Protocol for Assessing FCV-Sensitivity in PADs

Introduction

The **FCV-Sensitivity Assessment Protocol** is designed to evaluate the extent to which World Bank Project Appraisal Documents (PADs) integrate sensitivity to fragility, conflict, and violence (FCV) in their design and delivery. By using a structured set of criteria, the protocol assesses whether projects effectively address FCV-related risks, promote resilience, and avoid unintended consequences that could exacerbate fragility or inequality. The protocol is grounded in the framework outlined in the attached paper, focusing on five key characteristics of FCV-sensitive climate action.

The five characteristics assessed by the protocol are as follows:

1. **Consider How Interactions Between Climate and FCV Affect Program Delivery**: This characteristic evaluates whether the PAD identifies FCV-related risks that could disrupt project implementation and explores how climate-related impacts interact with drivers of fragility, conflict, and violence. The guiding questions focus on whether the PAD acknowledges these dynamics and designs measures to address them.
2. **Mitigate the Risk of Climate Actions Resulting in Maladaptation**: This characteristic examines whether the PAD includes safeguards to prevent unintended consequences, such as exacerbating vulnerabilities or reinforcing inequalities. It also considers whether adaptive mechanisms are embedded to address evolving FCV conditions and whether short-term needs are balanced with long-term resilience-building.
3. **Prioritize Climate Actions That Address FCV Root Causes and Enhance Peacebuilding**: This characteristic assesses whether the PAD includes interventions targeting the structural drivers of fragility, such as inequitable access to resources or weak governance, and whether it actively seeks to promote peacebuilding by fostering trust, social cohesion, or conflict resolution.
4. **Prioritize the Needs and Capacities of Vulnerable Regions and Groups**: This characteristic focuses on whether the PAD identifies vulnerable populations disproportionately affected by climate and FCV dynamics and includes measures to address their needs. It also evaluates whether the project ensures equitable benefit-sharing and avoids reinforcing inequalities.
5. **Encourage Coordination Across Development, DRM, and Peacebuilding Actors**: This characteristic looks at the extent to which the PAD demonstrates collaboration across sectors and outlines mechanisms to align actions, resolve mandate overlaps, and avoid duplication.

Scoring and Assesment

Each characteristic is broken into two or three guiding questions to provide a focused evaluation of key elements. Scorers are encouraged to review specific sections of the PAD, such as the **Key Risks**, **Implementation Arrangements**, and **Stakeholder Engagement Plan**, to identify relevant content and provide informed ratings.

The protocol uses a simple scoring system to rate responses for each question:

- **3 = Thoroughly Addressed**: The PAD explicitly and comprehensively incorporates FCV-sensitive measures aligned with the question.
- **2 = Moderately Addressed**: The PAD adequately addresses the question but may lack depth or completeness.
- **1 = Weakly Addressed**: The PAD references the issue but in a limited, superficial, or indirect way.
- **0 = Not Addressed**: There is no evidence in the PAD that the issue has been addressed.

By aggregating scores across the five characteristics, this protocol provides an overall measure of a project's FCV sensitivity. Each question is broken down into a series of questions with a scoring system between 0 (Not Addressed) and 3 (Thoroughly Addressed) for each question. Equal weights are used for each question – though it's worth noting that some characteristics have more questions than others. The total scores across all questions are then added up cumulatively to provide a high-level assessment of conflict-sensitivity.

Scorers should carefully consider both explicit and indirect evidence of FCV sensitivity, as some aspects may be embedded in sectoral or contextual activities not immediately labeled as FCV-related. The provided guidance for each question includes core issues to consider, traits of high and low-scoring responses, and relevant sections of the PAD to focus on during the evaluation.

This structured approach ensures that projects are consistently evaluated, supporting efforts to enhance FCV-sensitive design and maximize project impact in fragile and conflict-affected settings.

Characteristic 1: Consider How Interactions Between Climate & FCV Affect Program Delivery

1. Does the PAD explicitly identify risks to project implementation from FCV-related barriers (e.g., security risks, institutional weaknesses, or strained community relations)?

Core Issues to Consider:
This question evaluates whether the PAD anticipates and addresses FCV-related risks that might disrupt the project. Relevant traits include identifying barriers such as insecurity, governance deficits, corruption, or social mistrust, which can impede delivery. Look for explicit acknowledgment of risks in areas like service delivery, stakeholder engagement, or operational access.

High vs. Low Scores:
A high score reflects clearly identified risks, with detailed mitigation strategies such as capacity-building for weak institutions or contingency plans for conflict-prone areas. A low score reflects minimal or superficial identification of FCV risks, with no evidence of how these might impact implementation or how they would be mitigated. **Key Sections to Review:** The **Key Risks** and **Implementation Arrangements** sections.

2. To what extent does the PAD seek to identify the specific pathways through which climate impacts interact with FCV dynamics?

Core Issues to Consider:
This question assesses whether the PAD explains how climate-related risks (e.g., floods, droughts, storms) exacerbate or intersect with drivers of fragility, conflict, and violence (FCV). Discussions should highlight how resource scarcity, governance challenges, or displacement may emerge or worsen due to climate impacts. The PAD should provide evidence-based, context-specific analysis.

High vs. Low Scores:
A high score reflects a detailed and localized analysis of climate-FCV interactions, linking specific climate risks to governance failures, social inequalities, or resource disputes. A low score indicates either a lack of analysis or only generic references to climate-FCV interactions. **Key Sections to Review:** The **Country Context** and **Sectoral and Institutional Context** sections.

Characteristic 2: Mitigate the Risk of Climate Actions Resulting in Maladaptation

1. Does the PAD incorporate specific safeguards to ensure project interventions do not exacerbate FCV-related vulnerabilities or create new sources of tension?

Core Issues to Consider:
Safeguards are essential to avoid unintended consequences, such as reinforcing inequalities, triggering disputes, or excluding marginalized groups. Strong safeguards might include grievance redress mechanisms, stakeholder engagement frameworks, or equity-focused interventions embedded in the project design.

High vs. Low Scores:
A high score reflects robust safeguards that are clearly articulated and tailored to the context, with mechanisms to prevent exacerbating FCV dynamics. A low score reflects the absence of safeguards or only generic provisions without actionable strategies. **Key Sections to Review:** The **Environmental and Social Management Framework (ESMF)** and **Key Risks** sections.

2. To what extent are adaptive mechanisms embedded into the project to accommodate evolving FCV conditions in the country or region?

Core Issues to Consider:
Adaptive mechanisms ensure the project can respond to changing FCV conditions, such as escalating tensions or shifts in governance capacity. Traits to look for include contingency funds, scenario-based planning, or flexible implementation arrangements that allow for mid-course corrections.

High vs. Low Scores:
A high score reflects well-defined adaptive mechanisms embedded into project design, with clear strategies for adjusting to FCV-related changes. A low score reflects rigidity in project design, with little

consideration for how evolving FCV dynamics might affect implementation. **Key Sections to Review:** The **Implementation Arrangements** and **Monitoring and Evaluation** sections.

3. Does the PAD show evidence of explicit efforts to balance immediate needs with long-term resilience-building in a way that avoids maladaptive outcomes?

Core Issues to Consider:
Balancing short-term relief with long-term resilience is critical to avoiding maladaptive outcomes, such as dependency or unsustainable practices. Relevant traits include incorporating climate-smart infrastructure, institutional capacity-building, and participatory planning that integrates resilience goals into immediate recovery efforts.

High vs. Low Scores:
A high score reflects a thoughtful balance, ensuring immediate actions (e.g., emergency housing) are designed to contribute to sustainable, long-term outcomes (e.g., disaster-resilient communities). A low score reflects a narrow focus on short-term objectives without considering long-term impacts. **Key Sections to Review:** The **Project Components** and **Sustainability** sections.

Characteristic 3: Prioritize Climate Actions That Address FCV Root Causes & Enhance Peacebuilding

1. Does the PAD include interventions that explicitly address root causes of FCV, such as inequitable access to resources or weak governance?

Core Issues to Consider:
This question evaluates whether the PAD includes interventions targeting structural FCV drivers, such as governance deficits, social exclusion, or resource competition. Examples include equitable resource allocation mechanisms, measures to strengthen governance, or addressing systemic inequalities.

High vs. Low Scores:
A high score reflects clear identification of FCV root causes and actionable measures to address them, such as implementing transparent land management systems or promoting equitable service delivery. A low score reflects vague or generic references to FCV root causes without actionable strategies. **Key Sections to Review:** The **Project Components** and **Lessons Learned** sections.

2. Does the project actively seek to promote peacebuilding, such as fostering trust, social cohesion, or conflict resolution?

Core Issues to Consider:
This question focuses on whether the project incorporates peacebuilding activities that foster trust, reduce tensions, or promote social cohesion. Relevant traits include community dialogue platforms, participatory governance structures, or conflict resolution mechanisms.

High vs. Low Scores:
A high score reflects intentional peacebuilding measures embedded into the project design, such as creating forums for dialogue or training community leaders in conflict resolution. A low score reflects the absence of peacebuilding components or vague references to community engagement without conflict resolution elements. **Key Sections to Review:** The **Stakeholder Engagement Plan (SEP)** and **Project Components**.

Characteristic 4: Prioritize the Needs and Capacities of Vulnerable Regions and Groups

1. Does the PAD explicitly identify vulnerable populations (e.g., women, displaced persons, minorities) and include measures to address their specific needs?

Core Issues to Consider:
This question assesses whether the PAD identifies and addresses the unique needs of vulnerable groups disproportionately affected by climate and FCV dynamics. Examples include gender-sensitive programs, targeting displaced populations, or addressing challenges faced by marginalized communities.

High vs. Low Scores:
A high score reflects clear identification of vulnerable populations and tailored interventions, such as ensuring equitable access to services or implementing livelihood programs. A low score reflects generic or absent references to these groups. **Key Sections to Review:** The **Project Beneficiaries** and **Environmental and Social Risk Analysis** sections.

2. Are mechanisms included to ensure equitable benefit-sharing and avoid reinforcing inequalities?

Core Issues to Consider:
This question evaluates whether the PAD incorporates mechanisms to ensure that project benefits are equitably distributed and do not exacerbate existing inequalities. Mechanisms might include transparent resource allocation, grievance redress systems, or targeted outreach to marginalized groups.

High vs. Low Scores:
A high score reflects well-defined mechanisms such as quotas for women's participation or frameworks for equitable resource distribution. A low score reflects the absence of mechanisms or insufficient attention to equity concerns. **Key Sections to Review:** The **Environmental and Social Standards (ESS)** and **Project Components** sections.

Characteristic 5: Encourage Coordination Across Development, DRM, & Peacebuilding Actors

1. Does the PAD demonstrate evidence of active collaboration with stakeholders across sectors (e.g., humanitarian, peacebuilding, disaster risk management)?

Core Issues to Consider:
This question examines whether the project actively seeks collaboration with stakeholders from multiple sectors to address overlapping climate and FCV challenges. Effective coordination might involve partnerships with local NGOs, international agencies, or government entities.

High vs. Low Scores:
A high score reflects clear multi-sectoral collaboration, such as joint planning with humanitarian organizations or aligning efforts with national DRM strategies. A low score indicates limited or vague references to potential partnerships. **Key Sections to Review:** The **Implementation Arrangements** and **Stakeholder Engagement Plan (SEP)** sections.

2. Does the PAD outline mechanisms to align actions, resolve mandate overlaps, and avoid duplication across relevant actors?

Core Issues to Consider:
This question assesses whether the PAD includes systems to ensure alignment of project activities with those of other stakeholders, addressing potential conflicts or overlaps. Mechanisms could include coordination bodies, clear governance structures, or formal agreements.

High vs. Low Scores:
A high score reflects robust alignment mechanisms, such as cross-sectoral working groups or shared implementation frameworks. A low score reflects a lack of coordination strategies or attention to mandate conflicts. **Key Sections to Review:** The **Key Risks**, **Implementation Arrangements**, and **Project Components** sections.

Process for Unbiased Manual Annotation

To ensure neutrality, objectivity, and consistency in evaluating FCV sensitivity, the following structured protocol will be implemented:

Step 1: Annotator Training

Annotators must develop a clear understanding of the FCV framework, guided by:

- The **FCV-sensitive Climate Action Report** as a foundational reference.
- **Sample annotated PADs** to illustrate the application of the framework.
- **Discussion of ambiguous cases** to refine assessment guidelines and enhance scoring consistency.

Step 2: Guided Annotation Process

Each characteristic is evaluated based on two or three guiding questions that focus on key elements of FCV sensitivity. Annotators are encouraged to review specific sections of the PAD—such as **Key Risks,**

Implementation Arrangements, and Stakeholder Engagement Plan—to identify relevant content and provide informed ratings.

Annotations will follow a structured scoring system:

- **3 = Thoroughly Addressed:** The PAD explicitly and comprehensively incorporates FCV-sensitive measures aligned with the question.
- **2 = Moderately Addressed:** The PAD adequately addresses the question but may lack depth or completeness.
- **1 = Weakly Addressed:** The PAD references the issue but in a limited, superficial, or indirect way.
- **0 = Not Addressed:** There is no evidence in the PAD that the issue has been addressed.

Equal weights are assigned to each question; however, some characteristics may have more guiding questions than others. The total scores across all questions are aggregated to provide an overall measure of a project's FCV sensitivity.

Step 3: Blind Double-Annotation & Discrepancy Resolution

- Each document is independently annotated by two reviewers.
- Annotators do not see each other's scores to prevent bias.
- If significant discrepancies arise, a third reviewer will be assigned to resolve differences and establish a final consensus rating.

Step 4: Standardization & Consistency Review

To ensure the reliability of annotations:

- A **random sample of 10%** of PADs will be rechecked for consistency.
- Adjustments to scoring criteria will be made if systematic inconsistencies emerge.
- An **annotation handbook** will be maintained, providing detailed guidance on scoring rationales and clarifications for ambiguous cases.

Step 5: Cross-Validation with LLM Outputs

Once the LLM is trained to classify PADs based on FCV sensitivity:

- The model will annotate the full dataset.
- A subset of **10% of LLM-classified PADs** will undergo manual review to assess agreement.
- Confidence in classification will be evaluated through **log probability (logprob) distribution analysis**:
  - **If randomness is NOT allowed:** A threshold will be established based on the overlap in "yes" and "no" logprobs.

- o **If randomness is allowed:** Bootstrapping will generate a confidence interval for each response, and manual review will be conducted when confidence intervals for "yes" and "no" overlap, indicating uncertainty in classification.

This structured approach ensures that projects are consistently evaluated, supporting efforts to enhance FCV-sensitive design and maximize project impact in fragile and conflict-affected settings.

Ensuring Unbiased LLM Classification

- **Train the model** on **balanced data**, ensuring representation of:

  - o Various **geographical regions**.

  - o Different **types of climate projects**.

  - o A mix of **high, medium, and low FCV-sensitive projects**.

- **Monitor model drift** by periodically **re-checking** LLM-classified PADs.

- **Refine classification if needed**, using iterative fine-tuning.