



UNITED NATIONS  
UNIVERSITY

United Nations University  
Office in New York  
Policy Brief - March 2016



---

# Assessing the Impact of DDR Programmes: Possibilities and Challenges

Franziska Seethaler<sup>i</sup>

---

i. [fms2143@columbia.edu](mailto:fms2143@columbia.edu)

© 2016 United Nations University. All Rights Reserved.  
ISBN 978-92-808-9029-7

## EXECUTIVE SUMMARY

Though monitoring and evaluation (M&E) exercises are increasingly incorporated into the design of disarmament, demobilization, and reintegration (DDR) programmes, thorough assessments of DDR programmes often are not systematically and effectively conducted. Ill-conceived assessments, however, can yield wildly inaccurate impressions of the effectiveness of these programmes, waste scarce donor resources, and prevent the necessary modification of DDR programmes — to the detriment of programme beneficiaries and donors alike.

This policy brief provides an overview of the state of impact assessment in the DDR community and offers practical recommendations to improve it. Drawing on 28 interviews with DDR scholars and practitioners from various organizations and an academic and programmatic literature review, this policy brief first provides an overview of current trends and types of assessments. It highlights knowledge gaps and limitations that, despite a growing critical awareness of the importance of assessments, continue to exist. Next, the policy brief discusses challenges faced in conducting rigorous impact assessment of DDR programmes. These include the lack of an early integration and prioritization of impact assessment, short-term budget cycles in DDR programming environments, and weaknesses in measurement approaches.

The policy brief concludes by providing recommendations on practical steps that the various actors in the DDR community can take to enhance the rigour and cost-effectiveness of DDR assessments. First, it recommends that donors and Member States invest in strengthening the M&E expertise of the DDR human resource pool, through training and other measures. This would allow for more high-quality assessments, more standardization of evaluations across programmes, and improvement of the cost-efficiency of M&E efforts. Second, the paper considers how new technologies may facilitate effective M&E in non-conducive environments. Finally, the paper stresses the need for longitudinal assessment — for example, through cooperation with universities and research institutes. The paper closes with a reminder that improving the effectiveness of DDR programming will depend on there being latitude to report failure. Incorporating these recommendations into DDR programming will significantly enhance the quality, cost-effectiveness, and utility of impact assessments of DDR programmes.

## INTRODUCTION

For almost three decades, disarmament, demobilization, and reintegration (DDR)<sup>1</sup> programmes have been used as a stabilization tool to reintegrate former combatants into societies emerging from conflict. Over this time, the characteristics and requirements for these programmes have significantly changed: originally constructed as a post-conflict reconstruction tool, DDR programmes are today often used in contexts where there is no peace to keep or where offensive operations are ongoing.<sup>2</sup> With such complex stakes in play, critical evaluations of these programmes are all the more important to ensure effectiveness and to modify the programmes where necessary.

Monitoring and evaluation (M&E) has become a buzzword in recent years for learning and reporting exercises that help guide the DDR process, in particular in ever-more-volatile environments,<sup>3</sup> and that help enhance accountability and transparency to donors and Member States alike. Yet the United Nations has identified M&E as “one of the weakest areas of disarmament, demobilization, and reintegration programme management.”<sup>4</sup> In particular, impact assessments<sup>5</sup> of DDR programmes remain an understudied research area and are not systematically and effectively conducted, to the detriment of programming. However, without robust assessment it is difficult to assess the effectiveness and efficiency of these programmes, tailor them to local contexts to maximize the impact for beneficiaries, and ensure that scarce donor resources are wisely spent.

Against this background, this policy brief seeks to provide an overview of the state of impact assessment in the DDR community. It offers a critical evaluation of the metrics and methods currently used, highlights persistent challenges to impact assessment in volatile environments, and offers recommendations for more rigorous impact assessment of DDR programmes. Twenty-eight interviews were conducted with staff members of international organizations, including United Nations and World Bank personnel both at headquarters and in the field, DDR personnel from host-state programmes, as well as scholars working on DDR. A full, but partially-anonymized, list of interviewees is contained in an annex. While this policy brief will mainly focus on DDR programmes led by the United Nations Department of Peacekeeping Operations, programmes run or implemented by other institutions, such as the World Bank, the United Nations Development Programme (UNDP), or host-state-led programmes are also drawn on for comparative and illustrative purposes.<sup>6</sup>

The first part of the policy brief provides an overview of some of the most important assessments of DDR programmes and discusses important themes as well as empirical knowledge gaps in both the policy and scholarly literature. Substantiated with insights from DDR expert interviews, the policy brief subsequently discusses some persisting challenges in impact assessment of DDR programmes, including terminological inconsistency, a lack of prioritization and early integration of assessments into programming, budget constraints, inadequate methods and metrics, as well as ethical concerns with regard to data access and storage. The policy brief concludes by providing recommendations on practical steps that different actors in the DDR community can take to enhance the rigour and cost-effectiveness of assessments of DDR programmes.

## THE STATE OF THE ART

In both the policy and scholarly literature, attempts have been increasingly made to assess the impact of DDR programmes; a range of guidance documents have been published on this topic. Many of these assessments are geographically specific or focus on certain aspects of DDR programmes, such as the effectiveness of reinsertion benefits or vocational training, or DDR programmes' contribution to peace and stability.

In the following section, impact assessments of DDR programmes in both academia and programming will be delineated. This review is not exhaustive, but rather seeks to contextualize and analyse the state of the literature produced in this field and to highlight existing knowledge gaps.

### Policy literature

The policy literature on DDR programme assessments is characterized by mostly internal studies that are not available to the general public. For more detailed impact assessments or after-action reviews, organizations conducting DDR often make use of the expertise of scholars through cooperation with academic institutions.

The majority of publicly available papers are guidance documents, published by different organizations working in DDR. In light of the adoption of the UN's Integrated Disarmament, Demobilization, and Reintegration Standards (IDDRS) in 2006, the Department of Peacekeeping Operations (DPKO) has published an *Operational Guide to the IDDRS*,<sup>7</sup> as well as a *DDR Programme Management Toolkit*,<sup>8</sup> which includes "Standard Operating Procedures (SOPs) on M&E for DDR". Similarly, UNDP has published a *How To Guide: Monitoring and Evaluation for DDR Programmes*<sup>9</sup> and a *Practice Note on the Disarmament, Demobilization, and Reintegration of Ex-combatants*.<sup>10</sup> The World Bank, through its Transitional Demobilization and Reintegration Programme, has published country-specific M&E guidance for national DDR commissions of different African countries, including the *ICRS and M&E Manuals for the South Sudanese National DDR Programme 2013-2014 Pilot*.<sup>11</sup>

All of these documents provide an introduction to the basic principles of M&E in DDR programmes, highlight the similarities and differences between monitoring and evaluation, and provide guidance on how to develop an M&E strategy and framework, as well as adequate indicators. The documents also stress the need to plan for ex-post evaluations, which focus on the long-term success and sustainability of DDR, often only measurable after DDR programmes have been terminated.<sup>12</sup>

Whereas basic M&E exercises are increasingly conducted for reporting purposes, more comprehensive or even comparative impact assessments are only commissioned sporadically. This "paucity of impact assessment documentation

for the reintegration efforts stretching over the past decade"<sup>13</sup> has been criticized in particular by organizations working on the long-term reintegration component of DDR programmes. With regard to the methodologies utilized in these assessments, many evaluators use mixed-method approaches, mainly consisting of key informant interviews, benchmarks, and case studies.<sup>14</sup> Yet some researchers admit the limitations of their respective approaches, for example given biases that result from a limited accessibility of data or data being provided to the researchers by the same agencies that have also commissioned the assessments, often painting a very positive picture of the programme.<sup>15</sup>

The public availability of assessments depends on the institutional culture of the organization commissioning the evaluations. The World Bank, for example, through its Transitional Demobilization and Reintegration Programme's (TDRP) website,<sup>16</sup> provides access to some of its assessments.<sup>17</sup> The TDRP has also published on the long-term impact of some programmes — for example, on the long-term impact of post-independence DDR in Southern Africa.<sup>18</sup> Many assessments are time-bound and geographically focused on a specific programme, such as the *World Bank TDRP's Assessment of the 2013-2014 South Sudanese Reintegration Project Pilot*.<sup>19</sup> In other cases, assessments are carried out shortly after the end of programmes to demonstrate to donors that the money was appropriately spent.

Thorough assessment of a programme's long-term impact is challenging, given these impacts may only be measurable seven to twenty years after the programme's end. While they might shed light on certain aspects of DDR programmes, time-bound and geographically focused assessments conducted too soon after the end of DDR programmes are to the detriment of a comprehensive, longer-term analysis of a programme's complex impact.

In some cases, international organizations conducting DDR programmes are cooperating with researchers, providing them with access to their databases and facilities while hoping to benefit from their impact assessment expertise. For example, James Pugel, a research associate with the Centre for the Study of Civil War (CSCW) of the International Peace Research Institute (PRIO), was provided access to the DDR programme in Liberia for a UNDP-funded empirical study. He has also published articles on the reintegration and reconciliation efforts in Liberia in 2006 based on a nation-wide survey of ex-combatants.<sup>20</sup> Unlike previous assessment efforts mostly focused on the macro level, his country-wide, micro-level investigation sought to analyse the interim impact of the DDR programme, to highlight strengths and weaknesses of the programme, and to aid the direction of ongoing DDR efforts.<sup>21</sup> Based on interviews with 590 randomly selected, adult ex-combatants, Pugel concluded that ex-combatants were more successfully reintegrated with the help of DDR programmes.<sup>22</sup>

The Liberian case shows that scope, frequency, and even the findings of assessments may depend on who is commissioning the evaluations and what definitions they employ. This is particularly true in the case of vague terms like "reintegration". Based on its definition of reintegration, the *2007 Fifteenth Progress Report of the Secretary-General on the United Nations Mission in Liberia*<sup>23</sup> comes to quite a different conclusion than James Pugel, arguing that "the reintegration programme has failed to provide sustainable alternative livelihoods for ex-combatants."<sup>24</sup>

In addition, the commissioning of studies also greatly depends on donor interest and the institutional culture of evaluation in the respective organization. Organizations focused on the development aspect of DDR programmes, such as the World Bank and UNDP, seem to have more robust M&E systems, have more in-house M&E experts, and conduct evaluations more systematically. In just a few cases, cross-geographic assessments were conducted, such as in the case of the *2013 Evaluation of UNDP Reintegration Programmes*<sup>25</sup> and the *1993 World Bank Discussion Paper on Demobilization and Reintegration of Military Personnel*, a comparative assessment that drew evidence from seven DDR programmes and highlighted the need to integrate programme evaluation into each DDR programme.<sup>26</sup> Another comparative study was conducted by the World Bank's TDRP in cooperation with Randolph Rhea.<sup>27</sup>

### Academic literature

There is a growing body of academic literature on DDR programmes, coming mainly out of North American and European universities and think tanks. In general, it is important to note that scholars are operating on different timelines than policymakers who are often preoccupied with putting out fires or responding to budgetary reporting requirements. This enables scholars to ask different questions, to operate more flexibly, and to think about wider issues.

North American scholars, including academics such as Jeremy Weinstein, Macartan Humphreys, and Chris Blattman, are mostly producing quantitative studies, using randomized controlled trials and quasi-experimental designs. Their research focuses on early DDR experiences, such as Uganda, Sierra Leone, Liberia, and Burundi, and is centred around questions such as whether a correlation between DDR programmes and conflict incidence exists and whether ex-combatants have successfully reintegrated, focusing mainly on socio-economic reintegration.

Macartan Humphrey and Jeremy Weinstein, for example, examined the DDR programme in Sierra Leone, analysing which micro-level factors contributed to successful reintegration, thus complementing previous macro-level assessments on the contribution of peacekeeping operations to peace and stability.<sup>28</sup> Drawing on a survey of 1,043

combatants from the five warring factions in Sierra Leone's civil war and using propensity score matching between participants and non-participants of DDR programmes, they conclude that "past participation in abusive military factions is the strongest predictor of difficulty in achieving social reintegration (whereas) on economic and political reintegration, wealthier and more educated combatants face greater difficulties."<sup>29</sup> James Pugel's aforementioned study in Liberia, conducted in cooperation with UNDP, draws on Humphrey's and Weinstein's methodological approach and applies it to Liberia, concluding that ex-combatants who completed DDR programmes were significantly more likely to have a livelihood-producing activity, although no significant effects on poverty status were discernible.<sup>30</sup>

In contrast, some European scholars, including Jaremej McMullin, Lilli Banholzer, and Randolph Rhea, have focused more on political and sociological issues, such as gender, the reintegration of child soldiers, and political reintegration. Using qualitative methods, including key informant interviews, focus groups, and ethnographic studies, they generally claim less causal authority for their findings. Jaremej McMullin, for instance, conducted 200 key informant interviews in Namibia, Mozambique, Sierra Leone, and Liberia, finding that short-term assistance programmes often marginalize ex-combatants and reintegrate them into poverty, given limited economic opportunities in post-conflict countries.<sup>31</sup>

While academic impact assessments of DDR programmes are more focused on the long term, they too face methodological challenges. Pointing to the problem of spill-over effects, as well as selection and sampling biases, Macartan Humphreys and Jeremy Weinstein note that they were not able to unambiguously determine factors leading to a successful reintegration given limited possibilities to identify causal relationships. They call for more robust strategies for demonstrating the efficacy of demobilization and reintegration efforts. They underline that they are only using a second-best approach in their assessment, arguing that randomization would have provided enormous power for understanding the impact of external interventions.<sup>32</sup> In response to the above-mentioned problems encountered with selection biases, Michael Gilligan, Eric Mvukiyehe, and Cyrus Samii conducted quasi-experimental research on whether DDR programmes are effective in achieving their programmatic aims of economic rehabilitation, and if this translates into political and social reintegration, exploiting exogenous bureaucratic failures in service delivery.<sup>33</sup>

With regard to comparative studies of DDR programmes, a growing number of academics are investing in statistical assessments, drawing on large-n studies of past DDR programmes and increasingly using experimental designs.<sup>34</sup> Even though this body of work has refined assessment methodologies, these studies are mostly thematically focused, for example on the effect of vocational training

on income and employment<sup>35</sup> or the effectiveness of cash versus non-monetized incentives.<sup>36</sup> Other studies are geographically-focused on specific conflicts such as Burundi,<sup>37</sup> Liberia,<sup>38</sup> or Sierra Leone.<sup>39</sup> As Ann Bernes and Birger Heldt argue, “it remains unclear as to what extent insights derived from case studies are possible to generalize across the spectrum of DDR programmes,”<sup>40</sup> concluding that case studies need to be supplemented with “systematic evaluations that incorporate a large number of cases.”<sup>41</sup>

Several meta-studies were conducted by academics exploring the current state of the policy and scholarly impact assessment literature. In their meta-analysis of different DDR assessments over the past decades, Jonah Schulhofer-Wohl and Nicholas Sambanis analyse the contribution of DDR programmes to the overall objectives of peace and stability. They conclude that “no study to date has been able to isolate and measure the impact of DDR programmes on the peace processes.”<sup>42</sup> In another meta-analysis of impact assessments, James Pugel argues that a poor understanding of a theory of change, underlying causal connections, and clearly defined indicators is a clear shortcoming of impact assessments of DDR programmes.

Vague definitions, according to James Pugel, also have implications for the interpretations of the results of these assessments: “The absence of a clear DDR doctrine or widely agreed output and outcome indicators has contributed to ambiguity. The paradigmatic benchmark of DDR success – whether or not a state relapses into conflict – only takes us so far.”<sup>43</sup> He furthermore stresses that particular challenges are associated with impact evaluations in environments with ongoing hostilities and when excessive political pressure is exerted to demonstrate success. He concludes that the “UN struggles to inject more rigour into its programmatic interventions (and that) the impact evaluation literature must also seek to align itself more effectively to practitioner needs for relevance and ease of implementation.”<sup>44</sup>

### **Trends and persisting knowledge gaps**

While this review of the policy and scholarly literature is meant to be illustrative rather than exhaustive, commonalities and trends can still be distilled.

In the policy literature, it is easily discernible that even though M&E requirements are increasingly incorporated into DDR programmes and there is a growing critical awareness of the importance of assessments, assessment frequently consists of time-bound and geographically specific case studies with limited cross-geographic generalizability given differences in DDR mandates. More advanced assessments conducted in cooperation with academic institutions are few and far between. In addition, the scope, rigour, and public accessibility of these assessments depend on the institutional culture of the respective organization and donor interest. In many cases, the

assessments are limited to M&E exercises, conducted too soon after the programme to demonstrate real impact and focused on programme outputs for reporting purposes.

While methodological expertise and different timelines in academic impact assessments allow for more rigorous and longer-term studies, current assessments are limited geographically as well as by the methodological approach taken. A skilful combination of different methods and more cooperation between different actors in the DDR community will be necessary to overcome the constraints of a particular methodological approach and to thus enhance the rigour of impact assessment.

### **CHALLENGES**

Although “M&E” has become a buzzword in recent years, it was identified in one of the IDDRS modules as “one of the weakest areas of DDR programme management in the past, partly due to a lack of proper planning, a standardized M&E framework, and human and financial resources specifically dedicated to M&E.”<sup>45</sup> Without effective assessments, programming cannot be optimized to better meet the needs of beneficiaries. Additionally, weak M&E does not allow for an effective and targeted allocation of scarce donor money.

In the following sections, persistent conceptual and methodological challenges that inhibit effective M&E in DDR will be discussed.

#### **Impact assessment of what? Varying visions of the objectives of DDR**

Rigorous impact assessment is challenging. It requires a clear vision of the objectives of a programme and a sound understanding of the theory of change underlying programme design. It requires a rigorous approach that allows for the testing of the underlying assumptions and causal relationships that underpin DDR programmes. Yet in the case of DDR, two factors make coherent impact assessment particularly challenging.

First, DDR is typically a multiparty enterprise and, in the majority of the DDR programmes, different actors are executing different parts of the programme. In the case of Liberia, for example, the United Nations Mission in Liberia (UNMIL) was responsible for the disarmament and demobilization component of the programme, while UNDP was tasked with the rehabilitation and reintegration component, and a multitude of other organizations such as USAID, the World Food Programme, and the European Commission were involved in a donor or advisory capacity.<sup>46</sup> While the disarmament and demobilization aspects are thus classically the purview of UN peace operations actors, the reintegration part is frequently carried out by different implementing partners and does not fall within the scope of peace operation mandates.<sup>47</sup>

This might be problematic for several reasons: Programmers are confronted with the question whether it makes more sense to evaluate the different components separately or if an assessment of the whole DDR programme is advantageous. In addition, working across multiple units and entities — given inconsistency regarding definitions and priorities as well as varying budgeting cycles — may lead to difficulties. The disjointed nature of DDR programmes and their implementation across different bureaucratic and budgetary systems make it challenging to assess the complex and multidimensional impact of these programmes.

Second, the nature of DDR makes it in reality an umbrella term for highly context-specific programmes. Myriad activities fall within the scope of DDR programmes. In Haiti, for example, the DDR programme was essentially a violence reduction programme whereas in other DDR programmes the focus might be on the de-radicalization of combatants, such as for example in Somalia.

Depending on whether one is talking to security or development professionals, different answers will be given with regard to the objectives of DDR.<sup>48</sup> With increasingly context-tailored (yet vague) mandates, this divergence is likely to continue and therefore needs to be accounted for in assessments of DDR programmes. While it is, of course, necessary and beneficial that the recipients of these programmes receive programming that is as context-tailored as possible, the multitude of context-specific DDR programmes might make comparative assessments more challenging. Given these implications for the comparability of findings and knowledge development in general, the DDR community should use annual conferences and networks to work on common definitions and terminology.

#### **Barriers to early integration, prioritization, and systematic conduct of impact assessments**

Even though DDR is prominently featured in UN peace operations mandates, rigorous assessments are, in many cases, not systematically conducted and not integrated from the planning phases of DDR programmes.<sup>49</sup> Even though increasingly featured in guidance and mandates, “few people truly understand the necessity and importance of M&E.”<sup>50</sup> The outcome document of a March 2012 expert workshop on this subject, organized by the Geneva Centre for the Democratic Control of Armed Forces and the UN Department of Peacekeeping Operations, found that there is a “limited culture of evaluation of both Member States and the UN (...) and that the UN system lacks capacity to conduct impact assessments. Specialist expertise is required to design an assessment of the impact of a peacekeeping initiative.”<sup>51</sup>

In practice, M&E is often not recognized as a priority. Field missions are under-funded and faced with considerable pressure to respond to emergencies that arise in

the conflict environments they operate in. In crisis and post-conflict situations, “where the most urgent priorities are to get implementation up and running, M&E may consequently be overlooked or under-prioritized.”<sup>52</sup> Instead of an early integration and prioritization, impact assessments are generally only arranged for when accountability to donors or Member States is demanded,<sup>53</sup> and is often one of the first budget lines to be reduced in programme design, indicating that it is not viewed as a programmatic priority for senior management and in the donor community.<sup>54</sup> In order for programming to be designed based on triangulated, credible evidence rather than anecdotal evidence, a much heavier weighting must be given to M&E, which should be seen as adding more to the DDR community than purely information for programmers and donors regarding a specific programme. Consequently, there is not enough political pressure from donors, high-level political officers, and heads of missions to prioritize high-quality assessments and to change the institutional culture with regard to assessments, evoking negative consequences for donors and programme beneficiaries alike.

Rigorous impact assessment requires expertise within the programming institution.<sup>55</sup> First, sufficient expertise in M&E and impact assessment is required to design M&E systems and work plans for data collection, data analysis, and reporting.<sup>56</sup> Moreover, expertise enables practitioners to design a sound theory of change and to thus clearly articulate causal relationships and underlying conditions.<sup>57</sup> Finally, being conversant in impact assessment methods enables practitioners to determine whether the data they are using is robust, to use more advanced methods, and allows them to make choices about potential implementing partners, contractors, and external evaluators.<sup>58</sup>

While numerous guidance documents state that every DDR programme should have a dedicated DDR unit,<sup>59</sup> as one DDR expert at UN Headquarters emphasized: “at best, you have a M&E officer, an assistant, and a UN Volunteer, who jointly are responsible for data collection, the development of indicators, monitoring and evaluation, and the development of recommendations for the amelioration of the programme.”<sup>60</sup> Expertise in impact assessments therefore should be systematically deployed where needed, funded for, and enhanced through regular staff training.

The rigorousness of M&E efforts is, in many ways, proportionate to the resources invested in them, and the calls for more rigorous assessments are currently not carried into budgetary discussions. UN guidance documents, such as the UN Standard Operating Procedures, state that “the budget for M&E should normally be between 3% and 7% of the overall DDR budget.”<sup>61</sup>

However, as stressed in the outcome document of the aforementioned 2012 expert workshop, “there is no money earmarked for monitoring and evaluation (M&E) in

peacekeeping assessed budgets.<sup>62</sup> The document also underlines that the cost of M&E is negligible – only 1 per cent of the total programme budget according to the UN Evaluation Group – compared to the amount of money that is currently lost on peacekeeping activities that do not achieve their objectives. Significant funds can be saved by making activities “more relevant, effective, efficient, and focused on meeting stated objectives.”<sup>63</sup> In particular, assessments of longer-term reintegration processes are more burdensome, “requiring sustained commitment long after most donors had shifted to the next hot spot.”<sup>64</sup> As a consequence of underfunding and declining donor interest, M&E officers in some missions are not equipped to conduct longitudinal impact assessments.

The funding structure of DDR assessments has even wider implications on the way assessments are conducted. The results-based budgeting (RBB) system of the UN for example — with its short-term budget allocation and lack of a culture for multi-year programming — makes longitudinal assessments or assessment after peace operations exit the country difficult to conduct, as extra-budgetary funds for evaluations have to be acquired.<sup>65</sup> Longitudinal ex-post assessments of the long-term impact of DDR programmes, which is observable only years after the programme has ended, are therefore extremely difficult to conduct in the current budgeting framework. In addition, security constraints in the volatile conflict environments typically encountered in DDR can complicate thorough assessments, having implication for access, the collection of reliable baseline data against which to measure progress, and thus the ability to make the necessary modifications to programmes.<sup>66</sup>

### Methodological choices and challenges

Methodological challenges in impact assessment can arise with regard to the choice of metrics and methods, the chosen time frame for evaluations, and difficulties in measuring intangible impact or the effect of preventive programmes.

One issue frequently raised by the DDR experts interviewed concerns the question of whether a DDR programme’s impact is even measurable. This, of course, depends on the definition of output, outcome, and impact.<sup>67</sup> If impact is defined as the long-term effects and contributions to the declared objectives of DDR programmes, which often include references to “reducing the threat of armed groups (and making) space for stabilization activities”<sup>68</sup> or “creating the security conditions for ensuring sustainable economic development”<sup>69</sup>, then it is questionable, argue some experts, that an improved security situation can be attributed to the effects of a DDR programme alone. DDR experts sceptical of the possibilities to measure impacts advocate for measuring at least outcomes, instead of only outputs.<sup>70</sup> Even when you conclude that impact assessment is generally possible, persisting challenges include deciding on the “right” metrics that capture the complex impact

of DDR programmes, the inclusion of impact-related indicators, and the question of attribution versus contribution.

The choice of adequate methods and metrics depends on the subject of the research and is particularly challenging given the complexity of DDR programmes. Problems can arise when the evaluator gravitates towards easily measurable items, i.e., outputs, and neglects more complex aspects, such as recidivism rates and the effects of the incentive-based structure of DDR programmes on their effects.<sup>71</sup> Additionally, DDR programmes may even have the reverse effect by causing friction between the ex-combatants and the community, if the former are viewed by the latter as being the “perpetrators” who receive special assistance in comparison to the “victims” of the conflict. These context-specific dynamics, often based on subjective perceptions, as just discussed, are currently not incorporated enough in some DDR programme assessments.

Given the time pressures and security constraints, many assessments are limited to rudimentary evaluations carried out over a short time span and focused on assessing readily accessible and easily quantifiable metrics, such as the number of weapons collected. Yet measuring outputs<sup>72</sup> instead of focusing on outcomes<sup>73</sup> or impacts<sup>74</sup> can provide inaccurate impressions of an intervention’s efficacy.<sup>75</sup> Outcomes are more complicated because they are less quantifiable and require an agreement on the “intended systemic change that components of the programme are intended to produce”<sup>76</sup> and “measuring impact is a much harder proposition.”<sup>77</sup> Given the inaccurate results that a reliance on output-related metrics can yield,<sup>78</sup> adequately tailored metrics, more holistic approaches, and lengthened time frames are required in order to accurately measure a programme’s complex impact.

Yet even in cases where a change in condition is observable, can this change ever be unequivocally attributed to DDR programmes given myriad other factors at play? How can we know whether it was the effect of one specific project within the DDR programme, for example a specific vocational training, or a combination of programmes, programmes conducted by other organizations or different elements of multidimensional peacekeeping missions, or other confounding factors, such as an improving economy, that resulted in the impact observable?<sup>79</sup> Given that it is difficult to isolate the impact of DDR programmes from other processes and initiatives, “DDR programmes can never be seen as the sole bearer of peace, but only as a contributor”.<sup>80</sup> Instead of the claim of attribution — that an accomplishment was caused by DDR outputs or outcomes — it may be more realistic to demonstrate that DDR programmes “were contributory, amongst others, to an observed change with respect to the expected accomplishment”<sup>81</sup> (attribution vs. contribution debate). This is particularly true for the case of DDR programmes where preventive and often intangible effects, such as the

non-return to conflict, are the object of measurement.<sup>82</sup> In addition, given the impact of confounding and intermediary variables as well as knock-on effects, in complex and volatile conflict environments, this approach might yield a better picture of programmatic impact.<sup>83</sup>

### **Adequacy of assessment methods and ethical concerns**

The biggest divide between the academic and practitioner communities in the interviews related to the methods used in impact assessments of DDR programmes. Practitioners interviewed repeatedly underlined that methods used need to be “practical and pragmatic” and able to be rolled out on a larger scale in conflict contexts. In the 2012 DCAF/OROLSI Workshop paper, it was even highlighted that “Member States should accept that the most scientifically rigorous approaches are not always possible or desirable to implement in peace-keeping contexts.”<sup>84</sup>

The methods used by practitioners most frequently included perception surveys, focus group interviews, and livelihood mapping. This contrasted to the position common amongst academics who — in particular in the North American tradition — often call for more randomization and (quasi) experimental designs,<sup>85</sup> arguing that these methods allow for more robust claims about causal inference by systematically controlling for one or more independent variables and measuring the change in the dependent variable.<sup>86</sup> (Not all DDR experts agree on the value of experimental designs, however.<sup>87</sup>)

In particular, practitioners emphasize that withholding benefits from some communities is a politically sensitive issue, adding that one needs to be very cautious of the reputational effects these kinds of methods could evoke for the DDR programme. Increased criticism regarding a temporary denial of programme benefits can, for example, diminish a programme’s licence to operate and have negative repercussions for the entire DDR programme. Host states and implementing partners on the ground will thus often refuse to endorse these methods.<sup>88</sup>

A broader consensus existed among academics and practitioners regarding three points: First, the need to address any ethical concerns through the use of safeguards. Second, the need to use a mix of different qualitative and quantitative methods to help overcome the constraints of specific methods and achieve a more accurate and comprehensive picture of an intervention’s complex impact.<sup>89</sup> Third, the need to adapt methodological considerations to the specific context and changes within it.<sup>90</sup>

### **Issues of access, storage, and inter-agency cooperation**

This section will outline issues of database management, such as in the case of the DREAM database, as well as storage and privacy challenges associated with data collection. It will also discuss possibilities provided by modern

communication technologies and implications of their use for inter-agency cooperation.

At the centre of data collection and analysis is the case management system. The database commonly used in UN DDR programmes is DREAM, a generic DDR software for Disarmament, Demobilization, Reintegration, and SALW Control provided by UNDP.<sup>91</sup> This database “supports the initial registration of combatants, transitional payments and benefits, small arms and light weapons management and distribution, and the support to reintegration through training and work projects for each registered participant.”<sup>92</sup>

In interviews with different DDR experts, DREAM was frequently criticized for not being flexible enough to adapt to the specific context on the ground. Some interviewees debated the benefits and drawbacks of having a central, real-time DDR database to reduce the need to call the different field missions for information.<sup>93</sup> While the need for more innovation, more flexible databases, and the use of mobile technologies in data collection and analysis was reiterated, most practitioners agreed that no matter what innovations are made, they need to be placed in a sound methodological framework and proper safeguards need to be established.

With respect to rigorous information collection and management, mobile communication technologies can be a vital tool in data collection and analysis. Dr Vanda Felbab-Brown, for example, argued that non-state armed actors, like the Taliban or al-Shabaab, are very effectively using modern technologies — for example, to intimidate former combatants via phones.<sup>94</sup> Equally, mobile technologies could be more effectively used to report how many ex-combatants are still being contacted by their former commanders, whether they are still receiving their reintegration benefits, or if they still have their jobs.

However, there are also ethical concerns and privacy issues associated with the collection and storage of politically highly sensitive data with names, detailed information, and biometric data of former members of armed groups. These ex-combatants are often under significant threat, because armed groups might seek revenge or, in other cases, the national government might be interested in obtaining their personal data. Therefore, safeguards need to be implemented to protect the privacy and identity of individuals who take part in DDR programmes. When cooperating with external actors, for example, some agencies resort to sharing only anonymized, aggregated data to protect the privacy of the programme beneficiaries.<sup>95</sup> Further ethical concerns might arise with the use of mobile technologies in data collection and long-term tracking exercises. When surveying the long-term reintegration of ex-combatants with the help of phones, for example, ex-combatants might hesitate to agree to a long-term tracking of their activities.

What kind of implications do these issues of access and storage have for inter-agency cooperation? There was a consensus among the DDR experts interviewed that sensitivities need to be addressed, for example by providing different actors with varying levels of access when co-hosting the DDR Information System.<sup>96</sup> A couple of interviewees even conceded that in some cases the reason why data is not shared is not the sensitivity of the data, but due to political reasons (i.e., that faced with budgetary constraints, people want to demonstrate ownership over “their” data and its role in assessing impact and are therefore reluctant to share it). These “turf wars”, however, can produce sub-optimal approaches where political interests limit data sharing and create redundancies.

In general, there seems to be only limited data sharing, in particular of primary data, while analyses, such as economic livelihood mappings, might be shared between institutions with the right safeguards in place.<sup>97</sup> In order to allow for a good communication between academics and practitioners, they have to overcome conceptual and terminological opacity, speak the same language, and understand each other’s priorities, timelines, and limitations. Yet, Pugel argues, “problematically, academics and practitioners have yet to agree on a common definition or approach to reintegration and the interconnections between separate components. Without a coherent conceptual starting point to operationalize reintegration it will continue to be exceedingly difficult to measure.”<sup>98</sup>

### **Reporting and feeding lessons learned back into programme design**

High-quality data collection needs to be matched with rigorous data analysis and reporting. In this context, challenges persist with regard to how information is fed back into programming and the lack of latitude to report failures. The information produced currently in some cases is not absorbed by the donors and implementing partners and, subsequently, best practices and lessons learned are not fed back into programme design, so that valuable lessons are wasted. This can have negative implications when having to demonstrate accountability to donors and may also limit programmers’ ability to refine and adapt programmes for the beneficiaries where necessary. Lessons learned, one of the key functions identified in the IDDRS, can consequently not be identified and disseminated and, in addition, the institutional memory of an organization, which is particularly helpful given the fast staff turnover in the field, is often lacking.<sup>99</sup>

Drawing robust lessons learned from assessments of DDR programmes requires that not only positive but also negative lessons are learned. Yet the “UN’s weak culture of programme design and planning further compounds the capacity challenges.”<sup>100</sup> In many of the organizations conducting DDR programmes, there is currently no culture of reporting failure — or, even worse, an “institutional and

donor bias against honest reporting”<sup>101</sup> because highlighting where programmes have not achieved the desired objectives could damage career prospects.<sup>102</sup>

Another challenge pertains to the reporting of budgets used for DDR programmes. When programmes use less money than expected, they are not rewarded; instead, the budget is adjusted and they risk receiving less money to spend in the next budget cycle. Programmers are therefore under pressure to spend every penny, which is neither effective nor efficient, and may have distorting effects on an assessment’s accuracy of a DDR programme’s effectiveness. As a consequence, the institutional culture of organizations conducting DDR programmes will have to change to enhance the effectiveness and efficiency of these programmes and to improve the ability to draw accurate lessons from existing DDR programmes.

## **IMPROVEMENTS**

This section seeks to provide donors and senior management of international organizations with practical recommendations to enhance the quality, cost-effectiveness, and utility of impact assessments on DDR programming.

### **Expertise, training, and inter-agency cooperation**

Impact assessments need to be systematized, budgeted for, and built into programming from the conceptualization phase of DDR programmes. First and foremost, this necessitates practitioners being sufficiently conversant in research methods to conduct assessments in non-permissive environments and having a knowledge base in research methodology to choose adequate metrics and methods for the respective evaluation.

High-quality assessments based on empirics-led iterative programming cycles should be prioritized and valued by programmers and donors alike. A focus on adequate expertise will also enhance the rigour of assessments, as practitioners will be able to use more advanced methods, to more accurately and efficiently measure causal relationships, and to determine the richness of the data — for example, when using proxy indicators in cases where direct measurement is not possible. Practitioners conversant in advanced methods may also be in a better position to mitigate ethical concerns early: e.g., in cases where a fully random sampling or a division into treatment and control groups could evoke criticism based on ethical arguments, alternatives, such as stepped-wedge designs (which consist of staggered treatments over multiple time periods), can be chosen.<sup>103</sup> This can only happen when rigorous assessments are integrated from the planning phase of programmes and solid baseline data is collected, against which to later measure programmatic progress. Bringing in evaluators towards the end of programmes when accountability to donors is required is too late to conduct comprehensive

assessments and to produce an accurate picture of a programme's progress. These demands for more systematization and an early integration of assessments beyond ad hoc approaches all point to a need for adequate resources and political will.

Practical recommendations proposed by the DDR experts interviewed in this regard included the establishment of an M&E focal point, an M&E roster, or even a measurement section within DPKO to ensure an early integration, prioritization, and systematization of assessment efforts.<sup>104</sup> Several DDR experts proposed yearly meetings not only of the DDR chiefs but also for M&E officers, and more field-friendly guidance documents, such as a one-pager describing key concepts (e.g., the theory of change).<sup>105</sup>

Randoph Rhea highlighted the need for more stringent checks of M&E officer qualifications, for example through certifications provided by the Integrated DDR Training Group.<sup>106</sup> Donors and Member States should invest in tailored training programmes for practitioners carrying out evaluations to create a common knowledge base. This would allow for more high-quality assessments and more standardization of evaluations across programmes, and would provide DDR professionals with a background in research methods that they can also put to use in other peacebuilding positions. It will significantly enhance their ability to identify flaws in ill-conceived assessment, use advanced assessment tool, choose adequate external consultants, and make use of the latest technological innovations.

Evaluating DDR programmes can happen in two ways: First, through the enhancement of in-house capacity, for example in the form of an evaluation section within DPKO and DDR units consisting of experts with advanced methodological training. Second, through outsourcing assessments to external evaluators, which might be more expensive but has the added benefit of reducing biases and a conflict of interest.

With regard to cooperation with outside actors, collaboration with academic institutions can help to integrate longitudinal assessments from the beginning. A mutually beneficial cooperation should be incentivized, for example by providing scholars with long-term access to the field and programme data and by scholars sharing their methodological expertise and mentoring DDR practitioners on rigorous impact assessment. This would strengthen the quality of impact assessments (relating to robustness of research and independence of researcher), build capacity within practitioners, and reduce some of the burden on the programme team.<sup>107</sup>

Yet this kind of cooperation also requires that there is a better understanding of the language, needs, and timelines, which are currently very different in the academic and policy world. Donors and Member States should invest in

an annual forum of DDR scholars and practitioners to work towards common definitions and foster a culture of M&E. Currently the DDR Working Groups serve as a forum for policymakers, yet DDR field officers and scholars should be equally involved to enhance a mutual understanding of each other's language, needs, and timelines, which would allow for greater cohesion, comparability, and cross-fertilization of best practices between different entities. Incentivized cooperation between different actors in the DDR community — not only between programmers and scholars, but also with donors, contractors, and implementing partners — will allow for greater cohesion, comparability, and cross-fertilization of best practices between different entities.<sup>108</sup>

### **Effectively using technologies and enhancing innovation**

The volatile, non-permissive environments in which DDR programmes operate often require great flexibility and improvisation on the side of the programmers and programmes, which need to be flexible enough to be adapted to context-specific needs. In order to account for this flexibility in programming and to conduct more accurate, holistic assessments, flexibility, innovation, and a constant readjustment are also required with regard to the methods and tools utilized in impact assessments.

This also applies to the use of databases and modern communication technologies, in particular in non-permissive environments, where the use of mobile technologies becomes increasingly important. Many of the DDR experts interviewed underscored the need to be able to adapt information systems and databases like DREAM to the context-specific needs of each DDR programme. Given the fragmented nature of DDR programmes, it is necessary to enhance the interoperability of such databases for different users and to incorporate options to change and add context-specific variables, where required.<sup>109</sup>

In the Central African Republic in 2011, money was invested in the open-source multi-purpose software KoboToolbox, previously used in humanitarian assistance programmes, which allowed for more mobility and flexibility in data collection and analysis.<sup>110</sup> Moreover, mobile field units were used in the Central African Republic to verify more than 6,400 and disarm more than 5,000 ex-combatants at 21 sites in the different provinces by December 2011.<sup>111</sup> These highly mobile stations enhanced the speed, independence, and flexibility to react to new developments.

In addition, modern communication technologies have put more accurate and less expensive measurement<sup>112</sup> within reach and have potential utility in improving access to a larger number of beneficiaries.

Yet, while access to rural areas is improved and cost-effectiveness enhanced, biases and ethical concerns may arise and need to be dealt with: If not integrated in a sound

methodology, and if no robust baseline data is available, mobile technologies in both data collection and analysis can introduce biases — for example, by administering a written perception survey in areas with high illiteracy rates. Moreover, ethical concerns may be voiced when it comes to surveying the long-term reintegration of ex-combatants with the help of mobile technologies, for examples.<sup>113</sup> For these reasons, the use of modern information and communication technologies needs to be administered by a trained expert and placed duly under a “strict political and legal framework [that addresses] control and confidentiality concerns.”<sup>114</sup>

### **Prioritization of longitudinal assessments**

Longitudinal studies, focused on a programme’s complex impact, need to be viewed as an integral part of programme planning and be prioritized by senior management and donors alike. Instead of focusing on output-related indicators, which might have value for reporting purposes but might on occasion provide a highly inaccurate picture of an intervention’s impact, longitudinal studies will enable more accurate measurement of an intervention’s complex impact.<sup>115</sup> This can significantly enhance the efficiency of donor spending, identify and respond to harmful practices, and thus help tailor programmes for the benefit of programme participants.

Real commitment to rigorous measurement from the highest level — beyond M&E for simple reporting and auditing purposes — will also enhance the accountability for programmes, for example to the Special Committee on Peacekeeping Operations (C34).<sup>116</sup> This committee, reporting to the UN General Assembly, is tasked with reviewing all issues relating to peacekeeping, and thus has an interest in receiving the most accurate and comprehensive assessments of DDR programmes possible. Given the lack of a culture for multi-year programming, the onus is also on donors to invest in longitudinal studies and to prioritize assessments after the end of a peace operation. Senior managers must realize that we need to look beyond the life of the programme. For that reason, longitudinal assessments, beyond the output-related assessments for auditing purposes, need to be budgeted for, built into mandates, and prioritized by senior management.

There are potentially major cost savings available to donors through investment in rigorous assessments on the (long-term) effectiveness of DDR programmes. Yet prioritizing longitudinal impact assessment would also imply that the budgetary system would have to be reviewed — or reliable and predictable work-arounds found — because the short-term budget cycles of DPKO are counterproductive for long-term assessments.<sup>117</sup>

To allow for more longitudinal assessments, mutually beneficial cooperation with universities should be explored by donors and programme managers. Scholars could be granted more and longer access to DDR programmes to

conduct field research in DDR programmes in exchange sharing their methodological expertise with the local M&E team and mentoring DDR practitioners on rigorous impact assessment. Universities could be encouraged to find matching funds to those supplied by the institutional implementing partner.

Yet all efforts to make impact assessment more rigorous will be to no avail if the results of both programmatic and academic assessments are not fed back into programming. In this regard, there also needs to be a better management of expectations (with Member States, donors, and host states alike) of what DDR programmes can accomplish.

### **Fostering a culture of honest assessments and lessons learned**

Identifying and documenting lessons learned in order to feed these back into the programming cycle and to adapt programmes to better serve the needs of programme beneficiaries is a central element of impact assessment.<sup>118</sup> Yet too often the approach to assessment involves one-off studies conducted when problems arise instead of regular, comprehensive assessments that incorporate findings back into programming. Rather than adopting a bureaucratic survival mode and only conducting evaluations when donors demand it, systematic and honest evaluations should be demanded by donors and senior management. Iterative feedback loops, whereby impact assessment can drive programmatic and policy discussions and innovation, but also the dissemination of lessons learned and the institutionalization of best practices, should be encouraged to improve programming.<sup>119</sup>

In order to ensure that the right lessons are learned, the “limited culture of evaluation of both Member States and the UN”<sup>120</sup> needs to change. Both senior management and donors “need to be clear on their willingness to foster a learning story (and not just tell a ‘good story’)”<sup>121</sup> thus allowing for “more latitude to report failures.”<sup>122</sup> Programmes should not be defunded as a consequence of honest reporting, and highlighting failures should not end careers.

Confronting one’s own culture and allowing for, even demanding, honest assessments is vital, even if these assessments reveal inconvenient truths. Only if solutions to unflattering results are found and worked back into programming can necessary modifications be made that will enhance both the rigour and cost-effectiveness of DDR programmes and beneficiaries’ experiences with DDR programmes.

## CONCLUSION

Impact assessments are a challenging yet essential endeavour to evaluate the effectiveness of DDR programmes and to allow programme modifications for the benefit of beneficiary communities and financially-strapped donors alike. The challenges identified in the second part of this policy brief demonstrate that impact assessments conducted in the DDR community have yet to reach their full potential. More often than not, impact assessments are not integrated into programming from the planning phase, and are not prioritized by senior management and donors.

Short-term budget cycles and the lack of a culture of multi-year programming hinder the production of assessments that are able to accurately portray the complex impact of DDR programmes. Instead, assessments currently focus on output-related measures and are frequently only conducted when accountability to donors and Member States is demanded. Practitioners are often not sufficiently conversant in assessment methods to be able to adequately address possible methodological challenges and ethical concerns. These persisting problems continue to limit the quality and effectiveness of impact assessments.

In order to maximize DDR programmes' impact for the beneficiaries for whom they are conceived, senior management and donors of organizations conducting DDR need to encourage reforms to avoid unintended negative consequences of ill-conceived impact assessments. Not only should rigorous assessments be prioritized; they also need to be integrated from the planning phase and budgeted for. Adequate training opportunities for practitioners need to be created, and an emphasis should be placed on increased inter-agency cooperation to eliminate redundancies and enhance a mutually beneficial cooperation between different actors in the DDR community.

Yet all of these recommendations will be to no avail if problems and lessons identified are not fed back into programming and no latitude to report failures exists. Incorporating these recommendations into DDR programming, therefore, will significantly enhance the quality, cost-effectiveness, and utility of impact assessments of DDR programmes.

## ANNEX – LIST OF INTERVIEWEES

1. Robert Muggah
2. Vanda Felbab-Brown
3. Joshua Shuajo Mitrotti Ventura
4. Aki Stavrou
5. Alexandra Burrell Jung
6. Glaucia Boyer
7. Randolph Rhea
8. Jaremey McMullin
9. Jacopo Monzini
10. Nick Imboden
11. Tino Kreutzer
12. Wolf-Christian Paes
13. Joanna Richards
14. Dean Piedmont
15. Peter Romaniuk
16. Richard Bowd
17. Ingvild Gjelsvik
18. Senior DDR official, DPKO
19. Senior DDR official, DPKO
20. Senior DDR official, DPKO
21. DDR official, DPKO
22. DDR official, DPKO
23. DDR official, DPKO
24. DDR expert, UN WOMEN
25. DDR expert, UNDP
26. DDR expert, DPKO
27. DDR expert on Afghanistan
28. Independent DDR consultant

## ENDNOTES

1. According to the Secretary-General's note to the General Assembly (A/C.5/59/31) from May 2005, DDR may be understood as follows: "Disarmament is the collection, documentation, control and disposal of small arms, ammunition, explosives and light and heavy weapons of combatants and often also of the civilian population. (...) Demobilization is the formal and controlled discharge of active combatants from armed forces or other armed groups. (...) Reinsertion is the assistance offered to ex-combatants during demobilization but prior to the longer-term process of reintegration. (...) Reintegration is the process by which ex-combatants acquire civilian status and gain sustainable employment and income."
2. United Nations University, "UN DDR in an era of violent extremism", available from <http://collections.unu.edu/view/UNU:5532#viewMetadata> (accessed 4 October 2015).
3. See United Nations, "Operational Guide to the Integrated Disarmament, Demobilization, and Reintegration Standards", 2014, available from <http://www.unddr.org/uploads/documents/Operational%20Guide.pdf> (accessed 6 October 2015), p. 18.
4. United Nations, "3.50 Monitoring and Evaluation of DDR Programmes", IDDRS Module, Level 3: Structures and Processes, available from <http://www.iddrtg.org/wp-content/uploads/2013/05/IDDRS-3.50-Monitoring-and-Evaluation-of-DDR-Programmes1.pdf> (accessed 6 October 2015).
5. In this paper, impact assessment is defined as assessments that seek to measure the long-term effectiveness of programmes. In practice, the term impact assessment is often used interchangeably with the term Monitoring and Evaluation (M&E). The term M&E refers to assessments of how "various factors contribute to or detract from the achievement of the proposed outcomes and programme objectives, and measuring the effectiveness of outputs (...) M&E plays an important role in ensuring constant quality control of activities and processes and it also provides a mechanism for periodic evaluations of performance in order to adapt strategies (...) Monitoring is the systematic oversight of the implementation of an activity that establishes whether input deliveries, work schedules, other required actions and targeted outputs have proceeded according to plan, so that timely action can be taken to correct deficiencies. Evaluation is a timespecific activity that systematically and objectively assesses the relevance, performance and success of ongoing and completed programmes and projects." (see United Nations, "3.50 Monitoring and Evaluation of DDR Programmes", IDDRS Module, Level 3: Structures and Processes). Therefore, impact assessment is more closely related to the evaluation aspect of M&E.
6. Even though occasional references occur, this policy brief will not focus on US or NATO approaches to DDR, as for example employed in Afghanistan.
7. United Nations, "Operational Guide to the Integrated Disarmament, Demobilization, and Reintegration Standards".
8. United Nations DPKO, "DDR Programme Management Toolkit", available from [http://www.unddr.org/documnets/ddr-programme-management-toolkit\\_211.aspx](http://www.unddr.org/documnets/ddr-programme-management-toolkit_211.aspx) (accessed 6 October 2015).
9. UNDP, "How to Guide: Monitoring and Evaluation for Disarmament, Demobilization, and Reintegration Programmes", available from [http://www.undp.org/content/dam/undp/documents/cpr/documents/ddr/ddr\\_how\\_to\\_guide.pdf](http://www.undp.org/content/dam/undp/documents/cpr/documents/ddr/ddr_how_to_guide.pdf) (accessed 6 October 2015).
10. UNDP, "Practice Note: Disarmament, Demobilization, and Reintegration", available from [http://www.undp.org/content/dam/undp/documents/cpr/documents/ddr/DDR\\_Practice\\_Note\\_English\\_PDF.pdf](http://www.undp.org/content/dam/undp/documents/cpr/documents/ddr/DDR_Practice_Note_English_PDF.pdf) (accessed 6 October 2015).
11. Transitional Demobilization and Reintegration Programme, "ICRS and M&E Manuals", Republic of South Sudan National DDR Programme 2013-2014 Pilot Reintegration Project, available from [http://www.tdrp.net/PDFs/SouthSudanreportBook4\\_Complete.pdf](http://www.tdrp.net/PDFs/SouthSudanreportBook4_Complete.pdf) (accessed 6 October 2015).
12. DDR programmes provide a package of services. M&E typically focuses on the period during the delivery of the services, and rarely the subsequent period. Lessons from the field of criminal justice suggest that longer-term monitoring and evaluation be useful.
13. UNDP, "Sharing new ground in post-conflict situations: The role of UNDP in support of reintegration programmes", January 2000, available from [http://web.undp.org/evaluation/documents/postconflict\\_march2000.pdf](http://web.undp.org/evaluation/documents/postconflict_march2000.pdf) (accessed 6 October 2015).
14. See, for example, Abhijit Bhattacharjee and Hassan Ali Gadkarim, "Evaluation of Sudan DDR Programme, UNDP Sudan, Final Report, November 2012", available from <http://www.smallarmssurveysudan.org/fileadmin/docs/facts-figures/ddr/UNDP-Sudan-DDR-Evaluation-Nov-2012.pdf> (accessed 6 October 2015).
15. Ibid.
16. TDRP, "South Sudan DDR Pilot Reintegration Pilot Project Findings, Tools, and Lessons for Future Use", available from <http://www.tdrp.net/SouthSudanReintegrationPilotDocs.php> (accessed 6 October 2015).
17. In the case of the South Sudanese Pilot DDR Programme, for instance, these documents include the final evaluation of the programme, satisfaction surveys of ex-combatants, lessons learned, and tools for future use.
18. Guy Lamb, "DDR 20 Years Later – Historical Review of the long-term impact of the post-independence DDR in

- Southern Africa”, available from <http://www.tdrp.net/PDFs/DDR20YearsLater.pdf> (accessed 6 October 2015).
19. This document highlights transferable lessons. However, it is not only narrow in its scope but was also, like many other assessments, written too soon after the end of the project to draw meaningful conclusions about the programme’s long-term impact.
  20. James Pugel, “What fighters say: A survey of ex-combatants in Liberia, February-March 2006”, available from [http://www.operationspaix.net/DATA/DOCUMENT/904~v~What\\_the\\_Fighters\\_Say\\_\\_A\\_Survey\\_of\\_Ex-combatants\\_in\\_Liberia.pdf](http://www.operationspaix.net/DATA/DOCUMENT/904~v~What_the_Fighters_Say__A_Survey_of_Ex-combatants_in_Liberia.pdf) (accessed 6 October 2015).
  21. *Ibid*, p. 2.
  22. James Pugel draws conclusions about successful reintegration based on survey questions about self-reported and perceived acceptance, ex-combatants’ social networks, acceptance of conflict resolution mechanisms, and home community resettlement (see James Pugel, “What fighters say”, p. 5).
  23. “Fifteenth Progress Report of the Secretary-General on the United Nations Mission in Liberia”, UN Doc. S/2007/479, 8 August 2007. Further external evaluations on the DDR programme in Liberia were commissioned, inter alia, by UN DPKO, USAID, UNDP, UNICEF, and the European Commission.
  24. “Fifteenth Progress Report”, *op. cit.* p. 7.
  25. UNDP, “Evaluation of UNDP Reintegration Programmes”, available from <http://erc.undp.org/evaluationadmin/managementevaluation/viewevaluationdetail.html?evalid=6162> (accessed 6 October 2015).
  26. This study was written soon after the start of the programmes and, therefore, it was not possible to fully evaluate the programmes’ impact. In their assessment, Jonah Schulhofer-Wohl and Nicholas Sambanis highlight that numbers such as a 40 per cent dropout rate or a 28 per cent employment rate give little indication of the programmes’ real impact, because it is difficult to make claims from individual-level outcomes to broader policy level outcomes, which would also require the existence of a control group: see World Bank, “Discussion Paper on demobilisation and reintegration (D&R) of military personnel”, 1993, available from <http://siteresources.worldbank.org/INTCPR/882635-1120216203830/20626336/IDP0130.pdf> (accessed 6 October 2015).
  27. The *2014 Comparative Study of Ex-Combatant Reintegration in the Great Lakes Region* explores the dynamics and results of ex-combatant reintegration in Uganda, Rwanda, the Democratic Republic of Congo, the Republic of Congo, and Burundi.
  28. Macartan Humphreys and Jeremy Weinstein, “Demobilization and Reintegration”, in *Journal of Conflict Resolution*, Volume 51, Number 4 (2007), pp. 531-566.
  29. Macartan Humphreys’ and Jeremy Weinstein’s findings did not show a discernible effect of DDR programme participation on economic and political reintegration; see Macartan Humphreys and Jeremy Weinstein, “Demobilization and Reintegration”, in *Journal of Conflict Resolution*, Volume 51, Number 4, p. 531.
  30. James Pugel, “What fighters say: A survey of ex-combatants in Liberia, February-March 2006”, p. 2.
  31. Jaremej McMullin, *Ex-combatants and the post-conflict state – challenges to reintegration* (London, Palgrave Publications, 2013).
  32. Macartan Humphreys and Jeremy Weinstein, “Demobilization and Reintegration”, p. 533.
  33. Original survey data was used to study ex-combatant reintegration after Burundi’s 1993–2004 civil war. Their findings show a 20-35 per cent poverty reduction, moderate improvement in livelihoods, but only a “modest increase in propensities to report civilian life as preferable to combatant life, (...) satisfaction with the peace process or a more positive description toward current government institutions.”; see Michael Gilligan, Eric Mvukiyehe, Cyrus Samii, “Reintegrating Rebels into Civilian Life: Quasi-Experimental Evidence from Burundi”, *Journal of Conflict Resolution*, Volume 57, Number 4, (2010), pp. 598-626.
  34. See, for example, ECP, *Analysis of the Disarmament, Demobilization and Reintegration Programmes Existing in the World During 2006* (Madrid, Escola de Cultura de Paz, 2007); Gilligan et al., “Reintegrating Rebels”; Yuhki Tajima, *Understanding the Livelihoods of Former Insurgents: Aceh; Indonesia* (Washington DC, World Bank, 2010).
  35. Ian Lively, “Measuring intermediate outcomes of Liberia’s disarmament, demobilization, rehabilitation, and reintegration programme”, *Defence and Peace Economics*, Volume 25, Number 2 (2014), pp. 139-162.
  36. See, for example, Alpaslam Özerdem, *Post-war Recovery: Disarmament, Demobilization and Reintegration* (London, I.B. Tauris, 2008); Sigrid Willibald, “Does Money Work? Cash Transfers to Ex-Combatants in Disarmament, Demobilization and Reintegration Processes”, *Journal of Disasters*, Volume 30, Number 3 (2006), pp. 316-339.
  37. Michael Gilligan, Eric Mvukiyehe, Cyrus Samii, “Reintegrating Rebels into Civilian Life: Quasi-Experimental Evidence from Burundi”; Yuhki Tajima, *Understanding the Livelihoods of Former Insurgents: Aceh; Indonesia* (Washington DC, World Bank, 2010); Pyt Douma, *Reintegration in Burundi: between happy cows and lost investments*, (Den Haag, Clingendael Institute, 2009).
  38. Pugel, “What fighters say”.
  39. Humphreys and Weinstein, “Demobilization and Reintegration”; Kieran Mitton, “Engaging disengagement: the political reintegration of Sierra Leone’s Revolutionary United Front”, *Conflict, Security and Development*, Volume 8,

- Number 2 (2008), pp. 193-222.
40. Ann Bernes and Birger Heldt, "Preface", in Jonah Schulhofer-Wohl and Nicholas Sambanis, *Disarmament, Demobilization, and Reintegration Programmes – An Assessment*, Folke Bernadotte Academy Publications, available from [https://fba.se/contentassets/7da0c74e1d22462db487955f2e373567/ddr\\_programs\\_-an-assessment.pdf](https://fba.se/contentassets/7da0c74e1d22462db487955f2e373567/ddr_programs_-an-assessment.pdf) (accessed 6 October 2015), p. IX.
  41. Ibid.
  42. The authors' primary criticism is the failure to isolate the causality of DDR programming from other variables; see Jonah Schulhofer-Wohl and Nicholas Sambanis, *Disarmament, Demobilization, and Reintegration Programmes – An Assessment*, p. 2.
  43. James Pugel, "Measuring Reintegration in Liberia: Assessing the Gap Between Outputs and Outcomes", in Robert Muggah (ed.) *Security and Post-Conflict Reconstruction: Dealing With Fighters In the Aftermath of War* (London: Routledge, 2008), p. 70.
  44. Ibid, p. 77.
  45. United Nations, "3.50 Monitoring and Evaluation of DDR Programmes".
  46. Escola de Cultura de Pau, "Liberia", available from <http://escolapau.uab.cat/img/programas/desarme/mapa/liberiai.pdf> (accessed 6 October 2015).
  47. Randolph Rhea highlighted the role that institutional culture and mandates play in the design and execution of DDR M&E programmes. The World Bank for example is focused on establishing parity between combatants and their receiving communities (in order to get to a position where a development loan can be given). See also Interviews with Senior DDR Officer in DPKO and Dr Jaremy McMullin.
  48. This inconsonance is also reflected in the answers that interviewees provided to this questions: While security experts highlighted "conflict reduction" as the main objective of DDR, development professionals emphasized the need to "create a parity between ex-combatants and communities through reintegration". See further UN DPKO, "DDR in peace operations – a retrospective", available from [http://www.un.org/en/peacekeeping/documents/DDR\\_retrospective.pdf](http://www.un.org/en/peacekeeping/documents/DDR_retrospective.pdf) (accessed 6 October 2015), p. 2.
  49. See United Nations DPKO, "DDR Programme Management Toolkit", p. 3.
  50. Interview with senior DDR official in DPKO.
  51. DCAF and UN DPKO OROLSI, "Measuring the Impact of Peacekeeping Missions on Rule of Law and Security Institutions", Report of the Expert Workshop, convened on 12 March 2012, Permanent Mission of Switzerland to the UN, available from <http://www.dcaf.ch/Publications/Measuring-the-Impact-of-Peacekeeping-Missions-on-Rule-of-Law-and-Security-Institutions> (accessed 6 October 2015), p. 3.
  52. Several of the experts interviewed highlighted a disconnect between the field personnel and headquarters staff in this regard, stressing that headquarters staff needs to better understand the "already persisting challenges in the field," where programmes are set up under time pressure in conflict contexts where a thorough collection of pre-intervention baseline data is not feasible.
  53. Dr Richard Bowd underlined that robust assessments with the aim of providing real time data to enable effective programme adaptation are often not the focus of M&E. Rather, M&E is undertaken to gather data for donor reporting and compliance.
  54. Interview with Dr Richard Bowd.
  55. IDDRS guidance states that "all integrated DDR sections should make provision for the necessary staff, equipment and other requirements to ensure that M&E is adequately dealt with and carried out, independently of other DDR activities, using resources that are specifically allocated to this purpose."; see United Nations, "3.50 Monitoring and Evaluation of DDR Programmes", p. 2.
  56. Ibid, p. 1.
  57. One senior UN DDR official stressed that too few practitioners are familiar with designing theories of change. Moreover, one of the crucial steps highlighted in the UN DDR Programme Management Toolkit is that DDR planners need to be clear about the DDR results chain before they can determine SMART (Specific, Measurable, Achievable, Realistic, and Time-Bound) indicators and define capacities required for M&E.
  58. This is crucial even in cases where the evaluations are fully outsourced to external experts, which may be regarded as being "more objective and less prone to bias, and thus have more credibility" (UNDP, "How to Guide Monitoring and Evaluation for Disarmament, Demobilization, and Reintegration Programmes", p. 56). The UNDP "How to Guide Monitoring and Evaluation for DDR Programmes" highlights that the best choice is to "combine international evaluators with broad experience in DDR and evaluations, with national experts who have a better grasp of the local context and are better able to communicate with beneficiaries and partners" (see *ibid*, p. 57). Similarly, Integrity Research emphasized the crucial importance of filling key positions with local staff as they provide a project with invaluable access to programme sites and a more in-depth understanding of the context. This makes assessments more contextually relevant, conflict-sensitive, and, therefore, ultimately more sustainable and successful

- programming (see Integrity Research, "M&E: Challenges and Lessons from Syria", available from <https://www.integrityresearch.com/wp-content/uploads/2015-10-ME-in-Syria-Thought-Leadership.pdf> (accessed 6 October 2015)).
59. For large missions, three international staff and six nationally recruited M&E officers are suggested; see United Nations DPKO, "DDR Programme Management Toolkit", p. 9.
  60. In the worst cases, due to the high turnover of staff in the field, the lack of expertise available in peacekeeping missions may lead to a random staff officer being "baptised as M&E officer." Interview with senior UN DDR official.
  61. United Nations DPKO, "DDR Programme Management Toolkit", p. 10.
  62. DCAF and UN DPKO OROLSI, "Measuring the Impact of Peacekeeping Missions on Rule of Law and Security Institutions", p. 3.
  63. *Ibid.*, p. 7. This view was reflected by DDR experts, who pointed out disparities between mandating M&E and negative answers to funding requests from the budget committee. Demands were voiced from some UN DDR staff members to include M&E exercises of DDR programmes into the regular budget to ensure continuous financing of assessment.
  64. Robert Muggah (ed.), *Security and Post-Conflict Reconstruction: Dealing With Fighters In the Aftermath of War*, p. 75.
  65. As noted in the DCAF/OROLSI workshop paper, "results-based budget frameworks are useful in theory, but in practice they are inflexible instruments that lock peacekeeping missions into yearly planning and reporting. This provides a disincentive to evaluations that are intended to support timely readjustments within missions." See DCAF and UN DPKO OROLSI, "Measuring the Impact of Peacekeeping Missions on Rule of Law and Security Institutions", p. 3.
  66. James Pugel, for example, stresses the problem of planning for and conducting assessments in cases where "the majority of the country is inaccessible and 50 to 80 per cent of the infrastructure has been destroyed." See James Pugel, "Measuring Reintegration in Liberia", p. 76.
  67. It is important to note that the three phases of disarmament, demobilization, and reintegration are not to be equated with a programme's outputs, outcomes, and impact. Instead, each phase can have its distinct outputs, outcomes, and impacts. James Pugel, for example, stresses that outputs can be attributed to all aspects of DDR programmes and that one needs to "differentiate between easily quantifiable disarmament outputs and more complex reintegration outputs, as well as the causal relationships between outputs and outcomes" (see James Pugel, "Measuring Reintegration in Liberia").
  68. Security Council Resolution 2147, S/RES/2147, 28 March 2014.
  69. Security Council Resolution 1925, S/RES/1925, 28 May 2010.
  70. One DDR M&E expert argued that in the results-based framework, the accomplishments identified and derived from the mandate can be equated with the expected impact. See interview with senior UN DDR official.
  71. The incentive-based structure of DDR programmes — ex-combatants are offered assistance in exchange for disarming and standing down — can, for example, create situations where armed individuals volunteer for DDR programmes not once, but several times. This was observed by one of the interviewees in Afghanistan, and several other interviewees mentioned that it is possible to go through a DDR process several times, e.g., by providing a different name at registration.
  72. In the context of DDR, outputs are defined as the "specific products and services that emerge from processing inputs and activities through programme" (see United Nations, "3.50 Monitoring and Evaluation of DDR Programmes").
  73. Outcomes are defined as "the changes in conditions that the programme or project aims to achieve, outcome include the production of outputs and activities, and the contribution of partners" (see United Nations, "3.50 Monitoring and Evaluation of DDR Programmes").
  74. Impact is defined as the "positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended" (see *ibid.*).
  75. One senior DDR expert described outputs as "easily quantifiable and accessible measures for simple programme auditing purposes, which does not require much training and keeps the budget people happy." Interview with senior UN DDR official.
  76. Interview with senior UN DDR official.
  77. Interview with Dr Vanda Felbab-Brown.
  78. *Ibid.*
  79. See Michael Doyle and Nicholas Sambanis, "International Peacebuilding: A theoretical and quantitative analysis", *The American Political Science Review*, Vol. 94, No. 4 (Dec., 2000), pp. 779-801; and Michael Doyle and Nicholas Sambanis, *Making War and Building Peace: The United Nations since the 1990's* (Princeton: Princeton University Press, 2006).
  80. Interview with senior UN DDR official.
  81. United Nations DPKO, "DDR Programme Management Toolkit", p. 22.
  82. The Colombian Agency for Reintegration, for instance, defines "success" as a non-return to illegal activities, including illegal activities not linked to armed conflict, and claim a success rate of 76 per cent. Interview with and factsheet

- provided by Joshua Shuajo Mitrotti Ventura.
83. A more sophisticated methodology would require considering whether to focus on specific programmatic outputs or outcomes, or rather conduct a holistic analysis of the DDR programme's contribution towards overall objectives. In addition, some DDR experts interviewed emphasized that the problem of attribution can also become instrumentalized by programmers "fighting for turf" — i.e., wanting to demonstrate that their programme has yielded observable positive results.
  84. The *UN DPKO Standard Operating Procedures* recommend rapid small-scale surveys with random sampling ("probability sampling") or even mini-surveys with non-random sampling ("purposive sampling"), given the inability to conduct "technically complex and time-consuming surveys that may be more appropriate for academic research"; see United Nations DPKO, "DDR Programme Management Toolkit", p. 16.
  85. Quasi-experimental designs are similar to experimental designs, but lack the element of random assignments to treatment and control groups.
  86. Scott van der Stoep and Deirdre Johnston, *Research Methods for Everyday Life – Blending Qualitative and Quantitative Approaches* (San Francisco: John Wiley and Sons Publication, 2009), p. 106. Michael Gilligan, Eric Mvukiyehe, and Cyrus Samii, for example, used a quasi-experimental design to test the effectiveness of ex-combatant re-integration in Burundi by exploiting bureaucratic failures in the delivery of the re-integration benefits and halts in service delivery to measure programme effects; see Gilligan et al., "Reintegrating Rebels" p. 21.
  87. Interviews with Randolph Rhea and Dr Jaremev McMullin.
  88. Interview with UN DDR official.
  89. DPKO's Programme Management Toolkit states that "Triangulation reduces bias that can arise from relying on a single source or type of information, a major risk in volatile and complex environments typical for DDR" (see United Nations DPKO, "DDR Programme Management Toolkit", p. 16).
  90. This, however, raises questions regarding the externalities and comparability of results and ability to draw lessons learned across DDR programmes, given the context-tailored mandates of different DDR programmes.
  91. Biometric information as well as professional and military background information are usually entered in the database when registering ex-combatants.
  92. UNDP, "How to Guide Monitoring and Evaluation for Disarmament, Demobilization, and Reintegration Programmes", p. 22.
  93. Interview with senior UN DDR official.
  94. Interview with Dr Vanda Felbab-Brown.
  95. The aggregated data may, for example, provide insights into how many ex-combatants received reintegration packages in a certain area and how many of these were female.
  96. Interview with senior UN DDR official.
  97. Interview with Randolph Rhea.
  98. Pugel, "Measuring Reintegration", p. 78.
  99. United Nations, "3.50 Monitoring and Evaluation of DDR Programmes", p. 12.
  100. DCAF and UN DPKO OROLSI, "Measuring the Impact of Peacekeeping Missions on Rule of Law and Security Institution", p. 3.
  101. Interview with Dr Vanda Felbab-Brown.
  102. While this problem occurs not only in the UN system, but also with other international institutions, some other organizations are more tolerant when it comes to reporting negative developments.
  103. Franziska Seethaler, "Measuring the impact of security interventions: An introduction", UNU/SIPA Junior Research Fellowship Paper Series, Number 02, October 2015, available from [https://collections.unu.edu/eserv/UNU:3294/JRF02\\_MeasuringImpacts.pdf](https://collections.unu.edu/eserv/UNU:3294/JRF02_MeasuringImpacts.pdf) (accessed 08 February 2016), p. 7.
  104. Interview with UN DDR official.
  105. Interview with senior UN DDR official.
  106. Interview with Randolph Rhea.
  107. Interview with Dr Richard Bowd.
  108. It nevertheless is questionable whether full coherence among different actors in the DDR community is possible, given the variety of actors in the development and security fields.
  109. The DDR programmes with the most advanced databases were the ones with sufficient time, resources, and expertise to set up these tailored databases beyond the standard DREAM database. In the DDR programme in Sudan, a four-year delay in the programme start (2005–2009) allowed programmers to set up a sophisticated information management system and to collect baseline data: Interview with DDR Expert, UN WOMEN.
  110. Interview with Nick Imboden.
  111. Nick Imboden, Tino Kreuzer, Patrick Vinck, Phuong Pham, "DDR in the field: highly mobile data and information management", Powerpoint Presentation, December 2011.

112. Real-time automated data aggregation and data analysis with the help of software makes these new technologies cost- and time-efficient.
113. Given privacy concerns, these individuals will not consent to be tracked over an extended period of time. If long-term tracking is agreed to in exchange for assistance, this incentive-based structure, in turn, might introduce another bias in the study.
114. Vanessa Corlazzoli and Jonathan White, "Measuring the Un-measurable – Solutions to Measurement Challenges in Fragile and Conflict Affected Environments," Search for Common Ground/UK Department for International Development, March 2013, p. 15.
115. With regard to the funding of longitudinal studies, Robert Muggah argues that there has been "a widening engagement in DDR from multilateral and bilateral development agencies, foreign policy establishments and policy think tanks." (Robert Muggah, "Innovations in disarmament, demobilization, and reintegration research and policy", NUPI Working Paper 774, 2010, p. 11). Yet the commissioning of these studies greatly depends on the international interest and funding of these studies, which, as some DDR experts have highlighted, is declining.
116. Interview with UN DDR official.
117. It is ironic that when a peace operation actually contributes to the safety and stability in the country and the peace operation exits the country, there are no longer dedicated funds to invest in long-term assessment of the mission's impact.
118. United Nations, "3.50 Monitoring and Evaluation of DDR Programmes", p. 5.
119. As argued before, cross-pollination and accurate lessons learned can only be identified across different cases when taking into account the specific mandate and context of the particular DDR programme.
120. DCAF and UN DPKO OROLSI, "Measuring the Impact of Peacekeeping Missions on Rule of Law and Security Institutions", p. 7.
121. Ibid.
122. Interview with Dr Vanda Felbab-Brown.