

Evaluating Integrated Programming in Conflict-Affected Settings: Evidence, Learning, and Lessons from a Case Study in Pakistan

Hur Hassnain¹, Sarah Davies²
Humayun Khan³, and Zain Haideri⁴

Received: June 3rd, 2026
Accepted: Aug 30th, 2026

Abstract

This paper examines the performance and sustainability of integrated programming in conflict-affected settings through a case study of a multi-sectoral community development intervention implemented in Battagram District, Khyber Pakhtunkhwa, Pakistan, between 2018 and 2022. The intervention combined child protection, girls' education, economic resilience, social accountability, and community-based disaster risk management. Drawing on mixed-methods evaluation evidence, including survey data, 143 community change stories, key informant interviews, focus group discussions, and field observations, the article assesses whether integrated programming generated meaningful and sustainable change. The findings show substantial perceived benefits at individual and household levels: 76% of respondents reported that the change described had an immense impact on their quality of life, 76% considered the change lasting, 83% reported high social acceptance, and 85% believed the change would not have occurred without the project's contribution. Enterprise participants also reported income gains, with average monthly income increasing from PKR 291 to PKR 4,591 among women and from PKR 2,914 to PKR 15,569 among men. However, institutionalisation remained weak: 61% of respondents were unaware of the Child Protection Unit, 64% were unaware of Child Protection Committees, and 73% were unaware of Disaster Management Committees. The article argues that integrated programming can generate relevant short-term outcomes in fragile contexts, but durable change depends on deeper institutional anchoring, realistic programme scope, market alignment, gender-sensitive design, and adaptive monitoring and learning systems.

Keywords: integrated programming, sustainability, fragile contexts, Pakistan, community development, evaluation

¹ Independent Evaluation Specialist, Belgium

² Director Evaluations, I4D Consulting, United Kingdom

³ Programme Manager, CESVI, Pakistan

⁴ Evaluation Manager, I4D Consulting, Pakistan

Introduction

In fragile and conflict-affected settings, development challenges rarely occur in isolation. Poverty, weak public services, limited livelihood opportunities, gender inequality, restricted access to education, and exposure to recurrent natural hazards often interact and reinforce one another (Author, 2024). These challenges are also shaped by contested information environments, where evidence does not automatically translate into action and where community ownership, trust, and adaptive learning become central to effective evaluation and programming (Author & Author, 2026). This complexity has contributed to growing interest in integrated and area-based approaches that seek to address multiple dimensions of vulnerability within the same geographic and social context. Area-based programming is particularly relevant in such settings because it takes the locality, rather than a single sector or target group, as the primary entry point, allowing interventions to respond more directly to the interconnected realities of people's lives (Haider, 2021).

The rationale for integrated programming is compelling, but the empirical evidence remains more nuanced. A systematic review of 601 evaluations of integrated development interventions found that most studies assessed whether programmes achieved their intended outcomes, while only 26 used designs capable of testing whether combining sectors produced synergistic effects; of these, only seven detected such effects (Ahner-McHaffie et al., 2017). This points to an important evidence gap: the added value of integration itself, rather than simply the effectiveness of multi-sectoral interventions.

In practice, development organisations have increasingly turned to integrated community development approaches that aim to address several challenges at once (Independent Evaluation Office of UNDP, 2018; World Bank, 2020). The logic is clear: by combining interventions across sectors, programmes may be able to generate broader and more meaningful improvements in people’s lives. In fragile settings, however, this promise comes with practical demands. Integrated approaches may respond more appropriately to overlapping vulnerabilities, but they also require organisations to coordinate across sectors, engage with multiple institutions, and maintain sufficient implementation capacity. Recent evidence from humanitarian-development-peace programming similarly emphasises the importance of coordination, localisation, institutional alignment and sustained engagement in complex settings (Independent Evaluation Office of UNDP, 2024).

Despite this growing interest, important questions remain about what happens after integrated programmes generate visible short-term results. Do these gains continue once project support ends, or do they remain dependent on external inputs? Does addressing several issues at once strengthen outcomes, or can it stretch resources and weaken implementation depth (Author, 2023)? These questions are particularly important in fragile and resource-constrained settings, where short implementation periods, weak public systems and limited follow-up capacity can constrain the institutionalisation of project-supported mechanisms.

This article examines these questions through the case of the *Integrated Children and Women Centred Community Development* project, implemented between 2018 and 2022 in three union councils of Battagram District, Khyber Pakhtunkhwa, Pakistan. Battagram is a remote and



mountainous district with low socio-economic indicators, limited infrastructure, and recurrent exposure to natural disasters such as earthquakes, floods and landslides. Livelihood opportunities are limited, particularly for women, and access to education beyond the primary level remains a challenge for many families. Social norms also shape daily life, often restricting women's mobility and participation in economic activities.

The project aimed to improve the quality of life of women and children through five main components: strengthening child protection systems, improving the quality of girls' education, promoting economic resilience through skills and enterprise development, introducing social accountability mechanisms in public services, and building community-based disaster risk management systems. Its design reflected many principles widely promoted in development practice: it combined social, economic and institutional interventions, engaged community structures, and placed women and children at the centre. At the same time, it raised important questions about focus, implementation capacity and long-term sustainability.

The study asks two questions: (1) What changes were associated with the integrated intervention across household, community and institutional domains? and (2) What factors enabled or constrained the sustainability of these changes? Drawing on a mixed-methods evaluation conducted at the end of the project, the study uses quantitative survey evidence, community narratives, key informant interviews, focus group discussions and field observations. It is guided by the OECD DAC criteria of relevance, effectiveness, efficiency, impact and sustainability, while placing particular emphasis on community members' own accounts of

change (Chambers, 1994). This allows for a grounded understanding of how the project influenced people's lives and where its limitations became visible.

The paper argues that integrated programming can generate meaningful improvements in complex settings, but integration should not be equated with institutionalisation or sustainability. The Battagram case suggests that positive individual and household outcomes may coexist with weak community awareness and limited institutional consolidation. Sustainability therefore depends not only on the number of sectors addressed, but on the depth of engagement with local institutions, markets and social structures, the realism of programme scope, and the capacity to adapt implementation through monitoring and learning. In this sense, the study contributes to wider policy debates on how development programmes can balance ambition with focus, and move beyond short-term results towards more durable and meaningful impact.

Methodology

This article draws on the final evaluation of the *Integrated Children and Women Centred Community Development* project and adopts a mixed-methods case study design. The purpose of the methodology was not only to assess the results achieved, but also to understand the processes through which those results emerged, and the factors that enabled or constrained their sustainability. The evaluation was guided by the OECD DAC criteria of relevance, effectiveness, efficiency, impact and sustainability (OECD, 2010), which provided a structured framework for examining both programme performance and the conditions influencing longer-term change.

Given the multi-sectoral nature of the intervention and the complexity of the Battagram context, the study combined quantitative survey evidence with qualitative and narrative data. The

quantitative component included a structured survey administered through the Sprockler platform, which enabled the collection of numerical, perception-based and narrative data from project participants. This was complemented by qualitative evidence from 149 individual change stories, 24 key informant interviews, focus group discussions with community members across the three target union councils, and field observations. Together, these sources allowed the evaluation to examine reported outcomes across household, community and institutional domains.

The sampling strategy was purposive and component-based. Participants were selected to reflect the main intervention areas of the project, including child protection, girls' education, economic resilience, social accountability and community-based disaster risk management. The sample included women, men, children and other community members who were either directly involved in project activities or indirectly affected by them. Key informants were selected because of their knowledge of the project, the local context, or relevant institutional systems, and included government officials, teachers, community representatives and project staff. This approach was intended to capture variation in experience across beneficiary groups, sectors and institutional perspectives, rather than to produce a statistically representative sample of the district population.

A key feature of the methodology was the use of story-based approaches, particularly Outcome Harvesting and Most Significant Change (Wilson-Grau & Britt, 2012; Davies & Dart, 2005). These approaches allowed participants to describe changes in their own words and helped capture outcomes that may not have been visible through standard indicators alone. The change

stories were used to understand how participants experienced the intervention, what kinds of changes they considered most significant, and whether those changes were perceived as meaningful, socially accepted and likely to last. This was particularly important in a context where formal monitoring data were limited and where changes in confidence, agency, awareness and social relations were central to the intervention's intended effects.

Qualitative data were analysed thematically using an inductive approach. Interview notes, focus group findings, change stories and field observations were reviewed to identify recurring themes, patterns of change and explanatory factors. Particular attention was given to differences between reported individual and household outcomes, on the one hand, and the institutionalisation of project-supported mechanisms, on the other. Themes were then compared across data sources and project components to assess convergence, divergence and gaps in the evidence. Quantitative findings from the Sprockler survey were used to identify patterns in participants' perceptions of change, while qualitative material was used to interpret how and why those patterns emerged.

Triangulation was applied throughout the analysis to strengthen the credibility of findings. Survey responses, change stories, interviews, focus group discussions and field observations were compared to test whether findings were consistent across different sources and stakeholder groups. Where evidence differed, the analysis examined these differences rather than treating them as inconsistencies to be ignored. This was particularly useful for understanding the contrast between positive individual-level outcomes and weaker evidence of community awareness, institutional ownership and sustainability.

Several limitations should be acknowledged. First, the study relied partly on retrospective and self-reported accounts, which may be affected by recall bias or respondents' desire to present positive change. Second, the absence of a strong monitoring system during implementation meant that baseline and follow-up data were not always consistent or complete. Third, the purposive sampling strategy supported depth and diversity of perspectives but does not allow the findings to be generalised statistically to the wider district population. Finally, as the evaluation was conducted at the end of the project, it could assess perceived sustainability and early signs of institutionalisation, but not long-term outcomes several years after project closure.

Despite these limitations, the mixed-methods design provides a credible basis for assessing both the achievements and constraints of the intervention. By combining quantitative evidence with community narratives, interviews, focus group discussions and field observations, the methodology allowed the study to move beyond a narrow assessment of outputs and examine how integrated programming was experienced, interpreted and sustained in practice.

Findings

The findings present a differentiated picture of change. The project was well aligned with the needs of communities in Battagram and generated meaningful short-term improvements, particularly at individual and household levels. Participants reported improvements in income, confidence, awareness and engagement with community processes. However, these gains were uneven across components and were not always matched by stronger institutional systems or sustained community-level structures. The central pattern emerging from the evidence is therefore not one of simple success or failure. Rather, the project produced visible and valued

changes in people's lives, while struggling to embed all of these changes within local institutions, markets and community practice.

Relevance in a fragile and underserved context

The project was highly relevant to the realities of Battagram. The district faces overlapping challenges, including poverty, weak public services, limited access to education, restricted livelihood opportunities and recurrent exposure to natural hazards. Women and girls experience additional constraints linked to mobility, social norms, access to education and participation in economic life. In this context, the project's focus on child protection, girls' education, livelihoods, social accountability and disaster risk management responded to issues that were important to communities and consistent with the wider vulnerabilities of the area.

The relevance of the intervention was strongest at the level of overall design. By bringing together social, economic and institutional components, the project recognised that community needs were interconnected and could not be addressed through a single-sector response. The emphasis on women and children was also appropriate in a setting where these groups often face exclusion from services, decision-making and economic opportunities. At the same time, relevance at the design level did not always translate into equal effectiveness at the activity level. Some activities were not sufficiently adapted to local market conditions, community preferences or the capacity of local structures to sustain them after project support ended. This distinction between a relevant design and uneven implementation is important for understanding the findings that follow.

Individual and household-level change

The strongest evidence of change appears at the individual and household level. The Sprockler data show that many participants experienced the project as meaningful in their own lives. Of the 143 respondents whose stories were analysed, 76% reported that the change described had an immense impact on their quality of life. A further 10% gave a neutral assessment, while 14% reported no impact. Similarly, 76% characterised the change as lasting, compared with 17% who considered it a one-off change. Social acceptance was also high, with 83% reporting strong support for the change described.

Respondents also attributed a substantial role to the project in bringing about these changes. Eighty-five percent stated that the change would not have occurred without the project's influence, while 8% believed it would have happened anyway and 7% were neutral. These figures suggest that participants did not simply value the changes they experienced; they also saw the project as an important contributor to those changes.

These findings should be interpreted carefully. They are based on participant perceptions and retrospective accounts rather than a counterfactual impact design. They therefore do not provide a precise estimate of causal impact. However, they do offer credible evidence of reported change, perceived contribution and personal significance. In a fragile and resource-constrained setting, this form of evidence is important because it shows how people themselves understood the value of the intervention in relation to their everyday lives.

Livelihoods, income and economic constraints

Economic resilience was one of the most visible areas of change. Skills training and small enterprise support enabled some participants to start income-generating activities, improve existing livelihoods or contribute more regularly to household expenses. This was particularly important for women, many of whom had limited opportunities for paid work outside the home. Participants described being able to support children's education, meet household needs and gain a stronger sense of confidence and agency.

The enterprise data show measurable income gains among participants. The dataset included 84 respondents, of whom 55 were women and 29 were men. Participants received training in a range of trades and enterprises, including tailoring and embroidery, poultry management, carpentry, electrical work, plumbing, mobile repair and mushroom farming. Reported average monthly income among female participants increased from PKR 291 in 2018 to PKR 4,591 in 2022. Among male participants, average monthly income increased from PKR 2,914 to PKR 15,569 over the same period.

These income gains are significant, but they were not equally secure or consistent across all participants. The sustainability of livelihood outcomes depended heavily on the type of enterprise, the relevance of the trade, access to tools and raw materials, household responsibilities, and links to local buyers. Among enterprise participants, access to raw materials and tools was the most frequently reported constraint, identified by 31% of respondents. This was followed by social and cultural norms at 30%, access to buyer markets at 19%, and household responsibilities at 15%.

The livelihood findings therefore show both promise and limitation. Skills and enterprise support created meaningful opportunities for some participants, especially women, but training and start-up support were not enough on their own. Where market access was weak, social restrictions remained strong, or follow-up support was limited, economic activities struggled to grow beyond small or irregular income generation. This suggests that livelihood components in fragile settings need to be designed not only around training delivery, but around the wider conditions that allow enterprises to survive.

Community awareness and institutionalisation

The evidence is weaker when the analysis moves from individual outcomes to community awareness and institutional consolidation. The project established or supported structures related to child protection, education, social accountability and disaster risk management. However, awareness and participation in these mechanisms remained limited among surveyed respondents.

In relation to child protection, 61% of respondents reported that they did not know about the Child Protection Unit or its role, while 64% were unaware of Child Protection Committees. Awareness of child protection legislation was also limited, with 54% reporting that they did not know of any law protecting children in their communities. This suggests that although child protection structures were introduced or supported, their visibility and practical accessibility remained limited for many community members. Similar patterns were found in education and disaster risk management. Fifty-one percent of respondents were unaware of School Improvement Plans, and only 28% reported that they or another household member had

participated in them. In relation to disaster risk management, 67% were unaware of a union council- or village-level disaster risk management or rehabilitation plan, while 73% were unaware of Disaster Management Committees.

These findings point to an important distinction between establishing institutional mechanisms and institutionalising them. A committee, plan or unit may exist formally, but this does not necessarily mean that communities know about it, trust it, participate in it, or use it in practice. In Battagram, project-supported structures appear to have been created or reactivated, but their visibility, routine use and community ownership were less firmly established. This gap between formal establishment and practical institutionalisation is central to understanding the sustainability challenge in the project.

Education, child rights and social awareness

The project also contributed to education and awareness-raising, particularly around girls' education and child rights. Community discussions, school-level engagement and project-supported structures helped create space for dialogue on issues that are often sensitive or neglected. Teachers and community members reported increased awareness of the importance of girls' education, child protection, corporal punishment and early marriage. In some cases, improved interaction between schools, parents and community members was also reported.

However, the evidence suggests that these changes were stronger in awareness and dialogue than in system-level consolidation. While communities became more exposed to child rights and education-related messages, many participants still lacked clear knowledge of how to report or respond to child protection concerns. Formal mechanisms for case management also

remained limited. In education, the low awareness of School Improvement Plans and limited reported participation suggest that school-community accountability mechanisms had not yet become a regular part of local practice.

The education and child protection findings therefore follow the wider pattern of the project. The intervention helped raise awareness and encouraged engagement, but it did not always create the institutional depth required for sustained change. This does not diminish the value of the awareness generated, particularly in a conservative and underserved setting. It does, however, show that awareness-raising needs to be linked more deliberately to functioning systems, clear referral pathways and continued institutional support.

Disaster risk management and social accountability

Disaster risk management and social accountability were among the more difficult components to institutionalise. Although committees were formed and some training was provided, community awareness of disaster risk management structures remained low. Participation was also limited, particularly among women. The finding that 73% of respondents were unaware of Disaster Management Committees is especially important in a district exposed to earthquakes, floods and landslides.

Social accountability mechanisms faced similar challenges. Tools such as community scorecards were introduced, and some improvements in engagement with public services were reported. However, these mechanisms were not widely understood or consistently used by community members. This limited their ability to become part of routine interaction between communities and service providers.



These findings suggest that institutional components require a different level of time, trust and follow-up than short project cycles often allow. Disaster risk management and social accountability are not only technical activities; they depend on repeated interaction, clarity of roles, confidence in institutions and community willingness to participate. Without sustained support, such mechanisms may remain present on paper but weak in practice.

Structural barriers to sustainability

Across the project, several common factors constrained sustainability. The first was weak alignment with market realities. Many livelihood activities were relevant to participants' immediate needs, but not all were sufficiently connected to local demand, value chains or buyer networks. As a result, some income-generating activities remained small, irregular or dependent on continued support.

The second constraint was limited follow-up and weak monitoring, evaluation and learning systems. The project generated useful changes, but there was limited evidence of systematic tracking after training or initial support. This made it harder to identify emerging problems, adapt activities or understand which components were most likely to continue beyond the project period.

The third constraint was the continued influence of socio-cultural norms, particularly for women. Even where women gained skills or received enterprise support, restrictions on mobility, household responsibilities and social expectations limited the types of activities they could pursue and the scale at which they could operate. The data on reported constraints among

enterprise participants show clearly that economic sustainability was shaped not only by technical training, but also by gendered social conditions.

The fourth constraint was institutional capacity. Community structures were established or supported, but not always sufficiently embedded in local systems. This affected the durability of child protection, education, disaster risk management and social accountability mechanisms. Where communities were unaware of structures, or where roles and pathways were unclear, the likelihood of sustained use was reduced.

The breadth–depth challenge

A recurring lesson across the findings is that the project attempted to address many important issues at the same time. Each component was relevant, and the integrated design reflected the complexity of community needs. However, the breadth of the intervention also placed pressure on implementation capacity. Covering child protection, girls' education, livelihoods, social accountability and disaster risk management within one project created a wide agenda that required coordination across sectors, institutions and community groups.

This breadth was both a strength and a limitation. It allowed the project to respond to multiple vulnerabilities, but it also reduced the depth of support available to each component. In several areas, activities were implemented and appreciated but did not receive the sustained follow-up needed to become fully embedded. The findings therefore suggest a breadth–depth trade-off: integrated programming can improve relevance by addressing interconnected needs, but it can also weaken sustainability if the scope is not matched by time, resources, institutional capacity and adaptive learning.

Overall pattern of findings

Overall, the project achieved meaningful short-term results. Many participants experienced improvements in income, confidence, awareness and engagement. These changes were visible, valued and, in many cases, strongly attributed to the project. At the same time, the evidence shows that individual and household-level change did not automatically translate into stronger community systems or institutionalised mechanisms.

The main finding is therefore that integrated programming can generate important benefits in fragile settings, but integration alone does not ensure sustainability. Durable change depends on what happens after activities are delivered: whether institutions take ownership, whether communities know and use the mechanisms created, whether livelihood support is linked to real markets, whether gender norms are addressed, and whether monitoring and learning systems allow programmes to adapt over time. In Battagram, the project made a meaningful contribution, but its longer-term value was limited by the extent to which these conditions were only partially in place.

Discussion

The findings from Battagram offer a useful contribution to debates on integrated programming in fragile and conflict-affected settings. They show that integrated approaches can generate meaningful short-term benefits, particularly at individual and household levels, but that these benefits do not automatically translate into sustained institutional or community-level change. This distinction matters because much of the appeal of integrated programming rests on the assumption that working across sectors will produce broader, mutually reinforcing and more

durable outcomes. The evidence from this case is more cautious. Integration improved relevance and helped respond to overlapping vulnerabilities, but sustainability depended on the depth of implementation, the strength of local systems, market linkages, social norms and the capacity to continue learning after activities had been delivered.

Integration can generate outcomes, but not necessarily synergy

The Battagram case supports the rationale for integrated programming but also confirms the need for a more careful understanding of what integration can realistically achieve. Ahner-McHaffie et al. (2017) found, in their systematic review of 601 evaluations of integrated development interventions, that most studies assessed whether programmes achieved their intended outcomes, but far fewer examined whether integration itself produced synergistic effects. Only 26 evaluations used designs capable of testing synergy, and only seven found such effects. The implication is important: positive outcomes across several components should not automatically be interpreted as evidence that integration has produced added value.

This study does not test statistical synergy. Its contribution is different. It examines whether a multi-component intervention generated changes that were meaningful to participants, and whether these changes became embedded in local systems. The answer is mixed. The project produced visible gains in income, awareness, confidence and engagement, but these gains were not consistently accompanied by strong institutionalisation. This suggests that integration may increase the relevance of a programme by addressing several interconnected needs, but relevance alone is not enough. For integration to become more than the sum of separate activities,

components need to reinforce one another through shared systems, sequencing, institutional anchoring and sustained follow-up.

This finding is consistent with wider evidence from humanitarian-development-peace programming, which increasingly stresses coordination, localisation, institutional alignment and sustained engagement in complex settings. The Independent Evaluation Office of UNDP's (2024) synthesis of 65 evaluations and studies on the humanitarian-development-peace nexus similarly highlights the importance of working across institutional boundaries while maintaining attention to local actors and long-term engagement. The OECD and UNHCR (2024) review of 27 nexus initiatives also points to programming, coordination and financing as central challenges in translating joined-up approaches into operational practice. The Battagram case adds to this literature by showing that the challenge is not only coordination between actors, but also the practical institutionalisation of mechanisms at community level.

The breadth–depth trade-off

A second issue concerns the trade-off between breadth and depth. The project's five-component design was appropriate to the multidimensional nature of vulnerability in Battagram. Child protection, girls' education, livelihoods, social accountability and disaster risk management were all relevant to the local context. However, the same breadth that made the project responsive also placed considerable pressure on implementation capacity. Each component required different expertise, institutional relationships, follow-up arrangements and forms of community engagement.

The findings suggest that the project was strongest where activities produced direct and visible benefits for participants, particularly in relation to livelihoods and individual awareness. It was weaker where change depended on sustained institutional practice, such as disaster risk management committees, child protection structures, school improvement processes and social accountability mechanisms. This does not mean that the integrated design was wrong. Rather, it suggests that integrated programming in fragile settings requires sharper choices about sequencing, prioritisation and depth. Trying to do many important things at once can reduce the time and attention needed to make any one of them durable.

This point connects with earlier work on evaluation in fragile, conflict and violence-affected contexts, which emphasises the need for realism, conflict sensitivity, adaptive management and careful attention to local systems (Author et al., 2021). In such settings, programme ambition must be matched by implementation capacity. Fragility is not only a contextual label; it affects what can be delivered, how quickly change can be consolidated, and how much follow-up is required for results to last. The Battagram case therefore reinforces the argument that integrated programming should be designed around the capacity to sustain change, not only the ambition to address multiple needs. This also aligns with Patton's work on developmental evaluation (Patton, 2011), which argues that programmes operating in complex and uncertain environments require adaptive learning rather than fixed implementation models.

Individual change versus institutionalisation

The strongest contribution of this paper lies in the contrast between individual-level change and institutional consolidation. At the individual and household level, the evidence is



encouraging. Among respondents whose stories were analysed, 76% reported that the change described had an immense impact on their quality of life, 76% considered it lasting, 83% reported high social acceptance, and 85% believed the change would not have occurred without the project's contribution. These are important findings, particularly in a remote and underserved district where even modest improvements in income, confidence or awareness can matter greatly in people's everyday lives.

Yet the institutional evidence tells a more fragile story. A majority of respondents were unaware of several structures that were meant to support longer-term change: 61% did not know about the Child Protection Unit, 64% were unaware of Child Protection Committees, 51% were unaware of School Improvement Plans, 67% were unaware of local disaster risk management or rehabilitation plans, and 73% were unaware of Disaster Management Committees. This contrast is central to the article. It shows that participants may experience meaningful change without the wider system becoming strong enough to sustain or reproduce that change.

This finding also speaks to wider debates on evaluation use and evidence-to-action. Author and Author (2026), writing on translating evidence to action in a post-truth world, argue that evidence does not move into action automatically; it requires trust, interpretation, ownership and context-sensitive engagement. The Battagram case shows a similar dynamic at programme level. Project-supported mechanisms cannot be assumed to function simply because they have been created. They need to be understood, trusted, used and connected to the institutions and relationships through which people make decisions.



Sustainability as institutional and social embedding

The Battagram findings suggest that sustainability should be understood less as a final project outcome and more as a process of embedding. A livelihood activity is more likely to endure when it is connected to markets, tools, raw materials and household realities. A child protection structure is more likely to matter when communities know about it, trust it and understand how to use it. A disaster risk management committee is more likely to function when its roles are clear, participation is broad and links with public systems are maintained. Sustainability, in this sense, depends on whether project-supported changes become part of ordinary institutional and social practice.

This is where the case extends existing evidence on integrated and area-based programming. Haider's (2021) review of area-based programming in fragile and conflict-affected contexts highlights the relevance of locality-based approaches in settings where needs are complex and interconnected. The Battagram case supports that argument but adds a caution: locality is a useful entry point, but not a guarantee of institutional depth. Area-based programming can bring interventions closer to people's lived realities, but it still requires sufficient time, follow-up and institutional anchoring to become sustainable.

This also connects with Picciotto's (2007, 2014) wider argument that development evaluation must use methods and frameworks appropriate to contemporary development dilemmas, rather than relying on narrow or overly simplified models of attribution and effectiveness. In fragile settings, where institutions are weak and outcomes are shaped by social, political and economic uncertainty, evaluation needs to examine not only whether activities were

delivered or outcomes reported, but how those outcomes were produced, who owned them, and whether the systems around them became stronger.

Evidence, power and local ownership

The findings also have implications for how evidence is generated and interpreted in fragile settings. The positive change stories show that people valued many of the changes associated with the project. However, the weaker evidence on institutional awareness reminds us that reported benefit and sustained ownership are not the same thing. This distinction is important from a decolonising evaluation perspective. Earlier work on decolonising evaluation argues that evaluation needs to pay closer attention to whose knowledge counts, how evidence is translated, and whether learning is meaningful to communities as well as to donors and implementing agencies (Author, 2023).

In Battagram, the use of community narratives helped capture forms of change that may have been missed by standard indicators alone. At the same time, the findings show that community voice must be connected to programme adaptation and institutional learning. Listening to participants is not enough if the systems required to sustain change remain weak. This is particularly relevant for integrated programming, where different components may produce different types of value for different groups. For women, for example, income generation was not only an economic outcome; it was also linked to confidence, contribution to household needs and a sense of agency. But these gains remained shaped by social norms, mobility restrictions and access to markets.

Implications for policy and practice

The discussion points to several implications that are taken forward in the next section. First, integrated programmes should be designed for institutionalisation, not only implementation. Creating committees, plans or accountability tools is not enough unless communities know about them, use them and see them as relevant. Second, programme breadth should be matched with realistic implementation capacity. A multi-sectoral design may be justified in fragile settings, but only if sufficient time, resources and follow-up are available for each component to mature.

Third, livelihood support needs stronger links to local markets and value chains. Training and start-up support can create opportunities, but these opportunities are unlikely to last if participants lack access to tools, raw materials, buyers and continued advice. Fourth, gender strategies need to go beyond participation targets. Women's economic engagement depends not only on skills, but also on household responsibilities, mobility, social acceptance and the negotiation of local norms. Finally, monitoring, evaluation and learning systems need to be built into programme delivery from the beginning, so that implementers can identify what is working, what is not, and where adaptation is needed.

Overall, the Battagram case does not suggest that integrated programming is ineffective. Rather, it shows that integration requires discipline. In fragile and resource-constrained settings, the challenge is not simply to combine sectors, but to ensure that the combination is coherent, manageable and rooted in local systems. The policy lesson is therefore to move from integration as design ambition to integration as sustained practice. This means focusing less on the number

of components included in a programme and more on whether those components are sufficiently connected to institutions, markets, communities and learning systems to endure beyond the project period.

Policy Implications

The Battagram case offers several practical a way of connecting interventions to the institutions, markets, social norms and learning processes that determine whether change can last. First, programmes should design for institutionalisation, not only establishment. Creating committees, plans, community structures or accountability mechanisms is not enough if people do not know about them, trust them or use them. From the outset, integrated programmes need clear strategies for how project-supported mechanisms will become part of routine local practice, including who will own them, how they will be resourced, and how communities will access them after project closure.

Second, programme breadth should be matched with implementation capacity. Multi-sectoral designs are often justified in fragile settings because people's needs are interconnected, but covering too many components can reduce the depth and quality of implementation. Rather than treating integration as the inclusion of many sectors, programmes should focus on a smaller number of clearly connected priorities that can be implemented well, followed up properly and sustained over time.

Third, livelihood support should be anchored in markets and value chains. Skills training and enterprise inputs are unlikely to generate durable income unless they are linked to local demand, access to raw materials, tools, buyers and continued advice. Livelihood components

should therefore be informed by practical market assessments and designed around the real economic conditions participants face, especially women working within restrictive mobility and household constraints.

Fourth, gender strategies need to go beyond participation targets. Increasing women's access to training, services or community activities is important, but it is not sufficient where social norms continue to limit mobility, decision-making and economic participation. Programmes should engage families, community leaders and local institutions more deliberately, so that women's participation is supported by a wider enabling environment rather than treated as an individual responsibility alone.

Fifth, monitoring, learning and follow-up should be built into programme design from the beginning. Integrated programmes in fragile settings need practical systems to track not only activity delivery, but also whether outcomes are being sustained, adapted and institutionalised. This requires post-training follow-up, routine feedback from participants, attention to unintended constraints, and space to adjust implementation while the project is still active.

Finally, institutional components such as child protection, social accountability, school improvement and disaster risk management require longer time horizons than short project cycles often allow. These areas depend on trust, repeated engagement, clarity of roles and links with public systems. Where such mechanisms are important to the theory of change, donors and implementers should consider phased programming, transition support or follow-on engagement rather than assuming that structures created during implementation will continue on their own.



Overall, the policy lesson from Battagram is that integrated programming should move from ambition to disciplined practice. Sustainable outcomes depend less on the number of components included in a project, and more on whether those components are realistic, locally owned, institutionally anchored, economically viable and supported by learning over time.

Conclusion

This article has examined whether integrated programming in a fragile and resource-constrained setting can generate meaningful and sustainable change. The evidence from Battagram shows that integration can improve relevance and produce valued results at individual and household levels. Participants reported improvements in income, awareness, confidence and engagement, and the project responded to real and interconnected needs in the district. However, the same evidence also shows that positive short-term outcomes should not be confused with sustainable transformation.

The central finding is that integrated programming requires a second layer of integration. It is not enough to combine several sectors within one project design. For outcomes to last, interventions also need to become embedded in local institutions, markets, social norms and learning systems. In Battagram, livelihood support generated measurable income gains for some participants, but its durability was constrained by weak market linkages, limited access to tools and raw materials, household responsibilities and restrictive gender norms. Similarly, structures related to child protection, school improvement, social accountability and disaster risk management were introduced or supported, but community awareness and routine use remained limited.

The study therefore cautions against treating integration as an automatic route to sustainability. Multi-sectoral programming can be highly relevant in complex settings, but its effectiveness depends on disciplined prioritisation, realistic scope, adequate sequencing and continued engagement. Programmes that attempt to address many needs at once may produce visible outputs and valued individual changes yet still fall short of institutionalisation if local systems are not strong enough, or sufficiently supported, to carry the work forward.

The findings should also be read with some caution. The study draws on a final evaluation and relies partly on retrospective and self-reported accounts of change. Baseline and follow-up data were not always complete, and the design does not allow for a counterfactual estimate of impact. The evidence is therefore strongest in showing reported outcomes, perceived contribution, and early signs of sustainability or fragility, rather than long-term causal effects. These limitations, however, are also part of the reality of evaluation in fragile and underserved contexts, where data systems are often weak and where community narratives remain essential for understanding how change is experienced.

The broader lesson from Battagram is clear. Integrated programming should not be judged by the number of sectors it brings together, but by the depth of its connection to the systems that sustain change. This means designing for institutionalisation from the beginning, linking livelihood support to real markets, addressing gendered and social constraints, investing in monitoring and learning, and allowing enough time for local ownership to develop. In fragile settings, lasting change depends less on doing many things at once, and more on doing the right things deeply enough for people, institutions and communities to carry them forward.



References

- Author. (2023). *Details omitted for double-anonymized peer review*.
- Author. (2024). *Details omitted for double-anonymized peer review*.
- Author & Author. (2026). *Details omitted for double-anonymized peer review*.
- Author, Author, & Author. (2021). *Details omitted for double-anonymized peer review*.
- Ahner-McHaffie, T. W., Guest, G., Petruney, T., Eterno, A., & Dooley, B. (2017). Evaluating integrated development: Are we asking the right questions? A systematic review. *Gates Open Research*, 1, Article 6. <https://doi.org/10.12688/gatesopenres.12755.1>
- Chambers, R. (1994). Participatory rural appraisal: Challenges, potentials and paradigm. *World Development*, 22(10), 1437–1454.
- Davies, R., & Dart, J. (2005). *The Most Significant Change (MSC) technique: A guide to its use*. CARE International.
- Haider, H. (2021). *Area-based programming in fragile- and conflict-affected contexts*. Institute of Development Studies and Partner Organisations. <https://hdl.handle.net/20.500.12413/17110>
- Independent Evaluation Office of UNDP. (2018). *Evaluation of UNDP support to poverty reduction in the least developed countries*. United Nations Development Programme. <https://erc.undp.org/evaluation/documents/download/12300>
- Independent Evaluation Office of UNDP. (2024). *Humanitarian-development-peace nexus approaches in conflict and disaster settings* (Reflections Series). United Nations

- Development Programme. <https://www.undp.org/evaluation/publications/humanitarian-development-peace-nexus-conflict>
- OECD. (2010). *Quality standards for development evaluation*. OECD Publishing. <https://www.oecd.org/dac/evaluation/qualitystandards.pdf>
- OECD & UNHCR. (2024). *The humanitarian-development-peace nexus and forced displacement: Progress, insights and recommendations for operational practice* (OECD Development Policy Papers No. 57). OECD Publishing. https://www.oecd.org/en/publications/the-humanitarian-development-peace-nexus-and-forced-displacement_3e493170-en.html
- Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. Guilford Press.
- Picciotto, R. (2007). The new environment for development evaluation. *American Journal of Evaluation*, 28(4), 509–521. <https://doi.org/10.1177/1098214007306371>
- Picciotto, R. (2014). Have development evaluators been fighting the last war... and if so, what is to be done? *IDS Bulletin*, 45(6), 6–16. <https://doi.org/10.1111/1759-5436.12109>
- Wilson-Grau, R., & Britt, H. (2012). *Outcome harvesting*. Ford Foundation. <https://www.outcomeharvesting.net>
- World Bank. (2020). *World development report 2020: Trading for development in the age of global value chains*. World Bank. <https://doi.org/10.1596/978-1-4648-1457-0>