

Program Evaluation: Measures of Academic Progress (MAP)

Lesley Johnson

Walden University

Dr. Desiree DePriest

EIDT 6130: Program Evaluation

February 23, 2014

Table of Contents

Program Analysis: Measures of Academic Progress (MAP)..... 3

Evaluation Model Table..... 6

Evaluative Criteria 9

Data Collection Design and Sampling Strategy 12

Reporting Strategy Table 16

References 18

Program Analysis: Measures of Academic Progress (MAP)

The Measures of Academic Progress (MAP) test is a computerized adaptive assessment provided by the Northwest Evaluation Association (NWEA) that helps teachers understand their students' academic abilities (MAP, 2014). The NWEA states, "MAP assessments provide detailed, actionable data about where each child is on their unique learning path" (MAP, 2014). It's an adaptive assessment because it adapts the level of difficulty according to the student's correct or incorrect answers and finds the student's "learning level" at that time (MAP, 2014). There are four MAP assessments currently available: reading, language, math, and science. Colegio Americano de Torreon, or CAT as it's better known, uses the reading, language, and math assessments for their grades 3-9 students at their international school.

CAT began using the MAP assessment six years ago as a means to better identify areas of weakness in students' academic development. The MAP assessment is administered in the fall and spring of each year. After the fall assessment, teachers analyze the results of their class (or classes) and then create an action plan to focus on the areas of weakness in the hopes that the scores will improve in the spring. Teachers are also given the norm scores of students in the United States so they are able to compare student scores with that and differentiate instruction accordingly. However, since CAT is an international school where approximately 97% of our students speak English as their second or third language, we also compare our scores with the other ASOMEX (Association of American Schools in Mexico) schools because, as Beatrice Morales (personal communication, January 21, 2014), CAT's curriculum coordinator explained, until last year NWEA didn't have the norm scores for international schools with ELL (English Language Learner) students. Yet even with the international school norms, CAT continues to compare our students' scores with the other ASOMEX schools, creating an atmosphere of a high-

stakes test because the administrators like to use the MAP scores to compare our students with the other ASOMEX students to ensure our school is providing a competitive education.

Furthermore, each year CAT teachers are gathered in the gymnasium and the general MAP scores are reported so that the entire staff can see how our students have done. Math scores have always been above the norm and reading has been CAT's perpetual weak point with scores coming in below the norm (Morales, personal communication, January 21, 2014). Because scores are so public, teachers often feel the pressure to make their students perform; however, the MAP assessment is supposed to be another tool teachers can use to monitor student progress and for data triangulation (Kessler, 2014).

To add to the high-stakes feel, the MAP assessments are administered in a short period of time with students taking the three assessments within a two-day period while still attending their regular classes. For the 7-9 students, the tests often feel like their semester exams because a specific proctor is assigned to the room and students aren't allowed to leave until an hour and half has passed. Those procedures mirror CAT's exam procedures which can make the MAP assessment feel like it is a high-stakes assessment worth 30% of their grade, just like their semester exams. And lately administration has been pushing for teachers to add the MAP assessment as a graded assessment so that students will take the test "seriously" (Morales, personal communication, January 21, 2014).

Clearly, the biggest stakeholders in CAT's MAP assessment program are the teachers, students in grades 3-9, and administrators because the MAP scores directly affect them. Parents do receive their child's scores, but CAT is just beginning educate parents about what those scores mean (Morales, personal communication, January 21, 2014); therefore, they are not viewed as

major stakeholders at this point. However, it is the students who see their scores and are given a goal score for their next assessment (and it's those numbers that may comprise a grade for them in the future), teachers are expected to differentiate instruction to help students improve their skills and consequently perform better on the spring MAP assessment, and administrators who use the MAP scores to monitor student and teacher performance, so they are the major stakeholders.

As it is, it is unclear if the MAP assessments are appropriate for CAT. It's also unclear if the assessments are being implemented and the results interpreted in the most prudent manner. With such a high number of ELL students who feel as though all tests and assessments are high-stakes, the MAP scores could be flawed due to ELL comprehension issues and stressed out students. The stagnant scores also suggest an inability to use data in a meaningful way by both the teachers and administration. With all of these issues, one must question if CAT should be investing in the MAP program or if we should seek another assessment to help us triangulate student achievement data.

As a 9th grade Language Arts teacher at CAT, I am a part of this program and have participated in it for the past three years. I have felt the pressure to improve student performance and I have felt overwhelmed with the data the MAP assessment provides (it is very extensive). I will have to ensure that I am trying my best to put my own biases aside as I research if the MAP assessment is just as relevant for an ELL student as it is for a native English speaking student.

[back to Table of Contents]

Evaluation Model Table

Evaluation Model	Advantages	Disadvantages
<p>EXPERTISE AND CONSUMER-ORIENTED APPROACHES</p>	<ul style="list-style-type: none"> • I am an SME for Reading and Language, two of the areas of the MAP assessment. • I have worked with the MAP assessment for three years, so I am familiar with the information teachers and students receive. • I am a "connoisseur" of the MAP assessment (Fitzpatrick, Sanders, & Worthen, 2011, p. 128). • I have common sense (Fitzpatrick, Sanders, & Worthen, 2011, p. 129). • I can find other educators and administrators to interview to help me with the evaluation (Fitzpatrick, Sanders, & Worthen, 2011, p. 132). • I have a limited audience that I plan to share the results with (Fitzpatrick, Sanders, & Worthen, 2011, p. 143). 	<ul style="list-style-type: none"> • "Generally speaking, expertise-oriented evaluators should not be closely affiliated with the institution or product they are judging" (Fitzpatrick, Sanders, & Worthen, 2011, p. 135). Unfortunately, I am both... • It might be difficult to set aside my personal bias if I am the SME. • I'm not 100% up-to-date on the standards at my school because we might be shifting from the curriculum we've been using since my arrival in 2010 to the Common Core (Fitzpatrick, Sanders, & Worthen, 2011, p. 132). • It might be difficult to find a time to meet with other educators and administrators since many people have busy schedules and are working overtime this year.
<p>PROGRAM-ORIENTED EVALUATION APPROACHES</p>	<ul style="list-style-type: none"> • I bet I could find the objectives tied to the MAP assessment usage at Colegio Americano (Fitzpatrick, Sanders, & Worthen, 2011, p. 154). • The NWEA (Northwest Evaluation Association) has objectives listed for the use of the MAP assessment data. • It would be good to distinguish the difference between "implementation failure and theory failure" (Fitzpatrick, Sanders, & Worthen, 2011, p. 161). After all, either element could be affecting the MAP assessment at Colegio Americano. • Stakeholder input is an important part of the evaluation and I have access to many 	<ul style="list-style-type: none"> • It's unclear if the "purposes of [the MAP assessment] are specified" (Fitzpatrick, Sanders, & Worthen, 2011, p. 154). I doubt they are written down, but I'm sure if I talked to the administrators that they would have their own ideas and opinions. • It would be a gargantuan task to speak with all stakeholders, so I will have to carefully choose the stakeholders and ensure that I have a variety (i.e., middle school teachers and students, elementary school teachers and students, and administrators). • "Theory-based approaches are also criticized for oversimplifying the complexity of program delivery and context" (Fitzpatrick,

	<p>stakeholders (Fitzpatrick, Sanders, & Worthen, 2011, p. 162).</p> <ul style="list-style-type: none"> • I am an SME and "the process of developing program theory should, therefore, rely on a combination of input from stakeholders, theories and research from relevant social science studies, and the evaluators' knowledge and expertise" (Fitzpatrick, Sanders, & Worthen, 2011, p. 162). 	<p>Sanders, & Worthen, 2011, p. 167).</p>
<p>DECISION-ORIENTED EVALUATION APPROACHES</p>	<ul style="list-style-type: none"> • This is "among the oldest approaches to evaluation" (Fitzpatrick, Sanders, & Worthen, 2011, p. 184), so it has been proven to be tried and true. • "These approaches are meant to serve decision makers" (Fitzpatrick, Sanders, & Worthen, 2011, p. 172). I hope to be able to share this information with administration, or the "real" decision makers at Colegio Americano. • You don't have to "wait until an activity or program has run its course before evaluating it" (Fitzpatrick, Sanders, & Worthen, 2011, p. 178), and technically, MAP is ongoing at Colegio Americano. • The focus of this evaluation is about gathering information (Fitzpatrick, Sanders, & Worthen, 2011, p. 179). • "The evaluator is developing a personal relationship with the primary users to meet their needs and sustain their interest in the evaluation" (Fitzpatrick, Sanders, & Worthen, 2011, p. 181). I have personal relationships with many of the primary users. • "The CIPP model has been used extensively in school districts..." (Fitzpatrick, Sanders, & 	<ul style="list-style-type: none"> • It's unclear if MAP information has been consistently gathered throughout the six years the program has been in operation; however, I'm sure I can get something like two or three years worth of longitudinal data. • As an international school, Colegio Americano has a regular turnover of international teachers who stay for a couple of years and then leave, so there may be stakeholders who do not understand all aspects of the MAP assessments.

	<p>Worthen, 2011, p. 184), so it seems this form of evaluation has been proven to be effective in school settings.</p>	
<p>PARTICIPANT-ORIENTED EVALUATION APPROACHES</p>	<ul style="list-style-type: none"> • Involving "participating in the program, managers, staff, and other key stakeholders" will be important in the evaluation (Fitzpatrick, Sanders, & Worthen, 2011, p. 191). • Allows for more changes along the way as one learns from the participants, so it is less "preordinate with evaluation plans" (Fitzpatrick, Sanders, & Worthen, 2011, p. 193). • Observation is key for this kind of evaluation. • Stakeholders become more informed and the evaluator becomes more like "a negotiator to help stakeholders reach consensus on their diverse views and decide on next steps or priorities" (Fitzpatrick, Sanders, & Worthen, 2011, p. 198). • I already know about the organization (Colegio Americano) and its context (Fitzpatrick, Sanders, & Worthen, 2011, p. 205). • P-PE uses a smaller number of stakeholders (Fitzpatrick, Sanders, & Worthen, 2011, p. 210). 	<ul style="list-style-type: none"> • It's important to see the program in action, but that will be difficult (Fitzpatrick, Sanders, & Worthen, 2011, p. 191). I can observe teachers implementing their action plans, but I won't be able to see if those plans are effective because the spring assessment won't happen until April.
<p>Explain your choice of model for your program evaluation: I believe I will use the decision-oriented evaluation approach because I am using an ongoing program. Also, I believe I will have to use stakeholder interviews along with the data to complete my evaluation, something that is feasible for this evaluation. Observing the program in action would be far too large of a task, but I think the CIPP Evaluation Model will be a good fit for this MAP assessment evaluation and it will allow me to "give focus to the evaluation [because] experienced evaluators know how tempting it is simply to cast a wide net, collecting an enormous amount of information, only later to discard much of it because it is not directly relevant to the key issues or questions the evaluation must address" (Fitzpatrick, Sanders, & Worthen, 2011, p. 178). I think the model will help me focus on the important parts of the MAP assessment and its implementation at Colegio Americano de Torreon.</p>		

[\[back to Table of Contents\]](#)

Evaluative Criteria

As I continue to explore the MAP assessment program at Colegio Americano de Torreon (CAT), I will attempt to answer these five questions:

1. Is the MAP assessment appropriate for ELL students?
2. Should CAT include descriptive information about the students (i.e. how many years they have been at CAT, how long they have been studying English) with the final MAP assessment results?
3. Should some of our beginning students who have had limited English exposure be exempt from taking the MAP until they can show some English proficiency?
4. Should there be accommodations for student with lower English proficiency levels?
5. Should there be more direct instruction for the MAP results to be more accurate?

Clearly my main focus will be on the appropriateness of the MAP assessment for ELL students, a population that comprises 97% of CAT's students according to Beatrice Morales, CAT's curriculum coordinator (personal communication, January 21, 2014). From that main question I have derived other evaluation questions that will help me focus on specific elements of the MAP assessment, namely the characteristics of the learners. In my four years at CAT, I have never seen the number of years students have had English courses tied to the results. This is an important variable because CAT accepts new students who have limited English abilities (as long as they pass the entrance exam) and CAT provides a year-long Intensive English course for students who would like to start in 7th grade at CAT but have received the majority of their education in Spanish. Speaking from personal experience as a 9th grade English teacher, it is often the former Intensive English students who score lowest on the MAP assessment in 9th

grade; however, that would directly correlate with the number of years they have been studying English, so that is a statistic that should be tied to the MAP assessment if CAT is to utilize the assessment data to its fullest extent.

Similarly, this same issue leads one to question if students with a lower proficiency of English should wait to take the MAP assessment (especially if their teachers can tell through classroom assessments that they are behind their peers) so that those scores do not affect the group score norms. Also, if the MAP assessment is to measure academic progress, perhaps accommodations should be made for students who have a limited proficiency of English. Something like providing a dictionary wouldn't give students the answer, but it would help them understand new words that could have otherwise impeded the students' understanding of a question.

Finally, I will address the issue of using direct instruction to improve MAP scores. It's unclear if this practice would be ethical since the assessment is meant to be a measurement of the student's progress, but it is a proposal that administrators have made to improve our MAP scores. However, this kind of direct instruction may be a kind of accommodation for ELL students, so perhaps it would be acceptable.

These are all questions that revolve around the first question, "Is the MAP assessment appropriate for ELL students?" However, as I delve farther into the evaluation process, I will do my best to "remain flexible, allowing modifications and additions to the evaluation plan when they seem justified" (Fitzpatrick, Sanders, & Worthen, 2011, p. 315). Nevertheless, this evaluation will focus on the on the learners and not the implementation of the MAP assessment. While CAT may have some issues with implementation, the larger issue involves students' language abilities; therefore, the focus of this evaluation will be about the appropriateness of the

MAP assessment and if there are any ways to make the data more meaningful for students, teachers, and administrators.

The stakeholders who will need to be involved in determining the evaluation questions should be Beatrice Morales, CAT's curriculum coordinator and the administrator who most closely handles the MAP data and a sampling of the teachers who receive and use the MAP data. I believe only one administrator needs to be involved because she is the one who interprets the data for the other administrators since she has had the formal training. I will only use a sampling of the teachers who receive and use the MAP data because the MAP assessment is given in grades 3-9, so that means there are approximately 35 teachers and that may be too many stakeholders for this evaluation (especially since it is not school-sponsored and will ask them to invest their free time in a voluntary evaluation). Two teachers per grade level should give me an accurate idea of their issues and concerns and I will be able to exclude the newer teachers who are still getting acquainted with the program. Ms. Morales and the sampling of teachers will give me a clearer perspective of the scope and the issues that may be age- or grade-specific. I will also get the teachers to help me conduct some surveys that could be more telling about our newer students' progress. Their involvement will be key in reaching more stakeholders and as we know, “the single most important source of evaluation questions is the program’s stakeholders: its clients, sponsor, participants, and affected audiences” (Fitzpatrick, Sanders, & Worthen, 2011, p. 316). Ms. Morales and a sampling of the teachers and students will help me cover those bases and create a successful evaluation design.

[back to Table of Contents]



Data Collection Design and Sampling Strategy

The MAP assessment is a tool used at Colegio Americano de Torreon, or CAT as it is better known, to determine students' academic progress and to develop interventions to improve student performance. However, the MAP assessment is namely used and targeted for schools in the United States. According to Ms. Beatrice Morales, the curriculum coordinator at CAT, 97% of our students are ELL students and that could have an adverse affect on the MAP scores (personal communication, January 21, 2014). Nevertheless, many of the American Schools in Mexico use the MAP assessment as a measuring stick for their students' progress. In fact, Ms. Morales explained it has become customary to use the MAP scores as a means for comparison with schools in the U.S. and other American Schools in Mexico (personal communication, January 21, 2014).

Yet the question still remains if the MAP assessment is appropriate for the ELL learners at CAT or if another assessment should be used or if there should be some modifications made to the data collection process. After all, like most schools, there are new students every year, but those students may have had a limited amount of exposure to English. While there is little assistance for new elementary students, CAT does offer an Intensive English program for students entering middle school. Despite this assistance, former Intensive English students often continue to struggle and are often academically behind in comparison to their classmates. This begs the question if those students should be displaying their English proficiency before taking the MAP assessment or if there should be modifications made for them.

To answer those questions, two evaluation designs best suit our purposes: a time-series evaluation design and a case study. CAT has been using the MAP assessment for five or six years now, so we will be able to use that existing data in the time-series evaluation design to look

for "trends or changes over time" (Fitzpatrick, Sanders, & Worthen, 2011, p. 393). Fitzpatrick, Sanders, and Worthen (2011) even mentioned that "the results of a time-series design can be very useful at the beginning stage of a case study if the evaluator explores with stakeholders their interpretations of the ups and downs exhibited in the results" (p. 393). Gathering the important stakeholders (namely teachers and key administrators) at this point would be a nice way to hear their concerns and opinions of the MAP assessment, so that those could be taken into further consideration, perhaps even helping us focus our evaluation in a way that is more meaningful for all stakeholders. After all, teachers and administrators look at the MAP scores each year, but we rarely take a longitudinal look. This could inform us of trends or continuing issues our students have with the MAP assessment which would improve the intervention practices that teachers already plan and complete in a different way that would be more meaningful for the grade level as a whole.

After completing the time-series evaluation, a case study will give us an in-depth understanding of the MAP assessment and help us ascertain if the MAP assessment is appropriate for our high number of ELL students. Like Fitzpatrick, Sanders, and Worthen (2011) wrote, "Case studies... encourage a deeper exploration of the issues, recognizing that there are many different perspectives on each" (p. 390). Perhaps the case studies can even enlighten us to the common problems ELL students have and what other schools have done to address such issues. Fitzpatrick, Sanders, and Worthen (2011) wrote, "When absolute standards do not exist, the evaluator can begin with a review of literature seeking research or evaluation results on similar programs, as described earlier, and use those results as a stimulus for stakeholder discussion" (p. 335).

Consequently, because there is a large amount of data and varying experiences with the MAP assessment (mind you, CAT is an international school so the teacher turnover rate is high. Teachers usually stay for two or three years only.), three focus groups of stakeholders would be best. There could be a focus group for each level that uses the MAP assessment—lower elementary, upper elementary, and middle school teachers and administrators. That way there will be a small, relatively homogeneous group of people from whom I can gain in-depth information about their thoughts and experiences with the MAP assessment (Fitzpatrick, Sanders, & Worthen, 2011, p. 438). But this smaller group of people will help gauge the appropriateness of the MAP assessment for our high ELL population because I can find some more information specifically related to the evaluation questions. Additionally, a survey should be administered to the students who take the MAP test that explores their level of familiarity with English—that is, how long they have been studying the language and how comfortable they feel with the language as well as the student's age and grade level. That information can then be juxtaposed with the longitudinal study in the focus groups so that the group can determine if students with limited English exposure should be exempt from the MAP assessment until they can demonstrate English proficiency or if there should be accommodations made for students with lower English proficiency levels. Finally, with the student information and the MAP results from years past, the focus groups can explore if direct instruction would be fitting.

Nevertheless, there are existing biases that occur because there are teachers who believe their students don't take the test seriously enough since there isn't a grade tied to it so they may resist the idea that the data is meaningful; however, I can ask them to triangulate the MAP scores with their classroom assessment data to help them further understand any trends in our students' academic development. Having those kinds of discussions within the focus group could help

teachers see how the MAP data relates to their classroom data. Many teachers are untrained in data analysis, so it would be beneficial to spend some time discussing the numbers to at least get the stakeholders to consider their own biases and how that could affect the evaluation.

Finally, to best facilitate the discussion, the focus groups and student surveys should be conducted at Colegio Americano. That way if teachers or administrators need to access any classroom information or resources, they would be close. Colegio Americano does have some comfortable conference rooms that have circular tables that would work well for the focus groups, too. With some snacks and the appropriate technology (things like a projector, an Apple TV, and iPads for the participants), the environment would be conducive to conversation and an in-depth exploration of the MAP assessment (Fitzpatrick, Sanders, & Worthen, 2011, p. 439).

[back to Table of Contents]

Reporting Strategy Table

Stakeholder	Reporting Strategy	Implications	Stakeholder Involvement
Students	Newsletter	While this would be a report intended for students, it would be assumed that parents would read this as well. The newsletter should be brief and include some telling charts and graphs to assist comprehension and understanding.	Because these stakeholders are not the decision makers and merely participants, they deserved to know and understand the evaluation, but the proverbial ball will be in their court. These stakeholders will be expected to read the newsletter. The administrators may suggest alternative reporting messages for this population.
Teachers	Verbal Presentation	A PowerPoint presentation complete with visual aids such as charts and graphs will be presented to teachers who teach grades 3-9, the students who take the MAP assessment. The evaluators will present their findings and leave time for a question and answer session.	The question and answer session will allow for more involvement and it will ensure understanding
Administrators	Interim Reports, Verbal Presentation, and Personal Discussion	This group will attend the verbal presentation for the teachers. It will be pertinent for them to hear the questions their teachers still have so that they can assess what kind of assistance teachers still need. However, after that	Administrators will have been receiving interim reports throughout the evaluation process so that they will be informed and aware of the results in case changes need to be made in the evaluation design. Therefore, as a kind of

		<p>presentation, the evaluators would like to meet with the administrators, ultimately, the decision-making group.</p>	<p>summative evaluation, a personal discussion will be held with the administrators and the evaluators to decide the future of the MAP assessment.</p>
<p>Values, Standards, and Criteria: The NWEA's Measures of Academic Progress website offers many resources to help the evaluators and stakeholders understand the MAP scores. Students MAP scores will be measured against the norms from the U.S. only. If the norms are available for international schools (this is a newer development that is likely to be ready for the Spring 2014 MAP assessment), these may also be used to give the stakeholders two different pictures of student achievement. Similarly, if the norms have changed from year to year, those norms should be analyzed with CAT's scores. Nevertheless, the ultimate answer of if the MAP assessment is appropriate for CAT's ELL students will have to be answered by the administrators.</p>			
<p>Potential ethical issues: As a teacher at CAT, I will have to be careful to not let my own biases taint the evaluation results. I have been involved with the MAP assessment for three years and I have seen my own students' scores drop slightly. I will have to take care to be objective and fair. The best way to ensure this would be to have another evaluator working with me. This is a case when it would be ideal to have an "audit trail" (Fitzpatrick, Sanders, & Worthen, 2011, p. 98) of notes and observations so that I can see and my colleague can help me see if my personal views, experiences or expertise were affecting the evaluation. Also, if any of the data from the student surveys is shared, we should get the written permission of the parents. Even though the student surveys will be anonymous, teachers may be able to identify certain students due to combination of grade level and level of comfort with English. We will encourage staff to be discreet with the results.</p>			

[\[back to Table of Contents\]](#)

References

Fitzpatrick, J., Sanders, J., & Worthen, B. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Boston, MA: Pearson.

Kessler, M. (2014, January 6). *Applying MAP Data and Resources*. Symposium conducted at the Winter Workshop of Colegio Americano de Torreon, Torreon, Coah., MX.

MAP. (2014). In *Measures of Academic Progress (MAP) - Computer Adaptive Assessments / Northwest Evaluation Association (NWEA)*. Retrieved from <http://www.nwea.org/node/98>

[back to Table of Contents]