COOPERATION AND TRUST

Arguably, the success of Component 1's RCTs was largely dependent on the trust and respect between researchers and implementers. The research partnerships in Pakistan between RtP, Aga Khan, Texas Women's University was talked about in positive terms by those interviewed. RtP claimed that it worked well, in part because researchers and implementers had a clear understanding of their respective roles and maintained boundaries sensitively. This mutual working relationship took a little while to form but once all team members understood the interventions and the research approach to be taken, a co-operative ethos was forged. In Nepal, data collection firms were used, and, depending on which project they worked on, these were guided by senior researchers from either the South African Medical Research Council (SAMRC) and Emory University in the US.

The on-the-ground relationships between the senior researchers, data collectors and implementers was positive, with the data firms reporting they had learnt a lot. However, interviews with stakeholders outside of the programme, but who were very much engaged in using the research, brought up an interesting critical question. For example, one interviewee said:

> *I don't understand why researchers from South Asia weren't used? Why weren't our country universities engaged? We have many very respected social scientists in India in particular, why it was necessary to use researchers from the USA, I don't understand, and this must have been very expensive.*

(Interview conducted January 2020)

This quotation offers a point on which to reflect in terms of value for money but more pressingly with regard to ensuring that research is locally driven and rooted. This was not the case across all components, and the research in India for C1 was, in fact, conducted by an Indian team.. Across components there was a clear and transparent process for identifying lead researchers for projects. But there was clearly a degree of discontent among some actors regarding the appointments made, which could be down to who had the final say (and where they were from) regarding who was (and was not) deemed most appropriate for the various positions. It may also simply be down to a lack of communication among actors as to why selections were made.

The need for implementers to be guided on how to understand and respond to the results emerged in a number of interviews. In particular, how should implementers respond when results are not particularly positive or do not reflect what they feel to be the impact of their practice? Sharing results as they emerge with implementers was stressed as a key part of the ethical approach that is needed in research partnerships and can be taken as key lesson moving forward. The use of an RCT approach and the

rigour it potentially offers in generating quantitative insights is not necessarily going to support the beliefs held within an organisation regarding the benefits that their activities bring. In fact, it may well suggest the opposite. Preparing implementers for this reality should be embedded into the ethical protocols of RCT research from the outset. The implementers interviewed for this review all felt supported throughout the RCT, but key to this was the open exchange of information. As one key actor shared "findings must be shared throughout the RCT process and as they emerge so that implementers can prepare to respond and think about how best to use the results."

3.5 IMPORTANCE OF QUALITATIVE RESEARCH

Given the significant impact of the RCT work, the approach is undoubtedly critical in terms of generating the data of the scale and scientific rigour needed to leverage policy and donor support. However, one way of overcoming the challenges of RCTs is to better incorporate qualitative methods. As one researcher explained:

> *I think that there is a value of doing mixed methods, whilst RCTs definitely equal impact with policy makers but I think it's really critical if you are doing an RCT to absolutely have qualitative research because if you find out that something worked and you don't know why it worked or how it worked or still it worked better than others and how that relates to the context it makes it difficult to know how to bring the intervention to scale. I think having mixed methods which speak to each other is really important. Having a really good qualitative process from the start supports the evaluation of what works and what does not. I think what doesn't work is just as important to learn.*

> (Interview with Researcher, conducted August 2019)

When qualitative approaches were used, and in particular more open-ended methods, researchers felt a level of richness was added to the analysis and findings.

> *Actually, I really liked the design and we had an ethnographic process evaluation that went alongside it that utilised a lot of qualitative research and in- person participant observation. I guess the only methodology we could have included would have been non-verbal forms of research like asking men to like draw pictures or represent things visually through photo voice. That could have been a really nice addition, especially to get to some of these unspoken traumas like violence in daily life, or childhood abuse. I think nonverbal ways of communicating might have been better.*

> (Interview with Researcher, conducted August 2019)

It is interesting to note from this reflection that non-verbal qualitative approaches and methods may draw insights and experiences from men who are particularly hard to reach in research on sensitive topics, and when it comes to sharing personal experiences of trauma.

3.6 LESSONS ON METHODS

- A mixed approach works best (including in terms of integrating and connecting qualitative research with RCTs). Qualitative research needs to be valued more highly and invested in as a part of a mixed approach. In most mixed approaches, particularly those working on highly sensitive topics such as violence, using qualitative approaches first helps understand the ethical, logistical and a contextual difficulty that can be accounted for in tool design. Qualitative methods

can then be woven through the quantitative stage and used in follow up to test for sustainability and answer questions around how to bring an activity to scale.

- This will also serve to build important relationships between team members and with local communities and key gatekeepers, ensuring the necessary enabling environment is in place. Trust, cooperation and effective communication between various actors was identified as an important component of conducting effective research. This includes ensuring that, from the outset, implementers understand the potential negative outcomes of the research, and that communities involved in RCTs are aware that not all groups will benefit equally from the potential positive outcomes of the trial.
- The inclusion of *what does not work* in outputs is critical if clearer insights are to emerge regarding *why interventions do work* and also in order to identify how money should *not* be spent (because things do not work). However, managing how the 'does not work' element is to be communicated, needs careful consideration and planning, which will require the inclusion of implementers.

4 OPERATIONALISING RESEARCH METHODS

4.1 CAPACITY DEVELOPMENT

4.1.1 WORKING IN PARTNERSHIP FOR RESEARCH CAPACITY DEVELOPMENT

Across the portfolio of What Works, there are positive and challenging examples of working in partnerships, particularly when seeking work in ways that bridge the gap between North and South or between academic and practitioner worlds. The ethos of the programme stressed the importance of developing southern research capacity and supporting national researchers to build their profiles through co-authored or single authored publications. This focus was not necessarily consistently present at the start of the programme when we saw a much greater dominance of outputs authored by northern research leads, but by the mid to end point, as partnerships and relationships grew, more co-authored and single author papers by southern researchers began to emerge.

In contexts where teams of researchers managed to work closely from the beginning, the impact and results are clearly stronger and have far more likelihood of leaving a lasting legacy in the form of increased research capacity. In Rwanda, for example, one researcher commented that:

> *on the qualitative side I worked very closely with four different Rwandan qualitative researchers and we worked together. They conducted most of the data collection and were involved in analysing the data with me and were also involved in co-authoring papers. We strove to involve Rwandan programme partners in co-authoring papers which for all of them was the first time to be involved with publications. An important element was really building capacity especially on research writing. Our southern researchers, for instance, asked the team for capacity training sessions on writing blogs and on understanding randomisation so they could speak to local leaders on qualitative sampling.*

(Interview with Researcher, conducted August 2019)

4.1.2 IMPACTS OF RESEARCH CAPACITY DEVELOPMENT

The interviews brought to light differences in approach to capacity development. The model of capacity development was slightly different across components, and some component researchers felt uncomfortable with the pressure to document that capacity had been built.

There were mixed results: interviews reveal a mixed response with evidence of capacity development in many forms but perhaps with missed opportunity for more. In some contexts, senior in country researchers took responsibility to build the capacity of their team. As one senior researcher shared, "*It felt very uncomfortable to then ask southern senior researchers who had conducted training of colleagues if they felt they had seen their capacity built*" (interview conducted in January 2020), and some of the workshops, according to some KAs, 'lacked depth'.

Lessons emerging from the planning of research capacity development focus on ensuring that a full reflexive audit is conducted of where researchers feel their strengths and weaknesses lie and their priorities for professional development, these should then form the basis of a researcher development programme. That said, there are some very positive examples of local researchers and implementers who felt they had benefitted from working with more senior and experienced researchers and technical experts. For example, in Nepal the data collectors from Equal Access spoke of the growth in their methods as a result of working with the researcher from Emory University. It was acknowledged that relationships between researchers, data collectors and implementers is always potentially fraught: "*It always remains difficult between a researcher who comes from a research background, research institute, and an implementer, and a development worker. We should work on how to equalise that power dynamic*" (interview conducted in January 2020).

It should be noted that very few of the northern and senior researchers talked in interviews about how their capacity had been affected by the experience, yet clearly it will have had a dramatic impact, not least through the acquisition of new knowledge on cultural context. In terms of breaking down unhelpful and uncomfortable power dynamics, embedding an action learning approach that encourages the whole team, regardless of seniority, to reflect on learning and learning needs as part of an on-going process may go some way toward achieving this. This approach could be also used as a mechanism to check in on the health of the partnership and respond quickly to any challenges, problems and tensions.

4.1.3 LESSONS ON RESEARCH CAPACITY DEVELOPMENT

- The need for reciprocal learning through action led approaches[18] in which team members are able to reflect on where their gaps in knowledge and skills may lie, with strategies for development worked through in collaboration with funding research lead/partner. For this work, it is necessary to acknowledge that all members, regardless of seniority, will have gaps in knowledge and areas of weakness especially when a project and or context is new. Sharing this information potentially represents a useful step in the process of building strong partnerships that are inclusive and equal and not extractive.
- Wherever possible, the inclusion of senior in-country researchers to co-ordinate ongoing capacity development and to act as a learning bridge/knowledge exchangers between northern and southern team members would also help to recognise and promote southern expertise and break

---

[18] Scott, J., Averbach, S., Modest, A. M., Hacker, M. R., Cornish, S., Spencer, D., Murphy, M., & Parmar, P. (2013). An assessment of gender inequitable norms and gender-based violence in South Sudan: a community-based participatory research approach. *Conflict and Health, 7*(1), 4.

down assumptions that it is only "*white people that have knowledge*" (Extract from an interview conducted with a country stakeholder, January 2020).

4.2 BUILDING PARTNERSHIPS

4.2.1 TAKING A LOCALISED APPROACH

Across all three components, there was evidence that an inclusive approach to research was valued. However, at a more micro level, drawing on the individual experiences documented through interviews revealed the need to ask the question: what does (or should) an equitable partnership look like? There is acknowledgement that North-South relationships need to be reciprocal, but how this might play out in the realities of conducting fieldwork in complex contexts remains unclear. The comments in this section do not reflect a pattern across the programme, but rather highlight that there are issues that need to be addressed in some instances. More generally, they are a reminder to consider the importance of taking a localised approach across the board and at all stages, which in some instances has been successful within WW-VAWG.

In many interviews, in-country researchers and data collectors shared how much they had enjoyed their experiences of working with external partners and more senior colleagues. However, the extent to which long term capacity and profiles were built through and by the process, in some instances, is less clear. For example, in a few interviews it was felt that local knowledge was not drawn on at the beginning of the design phase and local researchers were not part of the analysis of the data. In one interview, a non-academic stakeholder and country advisor stated, "*expats should not have been sent straight into the field, they should first have come and sat down and familiarised themselves with the realities of GBV [Gender Based Violence] in this country*" (Interview conducted in January 2020).

Bypassing local knowledge can lead to inaccuracies in analysis and interpretation of data, which has specific implications when it comes to making programme recommendations.[19] In one instance, the use of a data collection company outside of the field country was raised by stakeholders interviewed. The decision seemed to confuse them as they felt there were a number of in-country data collection options that could have been used. There may or may not have been a good reason for this decision but this example raises the importance of clear communication at every stage of the research process and the need to make decision making as transparent as possible both within the immediate team and also with wider stakeholders.

Some key actor's expressed the view that the data and findings stopped short of making or helping practitioners understand the implications the data had for their programming. One KA stated this felt like "*a lost opportunity*" (Interview conducted in January 2020). This raises the question how far research can actually go when knowledge gaps are so huge? Here, managing the expectations of non-academic stakeholders in terms of what they might see as outcomes from research emerges as a lesson. Non-academic stakeholders are critically important in data collection. They act as gatekeepers and are part of the enabling environment needed in order to ensure data collection is safe. One senior researcher interviewed felt that a lack of recommendations might mean that "*some of the uptake claims may be overstated.*" In other words, how can uptake be measured if recommendations are not made?

[19] Reddy, R., Hollowell, C., Liong, L., Aresto, L. L. C., Grace, N., Joseph, M. B., Ker, J. A. M., Lado, J., & Kiden, M. (2019).
 Using feminist 'reflexive practice' to explore stress and well-being of local researchers in South Sudan, Gender & Development, 27:3, 555-571.

Measurement without practical guidance is really about communication and how widely the research is known and mentioned; but is this impact?

Similar comments were made in other interviews with non-academic stakeholders: the data "*didn't really tell us anything we did not already know, but this is not in any way to diminish the importance of it.*" Practitioners interviewed agreed that whilst many of the findings concurred with their on the ground experience the rigorous evidence gave strength to their claims and empowered them to leverage government and donors for greater support. New data will always show us something we did not know before. The issue emerging here is around how best to communicate findings to non-academic stakeholders. It is necessary to distinguish between what represents new insights whilst also ensuring the evidence to support what practitioners already know is packaged in a way that enables them to use it for maximum leverage.

4.2.2   ETHICAL COMMITMENT TO EQUAL PARTNERSHIPS

The benefits of an ethical commitment to working in equal partnership can be seen in many of the examples given above. In some instances, these relationships were not fostered, and a more extractive approach was taken. In one example, the relationships between the external northern university team and the organisation was difficult, with feelings of being "used" articulated by the intervention staff involved. The specific approach trialled, as well as the data collected, were referred to in key actor interviews as 'robust' and 'impressive'. However, problems emerged at the local level regarding the need for the approach to be commercially viable in order to scale. In this particular incidence there was positive evidence that the approach works, but the patent that it relied upon is owned by the northern research institution. This example raises some tricky questions. Firstly, and referring back to the section on partnership building, in this instance equal relationships were clearly not achieved, which provides further evidence of the need for clear protocols to ensure inclusion and mange shared expectations throughout. Secondly, a question emerges in relation to how scaling up can be achieved in a viable and sustainable way. Even when commercial viability is not a factor, an intervention that is proven to work may be taken to scale through donor funding but, as we know, this funding is often available only for the duration of a cycle with no guarantee of renewal.

Thus, the following questions arise: what do we mean by scale? and how do we measure it? If an intervention is simply too expensive regardless of the evidence that it works, it will not be economically viable and will, in reality, not be implemented at any significant scale. Ethically, then, it feels questionable to be trialling an intervention that will never realistically be scaled in the context it was tested due to these kinds of constraints. The perception given by one key actor involved in the data collection; "*Local people were used to collect data only, not part of the process beyond that. The use of a Microsoft package that local researchers had no training in limited their involvement and they were not offered training.*" (Interview conducted October 2019). Once again, this points to the need for firm ethical protocols about how to manage and behave in such partnerships.

4.2.3   LESSONS FOR PARTNERSHIP BUILDING

- Partnerships must be founded on ethical principles agreed at the start and by all involved. The action led approach recommended in the lessons from capacity development (see above) could be expanded to include regular 'check-ins' with all members, ensuring that everyone feels the included and respected.
- A localised approach that is implemented from the outset, including at the design phase, may help to maximise the impact of the emerging policy and practice recommendations. This could be

done through forums of researchers and practitioners who could work collaboratively with data to draw out what works and to generate clear policy and programming messages.

5   DATA QUALITY AND INNOVATION

5.1   IDENTIFICATION OF USEFUL TOOLS TO IMPROVE DATA QUALITY

The success of Component 2's research at the country level in South Sudan is particularly striking: firstly, because of the complexity and fragility of the working operational context; and, secondly because of the high regard in which the data is held. It was noted by a number of those involved in the data collection that "*The quality was of a very high standard and the lead researchers refused to (be) compromised on this*" (Interview conducted in January 2020). Interviews with IRC staff in country highlight that, at least in part, this success was down to the amount of time devoted to working with the data collectors ensuring they were trained in the approach and equipped with the necessary tools to work on VAWG. Component 2 used both a quantitative survey and qualitative methods, which allowed not only for prevalence data and factors that contribute to increasing VAWG to be identified, but also for a clear understanding of why to emerge. In particular, the link between IPV (intimate partner violence) and bride-price[20] emerged quite clearly in analysis.

Arguably data collection in Syria around cash transfers, and in Dadaab (Kenya) around GBV case management intervention, have had less impact, but both still represent innovation in research methods. In Syria, the use of mobile phones to collect data represents an innovative tool for use in conflict situations when more conventional on the ground data collection is not possible. Similarly, in Dadaab it was not possible to carry out an RCT for ethical and safety reasons. Instead a cross-sectional survey and qualitative interviews with the case workers were used. The approach revealed the extent of case workers vulnerability and the daily challenges they faced. Whilst the data cannot evidence the impact of the interventions in terms of reducing VAWG, it does provide new and valuable insights into the levels of vulnerabilities suffered, by both women and men, in refugee contexts. The outcomes point to the need to respond to these vulnerabilities in ever more holistic and sustained ways.

5.2   THE PRICE OF VAWG?

Component 3 also set out to conduct mixed method data collection in Ghana, Pakistan and South Sudan. In South Sudan the plans for qualitative data collection were halted due to the re-emergence of conflict and the associated ethical concerns over interviewing vulnerable women in even more volatile times. The goal of Component 3 was to evidence, through the development of a new costing model, that VAWG has a significant economic cost. Whilst costing models on VAWG exist[21], they tend to focus on industrial settings. The intention of the Component 3 team was to show that even in more fragile contexts, VAWG carries significant economic costs and should be addressed as such. The ideology of this component was very much focused on leveraging more resources for VAWG interventions with governments and women's organisations intended as the main users of the model and the data.

A key objective was to demonstrate that costing models for VAWG can be applied in fragile, non-industrial contexts. Yet some concerns were raised by advisors and researchers outside of the

---

[20] The practice whereby a bride's family receives payment from the groom or his family at the time of, or just before, marriage.
[21] Walby, S. (2009). *The cost of domestic violence: Up-date 2009.* Lancaster: Lancaster University.

component over the choice of country case studies. South Sudan, as an example, represents a context in which there is virtually no formal economic activity (i.e., little to no manufacturing or industrial production) making the possibility of capturing and measuring the cost of VAWG challenging. Component 3's response was that "there is always economic activity even in the most conflict ridden of contexts. People have to try and make a living." The component also expressed that whilst there were limitations in their approach and the that the figures reached may not be 100% understanding the economic implications of VAWG is not an unrealistic goal. However, when focusing in on the issues of precision and suitability of context from a research perspective, the overarching objective of the approach is ignored. Meanwhile, the impact of the research on policy is evident in South Sudan, and it has been cited in the county's new Gender Equality legislation.[22]

There is interest in this approach in other complex contexts too, in Pakistan, a member of the advisory board stated that "the cost of violence is difficult to measure and is a new concept for Pakistan. It is an area of deep interest to understand violence and the cost which will help with advocacy work." In light of this, the research of Component 3 presents interesting questions about how we might judge 'quality': If the data generates and supports the development of new approaches and models that are in turn taken up and used by actors seeking to push for greater resourcing and political commitment, should this not be seen as a marker of quality? Applying the adapted REF matrix to the outputs of Component 3 yields rating of 3* and 4* not least because of the ambition of the research its innovative approach and the reach the costing model seems to have already achieved.

5.3 USE OF ADVISORY BOARDS

Components 2 and 3 used an advisory board in all countries, which helped to foster strong relationships of cooperation. For example, in South Sudan which generated a sense of ownership and inclusion across what is a relatively small and tight-knit GBV community in Juba. In particular, the successful engagement of the Minister for Gender in South Sudan by Component 3 led to the crucial uptake of research results in the Gender Equality Legislation, as noted above. A Component 2 advisory board member stated, *"I feel a sense of ownership because we were called upon to review the findings and our views incorporated that made us feel honoured"* (Interview conducted in January 2020). In both components, the advisory boards were consulted at the time of tool design and also analysis and validation of findings. The boards also offered advice on how to ask questions and were sometimes proactive and engaged as part of the data collectors training.

5.4 LESSONS FOR DATA QUALITY AND INNOVATION

- Country level advisory boards help can really drive the both uptake but also the relevance on the study design and tools.
- Taking methodological risks pays off and should be encouraged.

6 CONCLUSION

In this thematic paper, we have sought to map the complexities and challenges of delivering rigorous VAWG research in complex contexts, drawing on the outputs of the rich multi-year, multi-country research

---

[22]UNDP (2017). *Legal Provisions Relating to Gender Equality & Sexual & Gender Based Violence In South Sudan*. Retrieved from: reliefweb.int/report/south-sudan/legal-provisions-relating-gender-equality-and-sexual-and-gender-based-violence

undertaken as part of the WW-VAWG programme. Though many challenges remain, there is also much to be hopeful about.

Though research is certainly being used by researchers who are part of the WW-VAWG community, the reach of the findings outside of the network is hard to assess and indeed it may be too early to see. The thrust of WW-VAWG with regard to uptake and applied impactful data means that the decision to focus outputs on policy makers and programming was rightly placed. That said, the number of citations documented in final reports, particular of Component 1 is impressive and indicates a degree of visibility and acknowledgement that means the body of work WW-VAWG represents cannot be (and is not being) ignored by others in the field. A number of research who are not part of WW-VAWG were interviewed to test the reach of the research. Not everyone had heard of the programme, but dissemination is still on-going, and as noted in the introduction to this paper, publications take time. Once published, there is typically a further delay in findings being picked up at discourse level. Researchers who were not part of WW-VAWG said, "I have definitely heard of it, it is important and exciting" and, "*As a body of work it has helped me conceptualise my work in a more global way*" (Interviews conducted in August 2019).

The key findings from this Thematic Paper on the WW-VAWG research process and outputs begin with the endorsement that, as a whole, the programme has delivered on its objectives: the global evidence base has been further built, we now know more than ever about the prevalence of violence, the main triggers and also what works to prevent it, and We also know much more about what does not work, which should be considered as equally important. We are at a moment in history when VAWG is considered a considerable global challenge, and donors are listening. As such, the money spent to end VAWG needs to work and the evidence to prove it has is critical. To this end, the programme has demonstrated the value of research as a means to leverage global support and resources. However, high quality research on its own may not be enough; political will and an engaged civil society with capable implementers are also crucial factors. We also now know where and how to do better research more effectively on VAWG. We now know how to position findings to support, drive and sustain these other factors that are needed to prevent VAWG in the longer-term, as well as which methods are most likely to result in the generation of data strong enough to make policy makers pay attention and act.

As with any good research programme, we are perhaps left with more questions than answers. For example, how can we scale the interventions that we know work? How can we assess whether interventions that work in one context are transferable? How can we measure sustainability and use these methods to feed into adaptive management approaches that can respond to change and ensure interventions continue to work? Now we know how prevalent VAWG is in fragile contexts, how might we work to prevent it in instances of humanitarian crises? How can we support women's organisations to use costing as a mechanism to leverage resources for their work? We also need to drill deeper into intersectional dimensions such as disability and age. It often remains unclear whether (and if so, why) interventions are more effective with some groups compared to others, which is linked to the question of how to reach and programme effectively the most vulnerable groups. Most importantly, how can we reverse the social norms that underpin VAWG once and for all, and do so across generations without leaving anyone behind? The foundations to answering these questions have been solidly laid, and the discourse now needs to shift to focus on the new knowledge gaps that have been identified as a result.