Based on our understanding of evolutionary theory and ontogeny, we believe that program change is something to be expected and encouraged. Ideally, throughout the program’s evolution, energy is spent on implementing the program, determining where it is in its development, and working to move it to the next phase of development. Each program follows its own unique path through this lifecycle, and any given program might move forward and backward through and between the phases as needed, although the general goal will be to progress through the phases over successive implementations. Programs mature as part of an ongoing trial-and-error learning process, and many times a program is discontinued or dramatically revised because it does not show evidence of affecting outcomes. In effect the idea of a lifecycle creates system pressure to move programs along and not allow them to become static or ineffective.

Programs in different life-cycle phases have different needs from evaluation, thereby implying parallel life-cycle phases for evaluation*. The ideas of symbiosis and co-evolution have important practical implications for program evaluation. This is similar to the bee and the flower, each needing the other in order to grow and evolve. In the ideal, program development and program evaluation are a symbiotic co-evolutionary relationship, where evaluation changes as the program matures, and both program administrators and program staff get needed information from the life-cycle appropriate evaluation data. This is a difficult ideal to achieve in practice. It requires that the evaluation systems be engineered in such a way that each stakeholder group’s incentive to participate in the evaluation is well understood. Phase appropriate evaluation encourages growth and maturation of the program. Our current model for this (not yet fully mature) is diaprammed below.

Most programs are not expected to survive through a complete maturity process. In many cases, experience and evaluation will show that a program is not sustainable for various reasons. It is possible for evaluation to reveal that a program is not achieving the desired outcomes, that there are negative consequences, or participation or funding was too low to maintain the program. Many programs will be retired then revamped in order to try another approach, thus facing another cycle of growth and starting the process over again.

<table>
<thead>
<tr>
<th>Program Phase I: Initiation.</th>
<th>Evaluation Phase I: Process and Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptualization through piloting. A Phase I program is a new program just starting up. These programs are changing a lot, and issues of initiation are likely to arise: identifying and training program staff, localizing the program to the immediate context, reacting to the unanticipated problems that arise, etc.</td>
<td>Initially, evaluation on new programs should be flexible and dynamic, emphasizing implementation and process assessment and providing rapid feedback that to refine the program model, “debug” the program procedures, identify barriers to high-quality adoption, and assesses participant response to the program. Ideally evaluation will be an integrated aspect of the program implementation, and debriefing between program staff.</td>
</tr>
<tr>
<td>Phase I A programs are in their initial implementation(s) either as a newly conceived program or as an existing program adapted from another context or from basic research.</td>
<td>Phase IA evaluations are typically post-only participant feedback and satisfaction or staff monitoring and observation.</td>
</tr>
<tr>
<td>Phase I B programs are new programs that have been revised and are being reimplemented. These programs have procedures that tend to change from one implementation to the next.</td>
<td>Phase IB evaluations are also typically post-only assessments, but focus on substantive outcomes, and are used for program development and assessment of internal consistency (reliability) of outcome measures. Informal observations are also used.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Program Phase II: Development.</th>
<th>Evaluation Phase II: Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase II programs are being implemented consistently and show empirical evidence of participant changes in outcomes. There been development of standardized program procedures or protocols.</td>
<td>This phase emphasizes the assessment of changes in outcomes (e.g., knowledge, skills, attitude, behavior, performance) that occur in association with the program. The major difference between the two sub-phases is where the change is being measured — within groups or within individuals. This phase should include the design of</td>
</tr>
<tr>
<td>Phase IIA programs are implemented accurately with respect to</td>
<td></td>
</tr>
</tbody>
</table>
their model and presented consistently over multiple implementations. The program model incorporates needed adaptations to local context and setting without requiring fundamental change to the model. Participant experience from one implementation to another is relatively stable (e.g., formal lessons or curricula).

**Phase IIIB programs** have formalized written procedures (e.g., preliminary protocol or implementation plan) outlining how to implement them.

**Observational procedures and measures of key outcomes**, assessing the consistency and construct validity of measures, looking at pre-post differences and examining the relationships among different observations. This is the realm of correlational studies.

**Phase IIIB evaluations** typically consist of a matched pretest and posttest of outcomes. Verify reliability and validity of change. Because participant identification is necessary to match pre and post outcomes and results are increasingly used for public accountability, the level of protection of participants increases and human subjects review and protection (informed consent, anonymity or confidentiality) is increasingly formalized.

**Program Phase III: Maturity.**
Mature programs are evidence-based programs. They meet the requirements of Phase II, are no longer changing significantly, and show evidence of being associated with changes in outcomes. Program planners and providers know what can be expected in implementing the program; there are relatively few surprises. Relatively few programs reach this level of maturity in terms of consistent outcomes or effectiveness.

**Phase III A programs** are associated with consistent change in relevant outcomes. They are aware of other similar programs; The program is routinized, frequently with a written protocol or process guide. The program is no longer dependent upon particular individuals for implementation.

**Phase III B programs** have been shown to be effective when compared with an alternative program or no program.

**Evaluation Phase III: Comparison and Control**
Evaluation focuses on the program’s causation and effectiveness. Evaluation uses comparison groups or variables and statistical controls for adjusting for uncontrolled factors. This is the realm of experimental and quasi-experimental designs and of more structured and comparative qualitative approaches. Performance of participants is compared with some standard expectation of performance or with outcomes of people who participate in alternative programs or no programs at all.

**Phase III A evaluations** show consistent change in outcomes. Use of design and statistical controls and comparisons (control groups, control variables or statistical controls).

**Phase III B evaluations** use experimental or quasi-experimental designs (randomized experiment; regression-discontinuity) for assessing the effectiveness of the program. Performance of participants is compared with some standard expectation of performance or with outcomes of people who participate in alternative programs or no programs at all.

**Program Phase IV: Implementation and Dissemination.**
The program is adapted for wider implementation (different sites, staff and participant) while still adhering to the essentials of the program model. Logistical issues regarding support of the program over a broader range of circumstances are addressed. The primary focus of this phase is on extending the program to other settings or populations of interest, pushing the ecological boundaries of the program as originally conceived into new niches or applications.

**Phase IV A programs** are implemented in multiple sites for the purpose of assuring that effectiveness holds in different contexts.

**Phase IV B programs** are in wide distribution under the assumption that evaluation evidence shows that they can be implemented widely with fidelity and with comparable effects. The primary focus of this phase is on extending the program to other settings or populations of interest, pushing the ecological boundaries of the program as originally conceived into new niches or applications.

**Evaluation Phase IV: Generalizability and Synthesis**
Evaluation is concerned with generalizability or external validity. It examines the consistency of outcomes across different settings, populations or program variations. This is the realm of secondary and meta-analysis or of program review approaches that seek general inferences about the transferability of the program.

**Phase IVA evaluations** are multi-site integrated assessments yielding large data sets over multiple waves of program implementation.

**Phase IVB evaluations** present a formal assessment across multiple program implementations that enable general assertions about a program in a wide variety of contexts (e.g., meta-analysis).