Evaluation of the Massachusetts Aggression Reduction Center’s K-5 Bullying and Cyberbullying Curriculum (2010-2011)

Background Information:

In September 2010, the Massachusetts Aggression Reduction Center published and made available to Massachusetts schools a Curriculum intended to help prevent and reduce bullying and cyberbullying among children in grades Kindergarten through Fifth Grade (the "MARC K-5 Bullying & Cyberbullying Curriculum"). This Curriculum offered 10 age-appropriate Lesson Plans for each of the 6 grades (K-5). During the fall of 2010, a pilot program was launched to systematically evaluate the potential impact of the Curriculum. This evaluation study was conducted in two elementary schools in Massachusetts. The study sample consisted of the children in grades Kindergarten through Grade Five who were given the Lessons Plans in their primary classroom, and who were subsequently observed by their teachers. Every participating teacher completed an Evaluation Sheet following a Lesson Plan, in which they recorded their observations of the children and feedback about the lesson. Thus, this evaluation measured teachers' subjective judgments about the impact of the Curriculum upon their students, in addition to other variables designed to shed light on why they rendered the judgments they did. The two schools described in this report both volunteered to participate.

Schools participating:

The two elementary schools offered very different venues, settings, and approaches. Both were K-5 schools in public school districts in Massachusetts, but one was located in an affluent suburb and the other was located in a working-class urban district. These two Districts and schools differed on several dimensions, including their setting (urban versus suburban), their socioeconomic status (social class), educational backgrounds of the parents in the District, and the myriad of other differences that accompany towns and schools that differ on socioeconomic status. In a sense, these two schools presented an excellent opportunity for comparison; but because they were the only two schools involved, we will not be able to tease

---

1 This report was written by Dr. Elizabeth Englander, Director, Massachusetts Aggression Reduction Center, Bridgewater State University, Bridgewater, MA. Contact us at MARC@bridgew.edu.
apart the separate and distinct effects of income, social class, and type of location (urban versus suburban).

<table>
<thead>
<tr>
<th>Town’s Median home price</th>
<th>% children receiving free lunch</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>School #1</td>
<td>$642,100</td>
<td>2%</td>
</tr>
<tr>
<td>School #2</td>
<td>$143,200</td>
<td>55%</td>
</tr>
</tbody>
</table>

**Variables Measured:**

For several reasons, it was decided not to utilize the number of bullying reports submitted to the school staff or administration as a measure of the Curriculum’s impact. The first difficulty with using such “official” reports is their interpretation. An increase in reporting rates can be seen either as success (e.g., the students and staff are now more aware and reporting more episodes) or as failure (e.g., the number of bullying incidents increased). Likewise, a reduction in the number of episodes reported could be viewed as either positive (e.g., fewer episodes) or as negative (e.g., fewer students reporting). Second, the time frame of the project (a few months) didn’t permit sufficient time for school administrators to thoroughly investigate and determine if all reported incidents were, in fact, bullying episodes. Finally, it’s been well established that too many bullying episodes are never reported to adults at all.

An alternative method for measuring the success of a curriculum is to survey the teachers or the students regarding their perception of the impact of the Curriculum. Granted, such perceptions are strictly subjective. Furthermore, surveying teachers and pupils places an additional burden on the school that has volunteered for a work-intensive project. In this project, it quickly emerged as unworkable to survey the children, since a fairly large proportion of the children involved were pre-literate and would require highly skilled interviewers. Ultimately, the decision was made to survey teachers about their observations of the children, and to survey them not only on their perception of the impact of the Curriculum, but also on the factors that led them to their conclusions.

Following each lesson administered, the teacher completed a Likert-type scale of 1 to 10. The variables that we measured were:

- Teacher's perception of the overall success of the lesson.
- Time of day the lesson plan was administered.
- School (#1 or #2)
- Grade (K-5)
- The engagement of children in the content covered in the lesson, as noted by the number of questions and opinions offered by students about that content.
- The interest level demonstrated by the students during the lesson plan.
• The difficulty students had in understanding the lesson.
• The teacher’s difficulty (if any) with the lesson plan: it’s clarity and ease-of-use

General Findings:

All grade levels participated in the Evaluation of the Curriculum. However, some grades submitted more evaluation sheets than others; Kindergarten teachers participated least, only submitting 10% of the 162 evaluations sent in. Grade 1 teachers submitted 26% of the sheets and were the most prolific in that sense. Grade 2 teachers submitted 20% of the sheets, and the remaining grades submitted 14% to 16% (see Figure 1).

![Diagram showing % of Evaluation by grade]

How Did We Measure Success and Impact?

All teachers were asked to rate their impression of the general success of the lesson plan they had just completed on a Likert-type scale of 1-10. Overall, with all lesson plans and grades aggregated, the average Success Score of the Curriculum was 7.2 out of 10, indicating a moderately high level of success. Scores by grade did not differ significantly, although the differences approached significance (see Figure 2).
Teachers were not asked to merely rate the success of the curriculum in isolation. They were also asked to rate variables designed to measure why they believed that the curriculum had an impact upon the students. Teachers were asked to assess the children on three factors to help determine the overall success of the curriculum:

- Level of interest during the lesson plan being evaluated;
- Level of engagement and participation during the lesson plan;
- Demonstrated understanding of the concepts being taught in the lesson plan being evaluated; and
- Level of teacher’s comprehension and understanding of the lesson plan.

The first of these variables was the children’s level of interest in the Curriculum as it was being presented. Teachers were asked to rate this interest level for each lesson plan presented (see Figure 3). Children’s interest was found to be strongly related to the success score in an Analysis of Variance (see Figure 4).
Another variable that was measured was the children’s level of talking, questions, and participation during the lesson plan (Figure 5). Children’s level of participation was significantly related to the overall success of the curriculum (Figure 6) although even teachers who rated the participation as low tended to rate the curriculum as moderately successful. Clearly, though, high participation was seen as an indicator of high success.
Below (Figure 7) is the distribution of scores received for each rating of the children’s comprehension and understanding of the content being evaluated. There was a notable double peak at scores of “10” (best comprehension) and “5” (moderate comprehension) (Figure 7). How well the children understood the curriculum was clearly related to how successful it was rated (Analysis of Variance) (Figure 8).
The final variable studied was the teacher's comfort level with the curriculum. It is important here to note that during focus groups with the K-5 faculty, several teachers expressed discomfort at the proposal that they would be teaching about cyber-behaviors and issues. This discomfort was expressed in the distribution, which showed a marked “bump” at the score 3 – a score suggesting a low level of understanding (Figure 9). Figure 10 shows that the better a teacher understood the lesson plans, the higher they scored the success of the curriculum; it is interesting to note that even teachers who rated their own comprehensive a diminutive 2 or 3 still judged the curriculum to be moderately successful (Figure 10).
Breakdown of variables by grade

Figure 11 compares the three variables related to the curriculum’s success scores, across grades by type. Children’s Interest and Children’s Engagement varied significantly across the grades ($F=2.45,p<.038; F=4.25,p<.001$), although differences across grades in Children’s Comprehension did not quite reach significance ($F=1.65,p=.153$). Interest and Comprehension followed the same pattern across grades observed in the overall Success Scores, wherein children expressed more interest/comprehension as the grades progressed, with the exception of Fifth Grade (where levels decreased). The pattern for Children’s Engagement was quite different; there, Kindergarteners and Third Graders were the most engaged, followed by Fifth graders, First graders, Second Graders, and finally Fourth graders.
Across most grades, interest, engagement and comprehension were relatively similar. One exception was kindergarten, where Engagement was much higher than Interest and Comprehension; and Grade Four, where Engagement was much lower (Figure 12).

The average teacher’s Difficulty score varied significantly between grades (Figure 13). Kindergarten teachers found the curriculum the easiest and Grade One teachers found it the most difficult (F=3.171,p<.01).
Comparison Between Schools

The two schools that participated in this pilot varied on a number of factors, most of which are typically associated with socioeconomic class. But in addition to the social class and suburban/urban distinction, teachers in the two schools may have differed on level of motivation. In School #2, all teachers who evaluated the Curriculum did so on a strictly voluntary basis. In School #1, in contrast, the school’s Administration volunteered the entire school for the pilot. This was a generous action to take, but focus groups readily revealed that individual teachers did not always feel supportive of the pilot. In focus groups, approximately one-quarter of the teachers from School #1 expressed some frustration with the pilot and the process. Primarily, they were not confident of their ability to deliver the cyberbullying portions of the Curriculum (although no technical skill is required).

School#2 scored the Curriculum higher in every variable excepting Children’s Interest, where both schools were essentially equivalent. All of these differences – none of which were statistically significant - could conceivably be due either to socioeconomic factors or to motivation factors. School #1 teachers did rate the Curriculum as more difficult (although again, that difference was not statistically significant).
DISCUSSION

Overall, the teachers participating in the pilot of this Curriculum judged the lesson plans to have had an impact upon their pupils. The strength of that impact could best be characterized as moderately strong. The factors that led teachers to draw these conclusions were the level of children’s interest in the content of the Curriculum, their engagement and participation during the delivery of the Curriculum, and their lack of difficulty in understanding the content.

A few variations on this general theme are worth noting. First, School #2 found the Curriculum to be slightly more impactful than School #1. This difference was not large or significant, but it was a trend. As briefly discussed above, these two schools varied on a number of variables, most notably their socioeconomic class, their setting (urban versus suburban), and the individual motivation of the teachers rating the Curriculum. Although we had no direct rating on teacher motivation, we do have a measure that is probably related to motivation: Teacher Difficulty (the level of difficulty the teacher had with a curriculum lesson). If School #2 teachers were more motivated, we might expect their Difficulty ratings to be much lower than those from School #1 – but in fact, that was only somewhat the case. Nevertheless, it is reasonable to assume that motivated teachers have more success with any Curriculum. It may be true that urban or less affluent communities will benefit more from this Curriculum – but only future research can fully establish that.
The results of this pilot do suggest that the Curriculum has an impact, but there are important limitations to consider. First, the measure of “impact” was the teacher’s observations of children’s behavior in the classroom, during the discussions that surround the use of this Curriculum. No measures of behavior in more unstructured settings, such as the playground, were taken. Of course, teachers regularly observe their students to assess if new content is understood and has been “absorbed” – this is a central part of teaching any topic. As discussed initially, there are sound reasons not to use official statistics from the schools involved, but a lengthier and more in-depth examination of this Curriculum could benefit from using teacher observations of children’s behavior in a variety of settings – although such observations would also suffer from the fact that much, if not most, bullying behaviors go undetected by adults.

A second limitation is the within-subjects post-test design utilized in this quasi-experimental study. A truly experimental design would have utilized randomly selected subjects, randomly assigned to either a control or experimental group. The quasi-experimental design utilized here is essentially a compromise that enabled us to evaluate the impact of the Curriculum, but our ability to draw definitive, absolute conclusions is limited as a result. One limitation is the possibility of maturation – that the children simply matured during the test period and it was the “growing up” that accounted for the impact, rather than the Curriculum per se. However, the possibility of maturation was mitigated here by keeping the time period of the test quite short (about 8 weeks). A second alternative explanation is the possibility that another variable in the school changed at the same time as the Curriculum, and that this other variable accounted for the impact. An example of such a variable might be increased adult sensitivity to bullying, due to their participation in the pilot program. This possibility is stronger, but if delivering anti-bullying curriculum increases the sensitivity of a school’s adults to bullying, such Curriculum could arguably be termed a success (albeit not in the manner intended by the Curriculum’s author(s)). In this pilot project, changes in the adults as well as in the children must remain as a real possibility.

A final limitation – not of this study but of the Curriculum itself – is the Curriculum’s introduction to cyber-concepts, which begins in kindergarten. Although children are utilizing technology at a very early age in Massachusetts\(^2\) and elementary school teachers are increasingly skilled in traditional content, many appear to feel less certain about their ability to teach cyber-content. During the development of the Curriculum, care was taken to keep cyber-content to a very basic and practical level, but significant anxiety was still expressed during focus groups. This was notable despite lessons that required no particular skills (e.g., in the Kindergarten lessons, children identify letters of the alphabet, and the teacher gives the children examples of words that start with that letter – “M” is for “Mouse” etc.).

\(^2\) MARC Research Report: Bullying in Grades 3-12 in Massachusetts. Online at: [http://webhost.bridgew.edu/marc/MARC%20REPORT-Bullying%20In%20Grades%203-12%20In%20MA.pdf](http://webhost.bridgew.edu/marc/MARC%20REPORT-Bullying%20In%20Grades%203-12%20In%20MA.pdf)
Recommendations:

- The Curriculum appears to have a moderately strong impact upon children in Grades K-5;
- We do not have data on children in other grades (e.g., utilizing the Grade 5 Curriculum in Grade 6);
- Schools utilizing the Curriculum are urged to review it carefully and resolve questions and anxieties in the teaching faculty prior to implementation (a Training Day for the Curriculum will be held at the Massachusetts Aggression Reduction Center at Bridgewater State University);
- Because high interest and participation were strongly related to positive impact of the Curriculum, teachers are encouraged to use classroom discussion skills to maximize such interest and participation, and to allot sufficient time to allow for class debate, discussion, and examples; and
- Schools are encouraged to contact MARC with any questions or concerns about the usage of the Curriculum.