Sartell-St. Stephen ISD 748

Equity Audit by Equity Alliance – Minnesota (EA-MN)

Evaluation and Recommendations for Data Additions/Clarifications

Prepared by:

David Switzer

Department of Economics

St. Cloud State University

This document is the result of a careful analysis of the data, graphs, and discussion included in the Equity Audit provided by EA-MN. After some initial notes on overall impressions of the deficiencies in the data, methodology, and reporting, this evaluation goes through each of the six dimensions listed in the audit and details all of the omissions and errors. In virtually every case, the omission/error makes it impossible to understand what was found, aggregates information so that it cannot be used for school-level improvements, or misrepresents the data to imply that there is a racial or gender disparity in the data that statistical analysis of the data shows does not actually exist. In short, the graphs are flawed and the results are frequently misinterpreted to imply problems that the data do not prove exist.

Here is the basic outline of this Evaluation and Recommendations document:

1. Summary of Recommendations for ISD 748 and EA-MN
2. Abbreviations for Common Issues
3. Evaluation of Graphs, Data, and Analysis of the Six Dimensions in the Equity Audit

This document was prepared by David Switzer and all recommendations are his and his alone. They are not intended to represent the views of any other individuals, nor of the larger Concerned Parents and Community of ISD 748 group.

Summary of Recommendations for ISD 748 and EA-MN:

1. All tables should be labeled clearly and numbered, so that it is easier to reference them in the future. For example, the first graph in Dimension 1 should be labeled "Figure 1.1" and a brief name and description should be provided. It makes it unnecessarily difficult to refer to figures in a report when figures are not numbered and multiple figures have the exact same name.
2. All data required to make inter-group comparisons should be provided. To do this, one needs the number of responses in each group, the mean, and the standard deviation of each. Only with this information can one determine if any reported differences are actually statistically significant. Since NONE of this information is every included in regard to any table, NONE of these tables can be shown to conclusively show that there are any differences across groups.
3. Qualitative results are shown by using 6-10 hand-picked quotes rather than a word cloud to indicate most common responses. We are to presume that the hand-picked quotes represent the larger sample of quotes, but that requires trusting a reporting party that clearly has a financial interest in making things appear one way. A word cloud or some other way of aggregating responses would be an improvement; an appendix that lists all of the comments would be preferred.
4. All rubrics used in the Classroom Walk-Throughs, including criteria and performance descriptors, should be included in the appendix, so the reader knows specifically what can be done differently to earn a higher score on the particular rubric criterion.
5. No information is provided about what determines a "positive" or "negative" comment. Examples of such comments are usually used in reports so the reader has a good idea of the methodology used.
6. Most of the time, results across all levels of school are aggregated together to provide "student" responses. There may be differences across grade levels, as is sometimes mentioned in the report, but data is usually not provided to support these claims. More problematic is that in many cases it would be extremely helpful to have the data aggregated at the school level, not the district level, so that action can be taken at the appropriate level; aggregating everything to the district level implies the problem is broader than it may in fact be, and makes targeted interventions impossible. All raw response data, with student names/IDs removed, would be the most effective way of truly understanding what can be improved at each school in the district.
7. Most "Overall" averages in graphs are NOT weighted by the size of each of the groups in the graph, but rather a simple arithmetic average of the rates of the different groups. It would be like taking a population of 1,000 people consisting of 900 males who had an average positive response rate of 70% and 100 females who had an average positive response rate of only 30%, and saying that the population average was 50%. The actual Overall average is [(900*.7) + (100*.3)]/1,000, which is 66%. All of these Overall

numbers should be recalculated correctly, as they provide a misleading comparison for the subgroups.

I would also note that little if any evaluation of Community responses is contained in this audit evaluation, as there is no way of verifying whether those who answered the survey are actually in the district and not someone outside the district or even possibly someone from EA-MN.

Abbreviations for Common Issues

To reduce the length of this report, I developed some shorthand notation to signify any issues with a specific table:

TITLE: There is no specific title for the graph that helps the reader know what they are seeing.

DESCRIPTION: There is no one-sentence description of the graph that lets the reader know exactly what they are seeing. Is it results from one specific survey question? An aggregate or average of the responses from multiple questions in the survey on a particular aspect of the dimension? There is no way to know.

METRICS: Numbers in the graph are not described in the methodology for the reader to know what they mean. For example, many graphs have three shades of blue in the background, indicating levels 1, 2, and 3, but it is unclear what these levels mean. The cutoffs for the different levels appear to be at 35% and 80%. Is there something actually behind this, or is this some distinction of levels created by EA-MN arbitrarily? Relatedly, it is often unclear what the answer choices were for survey respondents – were they rating statements on a scale from 1-3, and if so, what words accompany those levels (disagree, neutral, agree, for example)?

SUPPORTING DATA: There is not enough information provided (n, mean, SD) for the reader to determine whether differences in groups shown in a table are in fact statistically significant.

MISINTERPRETATION: When enough information is provided to determine statistical significance, interpretation of the results by EA-MN is in fact incorrect. In every one of the four observed instance of data misinterpretation, EA-MN uses differences across groups to imply a racial or gender disparity that statistical analysis shows does not actually exist.

Evaluation of Graphs, Data, and Analysis of the Six Dimensions in the Equity Audit

Dimension 1: Systemic Equity for Students (SES)

Graph on page 8

Issues: TITLE, DESCRIPTION, METRICS, SUPPORTING DATA

- The 1/2/3 have no meaning whatsoever. What are the percentages on the graph (the horizontal lines that differentiate between 1/2/3) actually mean? Where do they come from?
- Based on the explanation, this table appears to show the percentage of each group that "had not directly engaged in any activities to teach about race, culture, or understanding of others." That should be clear in the title, as there are likely other questions asked in the SES section.
- Were students asked if they engaged in activities to teach about race? This can be misinterpreted as actually creating the educational material, rather than "being taught about race." If the same survey, with the same wording, were given to people with dramatically different roles in school, it may be problematic to try to infer differences in responses, as groups may be interpreting the questions differently, as this example tries to illustrate.

First Graph on page 9 – first note that the title and labels are all exactly the same for this table and the one before, so clearly a description or title is necessary so the reader knows the difference between what these two tables are showing.

Issues: TITLE, DESCRIPTION, METRICS, SUPPORTING DATA

- There is no title and there is no clear number provided in the paragraph below that indicate exactly what the question here is. It appears it might be "believe that there is systemic equity for students". This should be in the title on the graph.
- Are the differences between BIPOC certified and white certified staff statistically significant? What are the counts (n) and standard deviations for each group? Without this information, one cannot determine whether BIPOC and white staff actually see things differently or not.

Reference is made to students in higher grade levels reporting low knowledge and awareness of opportunities for affinity or advocacy groups. This is data that we are never provided with, either in a table or graph. To prove that this difference is actually significant, we would need to be provided with the number of responses in all grades to this question, but this information is not in the report.

Second Graph on page 9 – What is the baseline for this? If 2.1% of students had positive comments and 9.6% of students had negative comments, does this imply that the remaining 88.3% of students had no comments whatsoever? One has to go to the appendix to see that this is in fact the case, and to find the number of comments in total. Then one has to do the math to calculate the exact number of comments to determine whether the difference in positive/negative comments is statistically significant; something the report never does; if it had, it would find that only the students are significantly different. Examples of methodology used to determine positive and negative responses would also be helpful here.

Comments on page 10: Without the larger group of data, one has no idea if these comments are representative of the student body overall. At what level of school do these comments come from? It would seem that if there are major problems in middle school but nothing in high school, this would be valuable information. But comments are never broken down by school, so we have no idea at what level to target any action – the obvious result is that the entire district would need the training at all levels, thereby increasing the amount of money provided to EA-MN. A more targeted approach may in fact be more appropriate here, but we will never know with this data.

Staff Survey Comments on page 11: References to "some staff" and "some" other staff that had other feelings. Without any numerical analysis on these comments, it is impossible to judge the size of these "some" statements.

Focus Group Graph on page 13: Despite there being almost 3 times as many positive comments by the community as negative, 3 of the 4 sample comments provided are negative, resulting in a biased sample.

Statements from students on page 14-15 should indicate which level of school it came from so more specific actions can be taken.

Classroom and Building Walk-Through Graph:

Issues: METRICS

- The rubric used, with specific performance indicators, is not provided for any of the six dimensions. What distinguishes between a level 0/1/2/3/4? How is 748 to know what exactly they should do to improve? We have no idea.
- In addition to the histogram for each area of the walk-through, simple averages should also be provided to make it easier to display all the information and summarize findings across dimensions.

Dimension 2: Efficacy of Programs for Students (EPS)

Graph on Page 18: Same issues as **Graph on Page 8** (I wish I could refer to graphs and tables by number, which would make this easier and the report more professional, but this is the only way of referring to a figure).

Issues: TITLE, DESCRIPTION, METRICS, SUPPORTING DATA

- There is zero discussion of exactly what this graph is referring to. It immediately starts talking about the graph on the next page, so we literally have no idea what was measured in this graph (whereas previously we could sometimes infer what the question was, roughly, from the evaluation of the responses in the subsequent paragraph. There is absolutely no way of knowing what this graph is showing.
- It is interesting that at the bottom of page 18, it says "The following graph (pg. 19)…" which indicates this is the only way EA-MN is referring to graphs and tables. The efficacy of their reporting would be helpful if they adopted a Figure 1.2, Figure 2.2, 2.3, etc. approach so that it is easier for both the reader and the reporter to know which figure is being referenced.

First Graph on Page 19:

Issues: TITLE, DESCRIPTION, METRICS, SUPPORTING DATA

- Discussion of the results indicate 71% twice for district leadership, but the only number for district in the graph is 80.95%. So what is this graph showing, exactly?

Second Graph on Page 19:

Issues: TITLE, DESCRIPTION, METRICS, SUPPORTING DATA

- Discussion of the results says 39% of BIPOC students responded a particular way, but the three BIPOC numbers in the graph are 75%, 61%, and 45%. So the discussion here is about something other than what the graph is showing. What exactly is the graph showing responses to? We have no way of knowing.

Comments by Students and Staff: Appendix B informs us that EPS comments were 6.3% positive and 2.8% negative for students, and 10.3% negative and 6.3% positive for staff. These specific numbers are not in the discussion about students, and instead we find out that "not all students felt this way." That kind of a statement is far too vague, as any negative response means not all students felt this way. The impression is given that, when numbers support the narrative that EA-MN is providing, they are mentioned; when do not, they are not mentioned, and the reader must refer to the appendix. Again, despite there being a more than 2-to-1 positive/negative comment ratio, of the 6 comments provided, 4 of them are negative, going 2-to-1 in the opposite direction. Staff comments are actually reflective of the overall numbers positive/negative.

Focus Group Graph on Page 22: Despite all groups having substantially more positive than negative responses (all outside the margin of error of 3% for the number of responses collected), the focus is entirely on the negative responses.

Classroom and Building Walk-Through Graph:

Issues: METRICS (as with all of the Walk-Through Graphs)

Dimension 3: Utilization of Resources (UR)

Graph on Page 27: Same issues as **Graph on Pages 8 and 18**

Issues: TITLE, DESCRIPTION, METRICS, SUPPORTING DATA

- The graph shows one number, but discussion references a variety of aspects: internet infrastructure, equitable use of resources across the district, support for WiFi, etc. This makes me believe that the graph is some average of all of these questions, but averages/aggregates can hide differences across questions, and it is important to have responses to specific questions for results to be useful for improvement.

Since this is the most highly-rated dimension for USD 748, there is very little attention paid to it in the report. No graphs of positive/negative responses by group, and the analysis on page 28 only focuses on the things that are wrong. It is refreshing, however, that the community responses here are all either positive about the school or negative about how the survey/audit is a waste of district funds.

Focus Group Responses: Again, despite positive-to-negative response rates being statistically significant, all attention is paid to negative responses. There is good information here about the other aspects of equity the district wanted EA-MN to investigate, such as ESL and income inequality issues. So while there are no graphs distinguishing anything except respondent type (student/staff/community), race, or gender, there is information in the audit about some of these other issues; it is all anecdotal, however, so of limited use.

Classroom and Building Walk-Through Graph:

Issues: METRICS (as with all of the Walk-Through Graphs)

Dimension 4: Curriculum Represents Equity (CUR)

Graph on Page 33: Same issues as **Graph on Pages 8, 18, and 27**

Issues: TITLE, DESCRIPTION, METRICS, SUPPORTING DATA

- It is unclear what this graph is showing. It appears that it is some measure of how well the curriculum represents equity (hence the name of the dimension), but unclear if that is one question or an aggregate of many. Clearly, there are multiple questions here, as the discussion references the percentage of people who believe students know about each other's families (is this a curricular matter?).

First Graph on Page 34:

Issues: TITLE, DESCRIPTION, METRICS, SUPPORTING DATA, **MISINTERPRETATION**

The report says: "The results demonstrate that a respondent's demographic affected if one was exposed to a diverse and culturally responsive curriculum for students and staff." No data is provided regarding staff to support this claim. As for students, the differences of 2.66%, 3.19%, and 2.33% for responses by BIPOC and white students (across the three levels of school, respectively) are NOT statistically significant for the sample size surveyed. The margin of error is 3% for the high school students and 4% for the other two groups (since fewer students responded at those levels). None of these differences is statistically significant, so a respondent's demographic DID NOT AFFECT if one was exposed to a diverse and culturally responsive curriculum for students.

- This is so far the most misleading graph in the report, as the claim about students is flat out wrong, and the claim about staff is not supported by any data contained in the report.

Second Graph on Page 35:

There are surprisingly no issues with this graph, except it is not numbered. The title explains what it is, the graphic and metrics are easy to understand and the discussion of the graph is consistent with what the data shows.

Focus Group Responses: Interestingly, Appendix A shows that 23.3% of students had negative CUR comments on the survey itself, while only 3.8% had positive comments – this is the largest disparity between negative and positive student comments for any dimension. This is switched entirely for the focus group comments, which are 2-to-1 positive for students. Perhaps there is an issue with the survey questions in this category, as when students are able to articulate things to each other in context (in the focus group), they are much more positive about the levels of equity and diversity in their curriculum.

Classroom and Building Walk-Through Graph:

Issues: METRICS (as with all of the Walk-Through Graphs)

Dimension 5: Student-Centered Leadership (SCL)

Graphs on Page 42: Same issues as **Graph on Pages 8, 18, 27, and 33**

Issues: TITLE, DESCRIPTION, METRICS, SUPPORTING DATA

Something does not add up here. The second graph says there is a 65.51% overall positive response rate. This graph breaks down types of staff and adds in Parent and Community (P&C) members. There were 544 P&C members responding to the survey, while only 472 staff members took the survey. The first graph indicates that P&C had an 87.07% positive response rate, while staff had only a 61.92% positive response rate. Adding these groups together would provide a collective 75.4% positive response rate overall. Yet the second graph says the overall average is 65.51%. Even if all the students are included in the "Overall" rating, that would make it 69.4%. So where is 65.51% coming from? EA-MN takes the average of the response rates across all the different staff demographic groups, completely ignoring the different sizes of those groups, and take a simple average. It is hard to express how wrong this "Overall" number is, and how it further erodes any credibility EA-MN has regarding their use of data.

And they do the same thing in the third graph, completely ignoring the fact that

a) There are no statistically significant differences between BIPOC Elem, white Elem, BIPPOC Mid, and white Mid. However, both BIPOC HS and white HS are significantly lower than the previous four.
b) High school students represent 42% of district students and 38% of survey responses, yet are weighed equally in the "Overall" rating because of the simple arithmetic weighting of averages.

MISINTERPRETATION: The discussion says that the student's racial identity had an impact on sense of community, but at every level of schooling, the difference between positive response rates of BIPOC and white students is less than the margin of error. Differences are all less than 1.5%, while margin of error is 3% or 4% depending on the school. In short, a racial difference DOES NOT EXIST HERE, yet discussion of results states that it does.

The ONLY statistically significant difference among students here is that all high school students, regardless of race, have less positive responses to SCL than the other two levels of schooling. This implies there is a problem in the high school that could use some fixing, but the problem is present for both BIPOC and white students.

Focus Group Responses: It would be instructive to see how the positive and negative responses broke down by school, since the previous data clearly indicates that there is a much more significant problem with SCL in the high school than in the other two schools.

Classroom and Building Walk-Through Graph:

Issues: METRICS (as with all of the Walk-Through Graphs)

Dimension 6: Inclusive Communication (COM)

Graph on Page 49: Same issues as **Graph on Pages 8, 18, 27, 33, and 42**

Issues: TITLE, DESCRIPTION, METRICS, SUPPORTING DATA

First Graph on Page 50:

Issues: TITLE, DESCRIPTION, METRICS, SUPPORTING DATA

- It is unclear what this graph is showing. It breaks down responses by race (BIPOC vs. white) and then shows the number of responses per location. Is this the number of "safe" responses? Is this the number of total responses? It appears it's the number of safe responses, but we don't know how many people responded to each location, which may differ across schools. Positive response rates should be used instead or in addition to the current graph to account for this.
- **MISINTERPRETATION**: Discussion states that BIPOC middle school students feel least safe in the locker room and the counselor's office. But in fact, the number of white students reporting feeling safe is also lowest in those two locations (except for Parking Lot and PROO, which I presume do not exist in some schools, since the numbers are either missing or extremely low overall). Thus, EVERYONE appears to feel the least safe in the locker room and the counselor's office, not just BIPOC students. Yet again, the data are misinterpreted to imply a racial problem that accurate evaluation of the data shows does not exist.

Second Graph on Page 51: MISINTERPRETATION. The graph is used to say that "gender also correlated to the students[sic] sense of safety in school." In fact, in this graph, sometimes the Female response bar is slightly higher and sometimes the Male response bar is slightly higher, but in every case but one (Classroom, where females feel safer), this difference is not statistically significant given the 3-4% margin of error in a survey of this size. The only gender difference this data shows is the male students feel less safe in the classroom.

Classroom and Building Walk-Through Graph:

Issues: METRICS (as with all of the Walk-Through Graphs)

Appendix A:

The graph showing the six dimensions is baffling. How is the overall ranking of the 6 dimensions determined. Why does it go up to 8? Is this the overall average for the questions in that section, divided by 10? That might make sense, as the average for UR is roughly 75%, and that dimension gets a 7.5 on the scale. This should be clarified so the reader knows how the numbers are calculated.