A Mature Approach to Assessment

Prof. Stephen J Zilora, Rochester Institute of Technology (CAST)

Stephen Zilora is an Associate Professor and Chair of the Information Sciences and Technologies Department at Rochester Institute of Technology. For nearly 15 years he has been researching and teaching software application technologies and development processes, areas in which he also has over 20 years of industry experience.

Prof. Jim Leone
Abstract

Educators can tend to focus on teaching rather than student learning. As such, assessing student outcomes is perceived as additional work not directly related to their craft and is an occasional exercise required to meet the demands of program accreditors. This attitude parallels that of software developers who see the need to deliver a software project on time and on budget as something that constrains their creative work. The Capability Maturity Model has been adopted by many software organizations as a framework to help change attitudes and improve the software development process. In this paper we show that the same principles can be applied to the assessment process to achieve the same goals.

1. Introduction

1.1 Attitudes

Educators regularly assess their students’ work with carefully constructed tools. There are many tools available such as tests, projects, and homework assignments. These tools are used based on the task at hand and new tools are developed as necessary. Such assessments assist the educator in determining the student’s degree of subject mastery. Formative assessments point out gaps in a student’s understanding and provide opportunity for improvement. Final, summative assessments involve the review of not just the final course average, but also the progress the student has made, the consistency of effort, and additional factors. Typically, the educator expends significant consideration to develop a complete picture of the student’s successful achievement of the course’s goals and objectives.

However, educators sometimes view the assessment of their own work as a fruitless burden imposed by an accrediting body. It is viewed as a purely summative judgment resulting from a process over which the educator has little control.

This attitude is due in part to most educators’ focus on teaching vs. learning. This is not surprising since educators are typically evaluated based on their “teaching” skills. Such evaluations may include a determination of their organizational skills, their clarity of delivery, and their fair and timely grading of student work. Their evaluations are a measure of their teaching, not student learning. However, while educators do and are encouraged (via evaluations) to practice “good” teaching, assessment is limited to measuring learning. Obviously, the two are not unrelated. And yet, teaching is just one of many ways learning can take place. Saulnier and others argue for a shift from a Teaching Paradigm to a Learning Paradigm, “The Learning Paradigm shifts what the institution takes responsibility for: from quality instruction (lecturing, talking) to student learning.”

This disconnect—educators practice and are evaluated based on teaching while assessment measures learning—significantly contributes to the perception of many educators that assessment is additional work for the sole purpose of satisfying some higher authority.

1.2 Process
Regardless of prevailing attitudes, most institutions have an assessment process in place for the purpose of accreditation. For many organizations, the use of the word “process” can be overly generous. For these institutions the assessment exercise is an ad hoc collection of data in an effort to satisfy a program reviewer. The attitudes in place undermine the value of the process as each assessment period becomes a standalone exercise.

The situation is not unlike the software development process in many companies. When a new project comes along, there is a flurry of activity to complete the work with little attention paid to past practices or past projects. As a result, developers end up working in “hero mode” putting in countless hours as looming deadlines draw near. In 1989 while at the Software Engineering Institute of Carnegie Mellon University, Watts Humphrey introduced the notion of the Capability Maturity Model to describe the state that existed in software development organizations. It includes the concepts of “levels” of maturity and “key process areas.” Many organizations have adopted this model now referred to as CMMI (Capability Maturity Model Integration) and used it to guide their maturation process. It is a framework that allows an appropriate, mature approach to software development to become part of the organization’s culture. Marchewka introduced the concept of “Assurance of Learning” coupled with CMM, a focus similar to that proposed by the authors. In that work, the emphasis was on curriculum development and program improvement with possible implications to accreditation as related to business school accreditation. The authors believe that this same framework can be applied to the assessment process to help institutes realize benefits from the assessment process, rather than view it as a necessary evil.

The CMMI was developed to form an overarching model that could unify the many existing maturity models. The adoption of the CMMI has been very successful with more than 1,400 appraisals (evaluation of an organization’s maturity level) done in 2012 and 7,200 since 2006. While the CMMI has exhibited success in unifying several maturity models, many independent maturity models, still based on the original CMM principles, have emerged. There are examples for curriculum design, e-learning, innovation, decision-making, and personnel assessment. The authors are proposing a CMM-based model for the curriculum assessment process.

2. Assessment

The word “assessment” takes on diverse meanings depending on the audience. As an educator, one can focus on the way knowledge and skill-sets are documented and achieved. In the STEM disciplines, the ABET definition is as good as any to initiate the link between CMMI and assessment:

Assessment is one or more processes that identify, collect, and prepare data to evaluate the attainment of student outcomes and program educational objectives. Effective assessment uses relevant direct, indirect, quantitative, and qualitative measures as appropriate to the outcome or objective being measured.

To the uninitiated, questions asked include, but are not limited to, how it is done (methodology), what it accomplishes (outcomes) and why it is necessary. To the educator who has been involved in curriculum efforts and/or engaged in discipline-specific or regional accreditation activities, the
answers are, for the most part, obvious. That fact not withstanding, however, the assessment process is fraught with problems and errors.

Some examples include incorrect or inappropriate development of assessment goals, objectives, or outcomes (connotations differ among these terms). Other common problem areas are weak or ineffective processes and problematic data collection (too much data, insufficient data, or simply inappropriate data).

Assessment has to do with the process of education. While professors are very good at “professing,” assessment gives us a measure of how effective that professing is. It is not about the actual material being delivered, but rather how that material is delivered and received. This notion can be met with embrace or disdain. The instructor who views teaching as a craft to be perfected will embrace assessment as a vehicle for improvement. However, an instructor who believes in the natural ability to teach may view assessment as an unnecessary burden.

This dichotomy is not unlike the attitudes that exist among software developers. Some see their profession as a craft to be perfected; others see it as a natural talent and a form of self-expression. Many software development organizations have successfully adopted the Capability Maturity Model (CMM) to instill an attitude and culture of craft perfection. The authors believe that the CMM can also be used to instill an attitude and culture of craft perfection for teaching and its assessment.

3. CMMI

3.1 Brief History

The development of the Capability Maturity Model (CMM) began in 1984 with the formation of the Software Engineering Institute (SEI) at Carnegie Mellon University, established by the Department of Defense to improve the quality of software-based systems. This work led to the publication of Watts Humphrey’s 1988 landmark paper describing the 5-level model. Work continued over the next several years and in 1995, SEI published a detailed description of the model and its implementation method in “The Capability Maturity Model: Guidelines for Improving the Software Process.”

Subsequent work resulted in a more comprehensive version of the framework, the Capability Maturity Model Integration. This version focused more on results than process, required that there be fundamental organizational support such as formal assignment of responsibilities and provision of training, and introduced the notion of a staged vs. continuous model.

While these differences are significant in terms of assessing an organization’s maturity or capability for software development, for the purposes of this paper we will focus on the idea of levels and progress from one level to the next.

3.2 Elements
The Capability Maturity Model defines 5 levels of maturity. Within each of levels 2 through 4 are several Key Process Areas (KPAs) as illustrated in Figure 1.

All organizations are, by definition, at least at Level 1. At this level, software development is done largely in a vacuum. The project team has its requirements and its tools and goes about its business developing the product. They may do a very good job or they may find themselves over-budget, late, or failing to meet all the specifications. Being at level 1 does not mean that an organization is incapable of producing good software—it means that compared to a more mature organization, a given effort may have more defects, require more rework, or have more schedule slippage.

Level 2 is about documenting the process. This level can appear to require fruitless activities—there is opportunity for improving the development effort of the current project, but there is no provision (in at this level) to incorporate any improvements into future work. However, it is not possible to evaluate a process if there is no documentation of it.

Level 3 is about planning. Shouldn’t the planning of level 3 come before the documentation of level 2? Ideally, yes. But from a maturity point of view, planning an activity comes at a later stage than performing an activity. Also, what is documented in level 2 can inform what is planned in level 3.

Level 4 requires active management of the plan thus ensuring that either the plan is followed or that a good reason for consciously deviating from the plan exists. Having a plan is good—in level 4 the plan is put into action.
Level 5 is the post mortem, or introspection level. This is a look at what was planned (level 3), how things went (level 4) and what the outcomes were. Most importantly, ways to improve the plans and management are sought. This is possible because of the existence of good documentation (level 2) of the plans and management. The importance of this level’s activities cannot be overstated. They make continuous improvement possible.

4. Assessment Maturity

In applying the Capability Maturity Model to the assessment process, we choose to retain the same 5 levels of maturity. While most of the Key Process Areas (KPAs) of CMM are unique to software development, some of them (e.g., training programs, organization process definition, and program change management) can be applied to a variety of domains. Further, the intent of the KPAs can be applied to non-software domains. It is this universality that gives rise to the maturity levels and KPAs for assessment that we illustrate in Figure 2 and describe below.

4.1 Maturity Levels

Just as with software development, all organizations are, by definition, at least at Level 1. At this level, assessment is done in an ad hoc fashion, if at all. For example, faculty teaching advanced courses may find their students unprepared and mention this to their colleagues who teach fundamental courses. Or, there may be an occasional effort to quantify student success for the purpose of a pedagogical study, but not necessarily to provide input into a process for continuous improvement. Being at level 1 does not mean that good teaching is not occurring. It means that there is no formal determination whether good teaching is occurring.

Level 2 is about conducting and documenting the process of formal assessment. Data is collected and analyzed, but without context and consistency it provides only a snapshot of the learning
success. At this level there is no pre-specified instrument for obtaining data nor is there any longitudinal or covariate analysis of the collected data.

Level 3, just as with software development, is about planning. At this level, measurements are mapped to learning outcomes and are collected on a planned schedule. The assessment process is planned and purposeful. Personnel will be assigned and given time and resources to carry out the assessment process.

At Level 4 a longer view of the process is taken. It is no longer about an individual round of assessments, but rather how any given assessment instance fits with other past and future assessments. At this level, the assessment plan is carefully managed to ensure the plan is followed or that a conscious deviation from the plan has a good reason.

Level 5 closes the loop. It is at this level that a review of both the results and the plan can result in necessary changes. There may be changes in the assessment methodology or in the teaching practices and curriculum. Most importantly, at this level what is sought are ways to improve all aspects of the assessment process and the outcomes.

4.2 KPAs

4.2.1 “Repeatable” KPAs

For an organization to be operating at Level 2, its process needs to be repeatable. In practice, this means that the process is fully documented so that it can be reproduced. The two key process areas for this level are Data Collection and Data Analysis. Documentation of data collection needs to include what data was collected, when it was collected, and how it was collected. This includes documentation of the courses and the measurement instruments that were used. It also includes documentation of the frequency (during the assessment period) and the manner in which the data was collected.

Documentation of data analysis must include any filtering of data that might have been done, a description of any statistical methods that were employed, and a statement of any thresholds that were used to judge the results.

4.2.2 “Defined” KPAs

For an organization to be operating at Level 3, its process needs to be planned. At this level, management commitment is essential due to the need for greater resources. The Key Process Areas for the Defined Level are Outcome Mappings, Scheduled Data Collection, and Assessment Training. Each of these KPAs requires a longer-term view of the assessment process. The Assessment Training KPA is one example. This KPA includes training on how to develop appropriate measurements for assessment purposes, training in data analysis, and training in the overall process. It is quite likely that only some faculty would engage in in-depth training on data analysis or measurement development. But it is essential that everyone understand what needs to be done and how to do it.
Part of the planning at this level is the schedule of data collection. Since not all outcomes will be assessed in every assessment period, a schedule needs to be in place to indicate when each outcome will be assessed. Additionally, a schedule of the assessment process phases (collection, analysis, review, action) needs to be defined.

Outcome mapping refers to the need to establish, a priori, what the possible outcomes will be as a result of the data analysis, i.e., what difference between the results and the threshold will dictate some action and what might that action be?

4.2.3 “Managed” KPAs

In the CMM appraisal, Level 4 (and Level 5) can be quite difficult to achieve. This level requires not only management support, but also an institutionalization of the model. It is at this level that the CMM becomes part of the culture of the organization. As such, the KPAs look at longer-term issues—issues that span assessment cycles. The Regular Data Collection and Regular Review process areas require that the assessment process be carried out on a planned (Level 3) schedule. The difference between Level 4 and Level 3 is the need for evidence, not just good intentions that the process is indeed being done on a regular basis.

The Historical Repository and Longitudinal Studies KPAs are crucial to the success of Level 5 activities. Not only must there be a database of some sort for tracking results over time, but there needs to be consistency in the measurements used and the process employed for well-designed longitudinal studies that are particularly important. At this level, an organization can/must objectively answer the question, “Are we getting better?” Tom DeMarco’s aphorism that you “cannot control what you cannot measure,” underpins Level 4. Data collected over time permits management of the process.

4.2.4 “Optimizing” KPAs

True maturity comes when an organization recognizes its weaknesses and works to eliminate them. Level 5 deals with process improvement. As such, the Key Process Areas deal with reviewing how well the process has performed and determining what can be done better. Curriculum Review, as the name implies, requires an examination of the curriculum. Can improvements be made to improve the assessment results? Methodology Review looks at the process of collecting the assessment data. Are the correct measurements being taken, are the thresholds appropriately set?

The area of Defect Prevention is borrowed from the software development field, but applies here as well. Perhaps an instructor was unaware and therefore did not use a particular test question that was needed for assessment. How can that be prevented in the future? Ideally, risk management would also be practiced at this level. That is, instead of waiting for a defect to occur and then changing the process appropriately, potential defects are identified and strategies to address the potential defects are put into place in a proactive manner.

Finally, Change Management must be considered. Any changes that are recommended as a result of the reviews must be implemented in a thoughtful manner. This includes a cost/benefit
analysis, approval via an established change control process, and a carefully planned process for change implementation.

5. Conclusion

The authors believe the proposed model for assessment maturity is a sound application of the CMM principles. Such a model provides two key benefits. The first is to serve as a framework for institutions to improve their assessment activities. Rather than expecting an institution to advance from ad-hoc assessment activities to a well-established, value-returning assessment process with no plan or guidance, the model lays out a path with defined stages for growth. The second benefit is intrinsic to the model. The layout of the CMM levels, together with the defined, required support of management, helps build a culture where assessment is seen as a valuable tool to improve learning.

The authors have not addressed the fact that in the CMM, KPAs are made up of goals. Further work will include a refinement of the KPAs as well as the establishment of goals for each KPA. Once goals have been established, the next step will be a test implementation. The test will consist not only of the application of these KPAs, but also the careful application of adoption strategies and considerations as outlined by Everett Rogers\textsuperscript{12}.

