REPORT:
TOWARDS MEANINGFUL AND MEASURABLE INDICATORS FOR
THE GIRLS NOT BRIDES (GNB) PARTNERSHIP

Aspen Planning and Evaluation Program
The Aspen Institute

August 29, 2014
Contents

Executive Summary.................................................................................................................................................. 1
Introduction ............................................................................................................................................................ 4
Rationale ............................................................................................................................................................... 4
Key Challenges and Technical Considerations .................................................................................................... 5
Methodology: Organizing and Ranking the Indicators ......................................................................................... 10
Explanatory Notes.................................................................................................................................................. 13
Recommendations for Next Steps ....................................................................................................................... 14
List of Consulted Sources .................................................................................................................................... 19
Executive Summary

*Girls Not Brides* (GNB) retained the Aspen Institute’s Aspen Planning and Evaluation Program (APEP) to help inform the secretariat’s understanding of indicators that organisations are using to measure progress toward ending child marriage, suggest useful indicators, and identify instances where better indicators are needed. This report, along with its companion materials, represents the first step in the process of developing a shared framework of indicators that will help the secretariat, national partnerships, and member organisations focus their work effectively and facilitate learning about both successes and challenges. More specifically, this framework will enhance the GNB partnership’s capacity to: build consensus among its 359 members; define what members should measure; demonstrate members’ impact; advocate for governments to take concrete actions; map and drive investment; define precise asks for data collection; and facilitate engagement with development frameworks like the forthcoming UN post-2015 goals.

The GNB secretariat gathered lists of indicators used by member organisations, United Nations and government agencies, as well as related reports and other documents suggesting assessment approaches and potential benchmarks. APEP used these materials to develop a *Preliminary List of Recommended Indicators* that map onto or complement GNB’s theory of change (ToC). APEP also delivered a *Report* that discusses gaps in available indicators, key challenges in measuring progress toward ending child marriage, and recommendations for next steps. To help ensure that the list of indicators is of practical use to GNB member organisations, APEP developed a *User’s Guide* containing examples, helpful tips about selecting and using indicators, and useful data sources.

HIGHLIGHTS

**Key Challenges and Technical Considerations**

Developing a shared framework of indicators for GNB’s diverse partnership is a complex task. The report discusses a number of key challenges in identifying appropriate indicators:

- **Balancing breadth vs. utility:** For a partnership as diverse as GNB, a key challenge is defining a practical scope for indicators while also reflecting members’ varied activities and impacts.

- **Discerning consensus:** Some indicators are commonly used (e.g., adolescent birth rate, the percentage of women ages 20-24 married before age 18). But the diversity of organisations and contexts in this field makes it difficult to build consensus around other key indicators.

- **Assessing the utility of indicators across contexts:** Indicators can be useful in some countries and even some sub-national regions or locales, but inapplicable or impractical in others due to differences in cultural factors and data availability.

- **Weighing new versus established indicators:** Well-known indicators typically come with the advantages of ample field testing and publicly available data. But newer or less well-known indicators can provide a finer-grained focus on specific issues relevant to child marriage and married girls.

- **Incorporating multiple levels of indicators:** GNB members’ work spans multiple levels, from individual-level change to international policy change. The larger the geographic area and target
population, however, the more difficult it is to isolate the unique contribution of a single organisation. APEP encourages members to pick a target population or target audience – and corresponding indicator – that makes it as easy as possible to identify their special contribution to change.

- **Capturing normative change:** Researchers and practitioners working to end child marriage have been hindered by a weak conceptual framework for defining and measuring different types of norms, such as descriptive and prescriptive social norms. The report and the list of indicators offer some background.

- **Distinguishing between indicators for program evaluation vs. program planning:** Indicators can serve *evaluation* purposes (helping track progress toward expected or desired outcomes), as well as *planning and advocacy* purposes (helping organisations prioritize, develop messages, or shift programmatic content).

**Preliminary List of Recommended Indicators**

APEP used GNB’s ToC documents to guide our analysis of the hundreds of indicators gathered by the secretariat and by APEP. We mapped indicators onto specific components of the ToC and ranked them by the extent to which they were:

- **Meaningful:** a valid, precise, and non-directional indicator of a given component in the ToC;
- **Measurable:** a feasible indicator to measure using available tools, methods, and data; and
- **Field tested:** a reliable indicator that has been tested and/or recommended by organisations within and outside the GNB partnership.

The resulting list of recommended indicators focuses primarily on the nineteen Outcomes and thirteen Results laid out in the ToC. These represent the most directly measurable components of the ToC that can be used to *evaluate* members’ progress and impact. The list also includes a brief section on measures useful for *planning and advocacy* purposes, as well as indicators for tracking progress on the ultimate goal of ending child marriage, captured by the *prevalence of child marriage*.

**Recommended Next Steps**

We offer some recommendations about next steps that GNB might take as it continues this valuable process of developing a framework of indicators to focus its efforts and record its progress. These steps overlap and interact with one another, so GNB can take complementary actions to address multiple recommendations at once.

- **Address gaps in the list of indicators:** We recommend that GNB solicit its members’ suggestions for ToC components that lack strong indicators. The list of indicators also draws attention to these gaps, explicitly encouraging members to share suggestions for potential indicators and data sources.

- **Clarify ToC language to make it easier to identify appropriate indicators:** We suggest that GNB organize one or more facilitated discussions to identify more precise focal points within some of the more broadly worded Outcomes and Results in the ToC. That will make it easier for the partnership to identify a select set of relevant indicators.
Facilitate a workshop discussion among GNB members to build consensus: Building on the above recommendation, GNB may wish to conduct a facilitated workshop to generate agreement among members around indicators that closely match their work, expanding consensus beyond the well-established indicators used across the broader development domain.

Work towards aligning the ToC and indicators framework: The GNB ToC documents represent a significant milestone in the partnership’s work. However, we discovered certain gaps in the ToC that GNB may wish to address. The ToC is a useful conceptual framework against which to map relevant indicators. We encourage GNB to “synch up” the ToC and relevant indicators, perhaps making modest adjustments to address gaps or tighten ambiguities in the ToC.

Consider how to incorporate indicators for capacity building, progress in policy advocacy, and social norms: APEP recommends including these indicators to allow members to track and demonstrate their progress in areas that have been overlooked, under-valued, or (in the case of social norms) lacked sufficient conceptual clarity.

Develop a time horizon for some outcomes and results: We suggest providing a realistic timeframe for certain Outcomes and Results, such as changes in norms, policies, and behaviours that may take years to achieve. Adding a timeframe to the ToC can help GNB members more accurately judge which indicators will yield useful information at a given point in the longer process toward desired changes.
Introduction

*Girls Not Brides* (GNB) retained the Aspen Institute’s Aspen Planning and Evaluation Program (APEP) to help inform the secretariat’s understanding of indicators that organisations are using to measure progress toward ending child marriage, suggest useful indicators, and identify instances where better indicators are needed. The ultimate goal is to develop a framework of indicators that will help the secretariat, national partnerships, and member organisations focus their work effectively and facilitate learning about both successes and challenges. The framework will also provide a way to standardize how impact is assessed, making it easier to aggregate impacts across member organisations and generate a larger picture of progress toward end goals.

This report represents the first step in the process of framework development. The GNB secretariat gathered lists of indicators used by member organisations, United Nations and government agencies, as well as related reports and other documents suggesting assessment approaches and potential benchmarks. The purpose of this report is to use those materials to develop an initial set of recommended indicators that map onto or complement GNB’s theory of change (ToC), highlight gaps in available indicators, and discuss key challenges in measuring progress toward ending child marriage.

This report includes the following:

- A brief rationale for the framework of indicators;
- A discussion of key challenges and technical considerations surrounding the measurement of progress toward ending child marriage;
- An explanation of our methodology for prioritizing among hundreds of potential indicators;
- Some explanatory notes to contextualize our preliminary recommendations for indicators; and
- A set of recommendations for next steps that GNB may consider taking as it moves forward with this valuable effort to develop a framework of indicators.

Accompanying this report is the Preliminary List of Recommended Indicators and a User’s Guide to help member organisations make use of the indicators. For reference purposes, a List of Sources consulted for this work is included at the end of this report.

Rationale

Based on conversations with the GNB secretariat and an informal memo prepared by *Girls Not Brides* USA,¹ a strong rationale for developing a shared framework of indicators emerged. The framework would serve a number of important purposes, strengthening the ability of GNB and its member organisations to:

- Facilitate engagement in a concrete way with other development frameworks, such as the United Nations post-2015 Sustainable Development Goals, and with donors investing in ending child marriage;²

---

¹ *Girls Not Brides* USA, “Memo re: Indicators to measure our progress toward ending child marriage,” from Meg Greene to Lyric Thompson, Carol Boender, Erin Kennedy, Milkah Kihuna, Doris Bartel, Lena Minchew, Nidal Karim, Jacqueline Hart, Ellen Travers, Ann Warner (June 23, 2014). (Document received from Girls Not Brides, July 2014.)

² For example, *Girls Not Brides* USA is engaging on the U.S. government’s strategy for adolescent girls and child marriage.
• Build consensus and focus efforts among organisations working to address child marriage;
• Define what members should be measuring;
• Demonstrate members’ impact;
• Advocate for governments to take concrete actions to track and address child marriage;
• Map and drive investment in efforts to end child marriage; and
• Define precise asks for data collection around indicators in the future.

This rationale for the value and utility of this framework informed our approach to organizing and prioritizing indicators, as well as our suggestions for next steps in this process.

Key Challenges and Technical Considerations

Developing a shared framework of indicators is a complex task. Here we describe seven key challenges in identifying appropriate indicators, along with their accompanying technical considerations.

1. Balancing Breadth versus Utility

The evaluation frameworks and documents gathered by the GNB secretariat and member organisations reflect their holistic understanding of child marriage and the challenges facing married girls. This comprehensive view is also evident in GNB’s ToC. A vast number of interlocking factors affect child marriage and the well-being of married girls. Potential indicators span health, economic, educational, and socio-cultural domains, and might track changes in everything from individual behavior to international policy. This report could attempt to encompass the full spectrum of relevant indicators of incremental progress toward the goals of ending child marriage and empowering married girls. Such a report would offer the requisite programmatic breadth, but at the expense of clarity and utility.

In articulating its own theory of change and results framework for ending child marriage, UNICEF underscores this concern:

The risk here is that development actors could try to comprehensively define and affect actions that in their view may result in a change in child marriage and try to intervene in whatever seems most compelling, without assessing what is within their purview of influence. To minimise this risk, the theory of change outlined in Figure 1 depicts only strategic and viable entry points for development programming—and especially ones where UNICEF in particular has strategic and added value. It thus avoids laying out a vast range of influencing factors and defining a comprehensive array of change pathways that may further complicate the argument and ensuing interventions.³

GNB faces the challenge of identifying a broad and diverse enough set of indicators to accommodate the numerous strategies and target outcomes on which organisations and agencies focus without including so many indicators that the framework becomes unwieldy and impractical. To do that, we need to clearly define what falls within – and perhaps even more important, what falls outside of – GNB’s purview. Otherwise, the number of potentially relevant indicators becomes nearly limitless.

2. Discerning Consensus among Diverse Actors

Ideally, the GNB framework of indicators would include indicators that experts and practitioners tend to agree are valid, useful, feasible, and so forth – that is, indicators around which there is some level of consensus. In this way, the framework would reflect the combined expertise, experience, and evaluation tools of organisations working to end child marriage and empower married girls. GNB has made an effort to identify such “consensus” indicators; soliciting input from its diverse member organisations and suggesting indicators and evaluation frameworks used by key United Nations and government agencies could reveal which indicators are used by multiple sources.

Our review of the documents provided by GNB (see the List of Consulted Sources) suggests some consensus around aggregate-level indicators used to document progress on global development goals, including the prevalence of child marriage, girls’ educational attainment, the adolescent fertility rate, unmet need for family planning, attitudes toward wife beating, and the prevalence of violence against women. These indicators fall into the more “established” category of indicators (discussed below). Organisations may gravitate toward these indicators because they are well-known and field tested. UN and government agencies also collect and provide national-level data for these indicators, making them easier to use. Some offer the potential for disaggregation to the sub-national level. These indicators may also be popular simply because they are obviously relevant to child marriage and the welfare of married girls.

Finding consensus around other indicators is difficult for several reasons. First, although APEP consulted many sources for this report, relatively few of GNB’s member organisations provided lists of the indicators they use. Therefore, the sample of members was too small to draw conclusions about any consensus within the GNB partnership. Second, the diversity of GNB members and other organisations working in related domains makes it difficult to identify consensus around indicators specific to ending child marriage and empowering married girls. For example, some organisations focus exclusively on child marriage, while others have a broader mandate that tangentially overlaps with child marriage. These two types of organisations may have different programmatic and evaluation priorities and therefore perceive different indicators as relevant. Even where two different organisations are using the same indicator, their “consensus” on the indicator may be predicated on different uses. For example, the Canadian Department of Foreign Affairs, Trade and Development and the United Nations Population Fund (UNFPA) may use the adolescent birth rate as an indicator for evaluating quite different programs.

The varied size and resources of member organisations pose other challenges to consensus around indicators. Smaller community-based NGOs may not have the necessary resources to develop and undertake evaluation activities, at least not on a continuous basis. Indeed, a 2011 report by the International Center for Research on Women (ICRW) found that relatively few evaluations of child marriage projects have been conducted (and these evaluations varied in the degree of rigor). In order to build or reflect consensus among NGOs working in the field, one would need to consult those who have not yet developed indicators. We suspect that there has been no coordinated effort to evaluate the relative strength of – and consensus around – different indicators that may be used to measure progress on ending child marriage. GNB has an opportunity to make a substantial and unique contribution to the field by building consensus among those working to end child marriage and empower married girls.

---

3. Assessing the Utility of Indicators across Contexts

A third key challenge revolves around identifying indicators – and their associated methodological approaches for gathering data – that are appropriate across countries, cultures, and contexts. Perhaps most obvious, cultural differences render some indicators useless (or at least highly unreliable and unproductive) in certain countries. For example, it may be possible to ask directly about certain sexual or reproductive health behaviors in one country, but entirely unacceptable or unfeasible in another country, necessitating the use of indirect or proxy measures.

Differences in the distribution of target populations could also mean that an indicator is useful in one context, but less so in another. For example, using a survey of a random sample of the population to assess female genital mutilation may make sense in countries where the incidence is high, but it will be less useful in countries where incidence is relatively limited or concentrated in particular areas.

Countries also vary in terms of the availability of reliable data. For example, the quality of (and access to) data on the number of gender-related deaths (e.g., police and court records, medical records, news reports) is likely better in some contexts than in others.

Finally, at a very simple, instrumental level, some indicators may be favored by one donor or by an influential GNB partner in a given country.

4. Weighing New versus Established Indicators

A fourth challenge involves weighing the pros and cons of well-established indicators – data for which are often collected by governmental agencies – versus newer or less widely known indicators that may have only been field tested on a limited basis or not at all. Well-established indicators, such as prevalence of contraceptive use, educational attainment, and attitudes toward wife beating, have a major advantage: data on these indicators are regularly collected by governmental agencies or NGOs and made publicly available (e.g., UNICEF Multiple Indicators Cluster Surveys, Demographic and Health Surveys, UNESCO Institute for Statistics). This eliminates the often heavy costs associated with primary data collection. These indicators have also typically undergone close scrutiny with regards to their reliability and validity, providing a strong case for methodological rigor.

On the other hand, as some GNB members point out, these large publicly available datasets have limitations. For example, they are not available in some countries with a high prevalence of child marriage, such as Afghanistan and other countries in the Middle East, as well as high-income countries. Large publicly available survey datasets like UNICEF MICS and DHS also do not consider important contextual factors, such as consent to marriage, dowry and bride price, and social norms. These surveys also tend to be biased toward women of reproductive age (defined as 15-49). This means that girls who are married before age 15 are not asked questions about, for example, domestic violence or household decision making.

The 2011 ICRW report cited above adds these cautionary notes:

... governments and multi-lateral agencies undertaking large-scale initiatives to eradicate poverty, expand economic opportunity, promote women’s empowerment and promote girls’ education have the potential to prevent child marriage because they are, in effect, addressing
some of its drivers. However, these broader initiatives have not always been developed with child marriage in mind, and therefore do not collect relevant data.

At the other end of the spectrum are less widely known indicators that have been developed in specific contexts. Some of these indicators offer a way to tap key elements, such as social norms or beliefs about dowry and bride price, that are specifically important to ending child marriage but are missed by broader, well-established indicators. The drawback: these less established indicators have rarely been field tested in multiple contexts or undergone tests of reliability and validity. Thus, their ability to capture meaningful change – and to do so across different settings – remains unknown.

5. Incorporating Multiple Levels of Indicators

Organisations seeking to end child marriage and support married girls can have impacts at multiple levels, from individual-level change to international policy change. Both the conceptual framework provided by the ToC and GNB members’ activities are arrayed across these levels, pointing to the need for indicators that are similarly diverse in scope.

A key consideration when determining the appropriate level of indicator to use is the trade-off between the scope of the impact and the degree of difficulty in determining the precise contribution that an individual organisation made in bringing about a change. In general, the larger the geographic area and target population, the more difficult it is to isolate the unique contribution of a single organisation. This is particularly true in “crowded” fields where numerous actors are simultaneously working toward similar or broadly overlapping goals. For example, if the national prevalence of child marriage in India were to drop by 20% over two years, it would be extremely difficult for an individual organisation to determine its particular contribution toward that change. Myriad other actors and dynamic contextual factors over time would also have contributed toward the observed change. On the other hand, if an organisation works closely with a group of individuals in a small community and then observes changes in their target attitudes or behaviors, it is more feasible to identify how the organisation’s activities may have helped to produce this change.

A similar trade-off exists for organisations advocating for policy change. It is often (though not always) easier to trace a line between the activities of a given organisation and a policy outcome when working at a local level. It is still difficult to do so, of course; but because the number of actors and influences shaping policymaking processes at a local level tend to be smaller than at a national, regional, or international level, the task of identifying one’s contribution can be relatively more feasible.

This is an important point for member organisations to keep in mind when they select potential indicators with which to measure their impact. In some cases, indicators are clearly located on a particular level. For example, an indicator that focuses on passage of national legislation requiring the free and full consent for marriage of both female and male parties will necessarily be a national-level indicator. Indicators that draw on data sources such as the UIS, UNICEF MICS, or DHS will also primarily be focused on the national level, though some degree of disaggregation by geographic region may be available in certain datasets (or primary data collection for a local target population can be independently conducted).

Other indicators are somewhat more flexible in their application across levels. For example, indicators that call for a survey of the target population could be applied at a program level (administered to program participants), a community level (e.g., within an intervention community), at a state or regional
level, or at a national level, depending on the scope of an organisation’s target population. Defining the scope of the target population and/or geographic area is a critical first step for member organisations when they are deciding which of the recommended indicators will most effectively help them document progress toward their goals.

6. Capturing Normative Change

Normative change represents an important objective for many organisations working to end child marriage and improve the lives of married girls. There are relatively well-developed theoretical frameworks for conceptualizing social norms, emerging from fields such as social psychology and communications. But research on child marriage has largely failed to connect with a strong theoretical foundation (Bicchieri et al. 2014). As a result, researchers and practitioners have been hampered in their ability to effectively measure social norms and assess the impact of efforts to change norms over time. What’s more, normative change (or inertia) is shaped by many different kinds of forces, making it difficult to tease out the contribution that a single organisation may have made.

When thinking about how to measure normative change, it is very useful to distinguish among different types of norms. There are three types that are particularly salient to child marriage. A descriptive norm describes what we think everyone else (or at least a majority of our community or social network) is doing. It is really people’s perceptions of the descriptive norm that matter here; when a person thinks that most other people engage in a particular behavior, that perceived descriptive norm can then influence the person’s own propensity to engage in that behavior. For example, if a father thinks that all the other fathers in his community are marrying off their daughters at a young age, he may be more likely to marry off his own daughter early.

A prescriptive norm describes what we think everyone else (or at least a majority) thinks should be done. This is distinct from what we think everyone else is actually doing. Again, it is people’s perceptions of the prescriptive norm that matters most; when a person thinks that most others believe people should engage in a particular behavior, that perceived prescriptive norm can then influence their own propensity to engage in that behavior. Returning to the same example, if a father thinks that all the other fathers in his community think that daughters should be married off at a young age, he may be more likely to marry off his own daughter early. Descriptive and prescriptive norms are distinct from a personal norm, which refers to what you personally think you should do, regardless of what others think or are doing.

Note that descriptive and prescriptive norms both have a “social” component – they are contingent on what one believes that others in one’s social network or environment do or think should be done. Descriptive and prescriptive norms can therefore be thought of as two types of “social norms.” Personal norms are not contingent on what others do or think, and therefore are not “social” in that sense. Of course, if the same personal norm is held by a majority of people in a community, this would likely be reflected in individuals’ perceptions of the descriptive and prescriptive norms.

7. Distinguishing between Indicators for Program Evaluation and Program Planning

This report and the accompanying preliminary list of indicators focus primarily on indicators useful for program evaluation and strategic learning. These indicators are intended to help users track their progress towards the outcomes and results in GNB’s ToC. GNB also asked APEP to recommend indicators that members and national partnerships can use to track changes in relevant trends at the
national level as well. It would be difficult to demonstrate the specific contribution of GNB members to changes in these national trends, as we discussed above. They are therefore less useful for evaluation purposes. But these national level indicators can serve a different purpose: program planning or advocacy. Here the purpose of indicators is not to measure members’ own impact, but to track changes that are relevant to making the case for government or donor action.

For example, national-level data on child marriage, girls’ educational attainment, women’s labor force participation, and other indicators of female empowerment can help advocates communicate the need for greater action on ending child marriage or empowering married girls. Member organisations may also want to use these broader indicators for purposes of program planning, using the data to help inform program priorities or strategies. We encourage GNB members to distinguish between their use of indicators for evaluation purposes (e.g., measuring the behavioral impact of a specific project or intervention) and their use of indicators for planning and advocacy.

We emphasize that indicators are useful for answering some, but not all, questions that the GNB partnership or its stakeholders might have. Used well, indicators provide useful information about progress made toward expected or desired results, and can help determine whether a shift in strategy or programmatic content may be needed. Indicators are less helpful for answering questions about why results were or were not achieved, or why unintended results arose. And they can only partially point the way toward actions that may need to be taken to improve results. Perhaps most important, indicators should not be used as definitive assessments of program success or failure; rather, they are usually best used to chart progress and to highlight areas that may need adjustment.

**Methodology: Organizing and Ranking the Indicators**

The GNB secretariat gathered lists of indicators used by national partnerships, member organizations, UN agencies, and government agencies, as well as related reports and other documents that suggest assessment approaches and potential benchmarks. The APEP team used GNB’s ToC graphic and the accompanying ToC brief to guide our analysis of the hundreds of indicators gathered in this way. Using a set of criteria (described below), indicators were mapped onto specific components of the ToC and ranked in terms of strength, with those we judged to be the strongest ultimately comprising the preliminary list of recommended indicators.

This report and accompanying list of indicators focus primarily on two levels of the ToC: **Outcomes** and **Results**. The nineteen outcomes and thirteen results laid out in the ToC represent the most directly measurable components of the ToC for assessing progress toward ending child marriage and empowering girls. Our preliminary list of indicators also includes a brief section on measures potentially useful for **planning and advocacy** purposes. As noted above, using indicators this way is different from using them to evaluate an organisation’s own progress, but it may help a partnership or member organisation set programmatic priorities or craft effective advocacy messages. We also include indicators for tracking progress on the ultimate goal of ending child marriage, as measured by the **prevalence of child marriage**. These indicators can capture progress towards GNB’s ultimate vision of eliminating child marriage at the global level. But they are also quite practical indicators for measuring progress at the program, community, district or regional level, or for advocacy purposes.

Other levels of the ToC are largely beyond the scope of this report. For example, the broader Vision and Impact articulated in the ToC represent the “North Star” (or for that matter, the Southern Cross) that
provides the ultimate destination towards which all members are helping to steer society. Arriving there will require a combination of efforts and interactions well beyond those of even this strong, diverse, and deeply committed partnership of organizations. We do not attempt to suggest indicators for the broadly stated Catalyzing Strategy, which encompasses an array of interlocking issue domains and levels of change that are beyond the scope of GNB’s activities. Similarly, we have not included indicators for the Strategies level of the ToC, as interventions and approaches are likely to vary substantially across organisations and contexts. All of these are appropriately included in the ToC as a reminder of what GNB seeks to achieve and how it intends to achieve it. But the predominant focus of our report on realistic indicators for assessing members’ progress and impact will, we hope, help GNB and its partners ultimately achieve the clarity, precision, and consensus in measurement needed to advance evaluation of their work.

Based on the ToC brief, we interpreted the Outcomes to refer to the attitudes/beliefs and environmental conditions (e.g., services, training/protocols, policies) that serve as key antecedents of the Results, which focus on behaviors. As articulated in the ToC brief:

The outcomes reflect the changes we hope to see as a consequence of our strategic activities. They reflect changes needed in the attitudes and conditions of girls, families and communities, services and laws and policies...

The results (or intermediate impacts) reflect the demonstrated changes in behaviour on the part of individuals, families and communities, programme implementers, and law enforcement officials as a consequence of surrounding changes in attitudes and conditions. The results are deliberately framed in an active manner, to emphasise whose behaviour has changed.

APEP has organized the indicators to align with this understanding of the various components of the ToC, pairing each indicator with the ToC component that we judged the indicator to best measure. Some of the components of the ToC are not mutually exclusive, so there are instances in which an indicator could reasonably be understood as capturing progress toward more than one component.

To identify appropriate criteria for ranking and selecting indicators, we reviewed some general guidance from measurement reports produced by major agencies in relevant fields. A United Nations report on indicators for measuring violence against women (VAW) offered a particularly salient set of recommendations, stating that indicators should:

- Summarize complex data;
- Be unambiguous and easy to interpret;
- Enable an assessment as to whether an improvement or deterioration has occurred;
- Be meaningful and relevant to policy makers, service providers and the wider public;
- Be capable of being supported by reliable and robust quantitative data;
- Be neither so many as to confuse, nor so few as to mislead;
- Be available at regular intervals and be comparable over time; and
- Be comparable between countries and population groups.

Since data collection efforts are costly and complex, the UN report also recommended that indicators should as much as possible:
- Use available data;
- Provide consistency in the use of the time period covered, and include both a longer period and a more recent period; and
- Ensure consistent identification of the same population sub-set.\(^5\)

Additional useful guidance was obtained from a compendium of VAW indicators produced by MEASURE Evaluation, a group of organisations funded by USAID and the U.S. President’s Emergency Plan for AIDS Relief (PEPFAR). Researchers are encouraged to identify indicators that are:

- **Valid**: Indicators should measure the aspects of the program that they are intended to measure.
- **Specific**: Indicators should only measure the aspect of the program that they are intended to measure.
- **Reliable**: Indicators should minimize measurement error and should produce the same results consistently over time, regardless of the observer or respondent.
- **Comparable**: Indicators should use comparable units and denominators that will enable an increased understanding of impact or effectiveness across different population groups or program approaches.
- **Non-directional**: Indicators should be developed to allow change in any direction, and not specify a direction in their wording (for example: an indicator should be worded as “the level of awareness” instead of “an increased awareness”).
- **Precise**: Indicators should have clear, well-specified definitions.
- **Feasible**: It must be possible to measure an indicator using available tools and methods.
- **Programmatically relevant**: Indicators should be specifically linked to a programmatic input, output or outcome.\(^6\)

These useful tips helped us refine our own selection of recommended indicators. We evaluated and ranked indicators in terms of the extent to which they are:

1. **Meaningful**: a valid, precise, and non-directional indicator of a given component in the ToC;
2. **Measurable**: a feasible indicator to measure using available tools, methods, and data;
3. **Field tested**: a reliable indicator that has been tested and/or recommended by organisations within and outside the GNB partnership.

We weighted the first criteria most heavily. An indicator is only useful if it can meaningfully capture change (either positive or negative) in a specific target outcome. The second criterion – how measurable an indicator is – encompasses a number of elements, including whether data on an indicator are collected and made publicly available by agencies and organisations, and the degree of difficulty in collecting data (e.g., through surveys, interviews, mapping, desk research) and obtaining information (e.g., from officials, agencies, service providers, community members).


We use the extent of field testing to help determine the reliability and cross-cultural applicability of indicators. Indicators that are being used by multiple organisations or in varying contexts (including standard metrics collected by UN agencies) are more likely to have well-established reliability and cross-cultural applicability, compared to indicators that have been used in a more limited set of contexts or have not yet been tested in the field. This criterion helped guide APEP’s choices when there were many possible indicators for an outcome, or there were multiple variations of an indicator. However, keeping in mind the value of balancing new versus established indicators (noted above), our preliminary list of recommended indicators does include some cases that we believe strongly meet our first two selection criteria, but could benefit from additional field testing or refinement in order to meet this third criterion.

In some cases, we had a long list of indicators mapped onto a single ToC component. For example, there were many indicators for VAW, female genital mutilation/cutting (FGM/C), avoidance of early pregnancy and refusal of sex, and a legal framework for protecting girls’ and women’s rights, among others. This is partly a function of conceptual breadth (the components are broadly worded), and partly a function of how much is already being done in the issue domain (e.g., VAW has received considerable international attention). There were also multiple variations of certain indicators – that is, different sources offered slightly different versions of essentially the same indicator. In these cases, we included the version that best met all three of our criteria.

Based on the criteria, we judged each indicator to be strong, moderate, or weak. The strongest choices appear in our preliminary list of recommended indicators. We provided an earlier draft list with an exhaustive list of potential indicators, ranked by strength, for reference.

**Explanatory Notes**

As seen in the accompanying list of recommended indicators, the nineteen Outcomes and thirteen Results are organized according to the ToC, grouped into components for Girls, Families and Communities, Services, and Law and Policy. A table of contents provides a guide for navigating the list. For each indicator, we aim to identify an appropriate data source that would allow users to measure progress against the indicator. Where applicable, we also provide brief notes on methodological details likely to be useful to users.

The number of indicators for each outcome or result varies. In some cases, we offer three or more indicators; in others, only one or two. The higher number of indicators for certain components is largely a function of their conceptual breadth; a few were so broad that they had to be broken down into sub-components. At the other end of the continuum, there are ToC components for which we found few or no relevant indicators. This seems to be partly a function of the specific language used in the ToC; in some places, the language made it difficult to identify appropriate indicators. We also recognize, however, that a dearth of indicators for certain components could also reflect either the limits of our pool of indicators or gaps in prior evaluation research. Where applicable and appropriate, GNB and APEP suggested a few indicators (either new ones or adaptations of existing ones) and offered tentative suggestions for data sources. We believe that attempts to fill these gaps should be guided by member input; all of these cases are highlighted in the list of indicators in order to encourage suggestions for potential additional indicators and/or data sources.

We note that not all of the indicators emerging from the collective suggestions of GNB partners map onto the ToC, in our judgment. As noted earlier, some of the broader indicators are not typically useful
for evaluation purposes. But they may help national partnerships and members plan, prioritize, or advocate effectively. We have included selected indicators in this category. For example, we suggest that the United Nations Development Programme’s Gender Inequality Index and its recently introduced Gender Development Index (or the component indicators that comprise these indices) may be helpful for making a broader case for donor or government action.

There are also some indicators that appear to be highly relevant to assessing progress toward ending child marriage, but are not very precise or direct measures of specific results or outcomes in the ToC. In particular, indicators measuring social norms, capacity building, and steps in the policy change process did not map well onto the ToC, yet are central to many organisations’ efforts. In consultation with GNB, we sought to incorporate these relevant indicators, either as proxies in the list of indicators or as examples included in the User’s Guide. Additional options for incorporating these indicators are provided in the recommendations for next steps.

Two final points merit emphasis. First, some of the indicators may be understood as measuring “outputs” more than “impacts”—that is, they assess an organisation’s progress on implementing activities more than they measure its influence on others’ attitudes or behaviors. This appropriately reflects the fact that some Outcomes focus on putting into place the necessary environmental conditions for attitude or behavior change, consistent with how Outcomes are described in the ToC brief cited above. For example, by documenting the number of girls’ groups, the percentage of girls enrolled in asset-building programs, or the number of key stakeholders reached by campaign messages, member organisations are able to measure progress toward creating the conditions for changing attitudes and behaviors.

Finally, we were unable to draw many strong inferences about consensus among GNB partners, in part because relatively few members submitted the indicators they use. As previously discussed, the greatest consensus seems to lie around indicators that tap global development goals or have publicly available data sets. For the prevalence of child marriage in particular, there was fairly strong consensus around the indicator measuring rates of child marriage (before age 18) among women ages 20-24. This indicator is included in the list, along with two others that were somewhat less widely cited but still highly relevant for assessing the prevalence of child marriage.

**Recommendations for Next Steps**

GNB has made valuable progress by defining its ToC and by taking the step reflected in this report – that is, seeking to define shared indicators. We offer some recommendations about next steps that GNB might take as it continues this process of developing a framework of indicators to focus its efforts and record its progress. These steps overlap and interact with one another, so GNB can take complementary actions to address multiple recommendations at once.

1. **Address gaps in the list of indicators.**

A useful initial step is to address gaps in the preliminary list of recommended indicators. The list explicitly encourages members to share suggestions for potential indicators and data sources. GNB may also wish to pursue a more direct “crowdsourcing” approach, soliciting suggestions for particular ToC components from all member organisations or from those most likely to be working in areas relevant to the gaps.
For example, RG6 (*Married girls increasingly access divorce, annulment, and child custody*) and OS4 (*Improved economic security for girls*) could benefit from suggestions for additional indicators. One component, RFC2 (*Community, traditional, and religious leaders increasingly support alternative roles for girls beyond marriage*), has no indicators at all, in part because it overlaps with another component (RFC1). GNB may wish to ask members for suggestions, including how to distinguish RFC2 from RFC1. This need for sharper conceptual clarity speaks to a broader recommendation, discussed in more detail below, about outcomes and results that need to be clarified in order for appropriate indicators to be identified.

2. Clarify ToC language to make it easier to identify appropriate indicators.

APEP recommends that GNB undertake additional efforts to define the meaning of specific language in the ToC components. For example, one of the target results for girls is as follows: *Married girls increasingly access and use services and supports of all kinds* (see RG5 in the list of indicators). This result is quite broad and could encompass any number of services or supports, making it difficult to identify a manageable number of appropriate indicators. As GNB USA members suggested in their informal memo, a useful solution would be “choosing three kinds of primary services/supports and measuring these systematically.” By more sharply defining focal points for this broadly worded result, GNB will be better able to identify a select set of relevant indicators.

This recommendation applies to several other outcomes and results as well, including:

- RFC2. Community, traditional, and religious leaders increasingly support alternative roles for girls beyond marriage (this overlaps with RFC1, as noted above);
- OS3. Health and education services establish protocols on identifying the warning signs and addressing the risks of child marriage;
- OS5. Increased commitment of programmes to prevent and mitigate risk of child marriage;
- RS1. Service providers take greater action to prevent child marriage and support the needs of married girls;
- OLP3. Governments develop supportive policy frameworks with adequate resourcing across Ministries to increase educational, economic and social opportunities for girls at risk of child marriage and married girls; and
- OLP5: Increased accountability and monitoring of national / regional / community institutions.

To be sure, this is a difficult task; it will challenge members to address the tension between achieving the breadth needed to encompass partners’ diverse efforts versus developing a well-defined framework of indicators that allows for replication, comparison, and aggregation. However, APEP believes that GNB would benefit from a process to clarify the intention of the language in the TOC. This need not entail changing the TOC graphic. The accompanying ToC brief could be revised to provide concrete supplementary details that sharpen the definitions and provide much-needed focal points. The result should be improved, more precise indicators that can then be tested by members in their real-world settings. This could illuminate the relative utility of a given indicator for varied GNB partners – NGOs, businesses, multi-lateral institutions, et al. – and at various levels of interventions – local, national, regional, and global.
If need be, APEP could facilitate a discussion among GNB members to clarify broadly worded ToC components. This process will help strengthen the basis on which members can choose the right indicators for their work.

3. Facilitate a workshop discussion among GNB members to build consensus.

Building on the previous recommendation, APEP suggests that representatives from member organisations engage in a discussion to collectively identify which indicators are most directly relevant to their work. As noted earlier, a lack of consensus partly reflects the diversity of the field, including variation in the priorities, program focus, and resources of different organisations. Consensus also needs to be built around indicators for less well-documented types of progress, such as capacity building or normative change. Based on our review of the indicator lists and reports collected by the GNB secretariat, it seems that there are valuable indicators either being used in related fields (e.g., abandonment of FGM/C) or in isolated cases (e.g., one of ICRW’s programs). But there is no generalized agreement about (or awareness of) these indicators because (to our knowledge) there has been no coordinated effort to create consensus around a measurement framework. GNB can generate this kind of consensus by convening its members and discussing which indicators make sense for its twin goals of ending child marriage and supporting married girls.

Not all indicators require consensus building, of course. Certain indicators of child marriage rates and other outcomes that reflect global development goals are already in common use: the adolescent fertility rate; girls’ educational attainment; unmet need for family planning; and rates of violence toward women, among others. These indicators can be useful for members’ evaluation purposes (if data are disaggregated or collected at a level where program impact can be discerned) or for planning and advocacy purposes.

4. Work towards aligning the ToC and indicators framework.

APEP respects the difficult consultative work required to arrive at the GNB ToC documents. These documents represent a significant milestone in the partnership’s work. Accordingly, most of the Preliminary List of Recommended Indicators follows the ToC’s structure closely. However, in undertaking this exercise, we discovered certain gaps in the ToC that GNB may wish to address. In particular, indicators for capacity building, policy advocacy progress, and social norms did not seem to map well onto the ToC – a point we discuss in more detail in the following recommendations.

When confronted with this mismatch between relevant indicators and the ToC, the impulse may be to allow the measurement framework to drift away from the ToC. Some flexibility in mapping indicators onto the ToC is justifiable and arguably necessary; but we caution against allowing GNB’s indicators to become completely detached from the conceptual anchor provided by the ToC. Indicators and a ToC are both key parts of effective evaluation; indicators complement and clarify the conceptual roadmap linking organisations’ activities to expected outcomes or results. Too many deviations between the two can lead to confusion. For example, advocates may find it more difficult to make their case to donors or governments if their indicators are not tied to a ToC. More generally, indicators ultimately need to be grounded in a coherent conceptual framework for GNB’s work; without this anchor point, it is far more difficult to reach consensus on which indicators should and should not be included. In order to maximize the complementarity of the ToC and relevant indicators, GNB may need to make modest adjustments to address gaps or tighten ambiguities in the ToC.
5. Consider how to address capacity building.

One significant gap in the ToC: capacity building, which is only partially—or obliquely—incorporated. Indicators for service-related capacity building could be included under the Services Outcomes, but the fit is not always a natural one. For example, capacity building for VAW and other health-related services that married girls need could be placed under *Increased access to health services for adolescent girls, married and unmarried (OS2)*—if we define “access” to include training, protocols, etc. Partners engaged in such capacity-building activities might feel that this work is otherwise not recognized or valued in the ToC. Others of the Services Outcomes could reasonably encompass capacity building, but they refer to mitigating the risk of child marriage, rather than addressing the needs of married girls—so here again the fit is awkward for certain indicators that seem worth including.

Members’ capacity to advocate for policy change merits its own set of potential indicators. Success in contributing to policy change outcomes is affected by many factors beyond the control of the advocate. But advocates can and should establish goals for improving their capacity to advocate, and indicators of progress towards these gains in capacity could be included in the ToC. UNICEF and UNFPA, for example, offer some potential advocacy capacity indicators in the context of FGM/C:

- *Proportion/number of intended national/regional advocacy stakeholders reached with key messages of Resolutions calling for elimination of FGM/C.*

- *Number of cooperative advocacy events among stakeholders including Member States, international, continental, regional organisations, NGOs and the diaspora networks on elimination of FGM/C.*

Both of these indicators tell us the level of effort or the number of “outputs” delivered. They don’t tell us about the impact of the related advocacy effort; but an increase in the number of stakeholders reached would reflect increased advocacy capacity for the member recording this change. Even the number of executive directors or senior officials from member organisations who attend an advocacy planning session can tell us something useful about levels of commitment to advocacy work within a partnership, an important indicator of advocacy capacity. Indicators such as these could be adapted to focus on child marriage and related policy issues, thus providing a way to capture growth in advocacy capacity.

6. Incorporate progress indicators for policy advocacy work.

Based on the ToC brief, it seems that policymaking is conceptually located at the Outcome level (i.e., part of the conditions necessary for behaviour change). The focus of the ToC is largely on the question of whether policies such as favourable property rights for married women exist, with a secondary focus on implementation. This makes it unclear where to place indicators that measure key steps in the process that leads up to passage of a law or a change in key regulations or administrative practices. The ToC could potentially include indicators of interim progress toward achieving positive changes in legislation, regulations, or other policy-related outcomes. This suggestion dovetails with the second recommendation above, which points toward the need for additional clarification of certain ToC components. As part of the process of more sharply defining focal points for ToC components, it may become clearer where to place these kinds of progress indicators. In the interim, APEP has included a link to its Advocacy Progress Planner in the User’s Guide accompanying the Preliminary List of Recommended Indicators; the Planner provides a web-accessible template for planning an advocacy
effort and guides a user to develop meaningful “benchmarks” or indicators of progress towards desired policy change and advocacy capacity objectives.

7. Integrate indicators for social norms.

The indicators we have received and collected for social norms do not align well, in our view, with the wording of the ToC components. The most natural place to include social norms indicators seems to be in the Family and Community components of the ToC. However, we find that the current formulation of these components makes this an awkward fit for most of the social norms indicators we have. In addition, the pool of indicators suggested by GNB members did not include certain social norms that seem to be highly relevant to child marriage, such as norms about dowry and bride price. We have included a few social norms indicators in the list as proxy indicators – that is, indicators of changes that are relevant to but not direct measures of progress on an outcome or result. Looking ahead, it may be useful for GNB to consider adjusting the ToC language in order to more formally integrate the measurement of social norms.

8. Develop a time horizon for some outcomes and results.

Building on the above recommendations for potential ways to revise and/or refine ToC documents, it might be useful to add a time horizon to some of the outcomes and results. The UNICEF ToC and results framework for efforts addressing child marriage, cited earlier, notes that:

"Success at the population level over the longer term is defined in terms of significantly reduced child marriage rates and the shifting of gender norms. For most countries this is likely to occur over a minimum of a 10-30 year period... In the shorter term, success in terms of attitudinal and behavioural change among a critical mass of girls, families, and communities can occur over a 5-10 years period... The main risk here is that even for subpopulations in most countries, tracking outcome measures on child marriage and norm change in less than a 5-10 year time frame is likely to convey deceptive results. For a 2-4 year project, it would be essential that success be measured in terms of the effective and efficient execution of the programming effort rather than the resulting population level change."

Providing a realistic timeframe for sweeping changes in norms, policies, attitudes, and behaviours can help GNB participants reap the benefits of the significant learning function these indicators can provide. Even our brief interaction with GNB and its members has conveyed their commitment to learning and improving as they work. Well-chosen indicators can facilitate timely learning about successes (what is working) and challenges (what is not working), which in turn informs timely adjustments to strategy – what is often called “strategic learning.”

Our recommendations are of course based on our reading of the available documents. We speak at a distance, and we have no claim to the kind of deep knowledge and grounded insights that the secretariat and members have. We present our preliminary list of recommended indicators and this set of recommendations as a first step in a valuable process. We are grateful for this opportunity to help GNB pursue its vital work.
List of Consulted Sources


Department of Foreign Affairs, Trade and Development (DFATD), Canada. n.d. Children and Youth Standard Common Indicator Menu (Draft). EDRMS No. 6607922. (Document received from Girls Not Brides, July 2014.)


The International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), Associates for Community and Population Research (ACPR), and Plan International Bangladesh. 2013. *Child Marriage in Bangladesh. Findings from a National Survey*. Plan International Bangladesh.


ICRW. n.d. Overview of TESFA Program Design and M&E Framework. (Document received from Girls Not Brides, August 2014.)
ICRW. n.d. Conceptual Framework. (Document received from Girls Not Brides, August 2014.)

ICRW. n.d. Overview of IMPACCT Research Methods. (Document received from Girls Not Brides, August 2014.)


Netherlands [more precise information unavailable]. Result Chain of the Four Result Areas of Dutch SRHR-policy. Overview of goals, targets, indicators and performance questions. (Document received from Girls Not Brides, July 2014.)


UNFPA. n.d. Monitoring and Evaluation. (Document received from Girls Not Brides, July 2014.)


UNICEF India. n.d. List of indicators. (Document received from Girls Not Brides, July 2014.)


