HANDLING MESSY DATA

Overview
Sometimes the most illuminating data is also the messiest: conversations. Conversations occur both in oral form (e.g., interviews, informal discussions, focus groups) and written form (e.g., journals, open-ended surveys, other expository writing). Because conversation data is not easily quantifiable (i.e., reduced to numbers), it's tempting to shy away from using them. But if we really want to understand and be able to address important issues related to school performance, we must venture into this type of data collection and analysis. This tool provides general guidelines for how to handle messy data and a template of an analysis table for the data.

Source
This tool was developed by WestEd facilitators.

Materials
- Multiple copies of collected data
- Example: Handling Messy Data – Notes
- Example: Handling Messy Data – Analysis Table
- Template T3.3d-1: Messy Data Analysis Table
- Colored highlighters, sticky notes

Time Required
Depends on the data source and amount collected.

Directions
This tool assumes you have already collected some “messy data” in the form of written transcripts or notes from focus groups, interviews, etc. or have compiled the written input from open-ended survey questions or journals. Make enough copies of your data for everyone involved in the analysis. Below are some general instructions for how to compile, summarize, and describe data from conversations.

What to Do with Conversation Data

1. Make sure all data is in written form.

It is difficult to detect patterns in conversations without having the data in a form that can be read and re-read. To get the best results from spoken conversation, capture the data in writing during or directly after the conversation. When conducting formal interviews (i.e., scheduled talks typically guided by pre-defined questions) it may be advisable to take notes during the interview or, depending on the circumstances, immediately after. If subjects are comfortable being taped, consider using a tape recorder and transcribing the conversation afterwards.
Whether taking notes or taping, you can help make interviewees or focus group participants more comfortable by assuring them in advance that nothing from the conversation will be attributed to specific individuals and that all information is subject to agreed-upon confidentiality rules (see section on confidentiality below). Also, before starting a focus group, request that participants not share what goes on in the focus group with anyone outside the room.

When putting the data into written form, make sure to organize them by interviewee or speaker. Although you will not be identifying a speaker by name (and perhaps not even by role type) in any subsequent report, it is important to have speakers identified when looking for patterns. Assign a number (e.g., Speaker 1, Speaker 2) or use abbreviations (e.g., E.A.) to denote speakers.

If the data are already in written form, as in open-ended survey questions, organize the written comments into a list format by question or prompt. Do not include the contributor’s name, but include other demographic information, such as gender, grade level, or department.

2. Read and re-read the data holistically.

This is done to get a flavor of what is being said across questions and interviewees or respondents. The challenge in using conversational data is to identify patterns, themes, and new insights. Use a sticky note to highlight places you know you’ll want to review more carefully. Before getting started with analysis, make sure you know the data well.

3. Capture patterns and themes in the data.

After reading through the data a couple of times, you will likely recognize some emerging themes (i.e., ideas, thoughts, or information that occurs repeatedly across speakers and/or across questions). The next step is to capture these themes. Using colored highlighters, assign a color for each theme. For example, if you are looking for information about how the faculty is utilizing the new text adoption, you might highlight in blue each time they mention a specific way they are using it. Or if you are looking for the reaction to the new text adoption, you might highlight in yellow each time they share information about this. Do not limit your search for information on a particular topic to an interviewee’s direct responses to a question about that topic; instead, look across the entire written document for information on that topic and highlight every mention, using the same color. See Example: Messy Data – Notes.

4. Note frequency of themes.

Look across speakers to see how many people said the same general thing about a given topic; record this number. To do this, use the Messy Data Analysis Table (provided below) to list the questions — or the topics of the questions — you are investigating. Note in short form (e.g., keyword) what speakers said about the topic or question at hand, differentiating who said what by putting a speaker’s initials or assigned number (e.g., Speaker 1) next to your note about his or her comment. Also, list the number of speakers who mentioned a given idea more than once (e.g., repeated this same information in response to another question). Add any notes or comments associated with the theme or pattern. See Example: Messy Data — Analysis Table.
5. Note different emphases given to themes.

When interviewing individually or in focus groups, you will see that speakers feel more strongly about some issues or ideas than about other issues or ideas. This is also important to capture. For example, all members of a focus group may mention that the new schedule doesn’t give them enough time to plan as a team on a daily basis. Yet in the same conversation all of them may enthusiastically report that the new schedule gives them more time to work with individual students and talk about how this is increasing students participation in class. One set of statements seems to support the new schedule while the other does not. But the collective enthusiasm shown about the schedule’s positive impact on students seems to suggest that this is more important to focus group members than the fact that there is less time for team planning. This type of information goes beyond the “how many” question to get at the emotional tone of the conversation. Capturing the emotional tone or emphasis is not always possible, especially if the data are in written form to begin with, but there are times when it is very clear. Focus on the times it is clear. On the analysis form, list any notes or comments associated with speaker emphasis.


The power of conversational data is that they allow us to find the unusual or unexpected. Sometimes an unexpected comment or idea may give more insight into the issue than all the patterns, themes, and numerical data put together. The focus here is not on the one person who is complaining and bitter about a change or on another single individual whose enthusiasm for the change is over the top. Rather, it is on data of any sort that may deepen understanding (see the Insights section of Example: Handling Messy Data – Analysis Table). Pick a color of highlighter you have not yet used and re-read the data looking for statements that are quite different than the patterns or themes you have already identified, that go beyond the superficial answer, that are on another topic but closely related, and/or that for some reason stick out to you. Probably you have already noted at least one outstanding idea during your initial reading. On the analysis form, note any insights in the space provided in a shorthand form that captures the gist. Make sure to note the page where the entire comment appears so you can go back to it later.

Anonymity and Confidentiality

If people want to participate candidly in any conversation, they must feel safe in doing so. How freely they speak is related to their feelings about the person who is conducting the interview or focus group and their understanding of whether their statements will be attributed to them or used anonymously. Depending on the situation, including the topic, people may be less willing to openly share their opinions and thoughts if they think they will be quoted afterwards or will be judged by the interviewer. There are different ways to address this issue. First, if you believe either the context, the particular situation, or the given topic is particularly controversial, it might be best to have someone from the outside come in and conduct interviews or focus groups, providing you with only the analyzed data and not the actual transcripts. If you choose this option, be sure to tell those being interviewec or participating in the focus group how their data will be communicated (i.e., anonymously). Another option is to pose your questions in written surveys and ask participants to type their responses so their handwriting won’t be recognized. Even if the issues are not controversial, you still must not divulge the identity of speakers to anyone, either in the analysis or in any other
communication. *In the analysis, only patterns and insights are shared, not who said what.* You must also be careful in your analysis that the speaker cannot be identified by process of elimination (e.g., if there is only one kindergarten teacher and the analysis is reported by grade level). Whatever the situation, speakers must remain anonymous, and if you hope for candor in any planner conversations, you must orally promise this in interviews, focus groups, and before administering a survey. Without confidentiality you are less likely to get honest responses and if you betray speakers' confidentiality, future attempts to get information will be severely compromised.
EXAMPLE: HANDLING MESSY DATA — NOTES

Teaching English Language Learners
K-3 Teacher Focus Group with 7 Teachers

1. To what extent would you say that you are familiar with the English Language Development standards and their relationship to the English Language Arts standards? How helpful have the standards been in guiding your instruction of English language learners?

   T1 — Looked them over but don’t know enough especially for doing the report card.
   T2 — A lot of classes but need to do more.
   T3 — A lot of reading but need to integrate more.
   T4 — Piles of things to look over.
   T5 — Still need to work on both.
   T6 — I didn’t realize there was supposed to be 30 minutes of English Language Development in the classroom.
   T7 — There is a standards-based report card. Wish there was a way to indicate progress for English Learners without making them so low.
   T3 — Do we have to put the standard on the lesson plans?

2. Have you participated in professional development related to English Language Development instruction?

   T2 — The university provided a class in the different Limited English Proficiency levels and what kids need to know at each level. The person teaching the class wasn’t a good English model and the class wasn’t very helpful.
   T3 — County office offered class on grouping and language level. Considered helpful.
   T1 — Looking for more school-wide. Teacher needs are still there — It looks like the same river but it’s not.
   T5 — The expectations for what teachers need to do have changed over the years.

3. Do you feel the instructional materials you have for teaching English learners are adequate? To what extent are you able to use grade-level materials with your English learner students?

   T6 — Have an English Language Development book in third grade.
   T2 — K doesn’t have one.
   T4 — The language or words are hard for all kids. (All other teachers agree)

4. In your experience, which teaching practices are most effective for helping English learners learn English? For helping them learn academic content? To what extent are you able to incorporate these strategies in how you teach?

   T5 — English Language Development — hands-on, phonics.
   T1 — Group work, modeling.
T3 — Repetition, kids repeat back.
T4 — Academic Content – computer technology since it is very motivating.
T2 — Frontloading the lesson with a small group and vocabulary building, partner work, drawing.
T6 — Incorporate these regularly.
T7 — I use them all the time.

5. What have you found to be the strongest aspects of your school’s English learner instructional programs? What areas are most in need of further development?

T5 — The combination of interventions, after school, aides all contribute to success.
T1 — Time on task – kids are always working on something meaningful.
T3 — There is no wasted time.
T4 — The staff level of expertise translates to high expectations for English learners.
T6 — The standards (English Language Development and English Language Arts) both need more work.
T3 — Some attitudes need to change.
T2 — There is some stereotyping among coworkers.
T1 — Hear a lot of complaining about those kids.
T7 — Staff need to put themselves in the children’s shoes — hear kids are lazy, they’re not trying hard enough.
T5 — Teachers are caught between making kids produce and a perceived lack of support for kids from home.
## EXAMPLE: HANDLING MESSY DATA — ANALYSIS TABLE

<table>
<thead>
<tr>
<th>Theme</th>
<th>Frequency</th>
<th>Emphasis</th>
<th>Notes</th>
</tr>
</thead>
</table>
| Don't know enough about standards (ELA and ELD).  
*These statements are highlighted in green.* | 6/7 — T1 twice; T3 twice; T6 twice |                                                                         |                                                                               |
| County office training helpful.  
*These statements appear in pink.* | 1/7 — T3  |                                                                         | When asked about professional development only one teacher mentioned something helpful. Another teacher mentioned something the university did as not helpful, and another said they need training schoolwide. Later T4 contradicts self, saying there is a high level of teacher expertise. |
| Language or words in instructional materials are hard for students  
*These statements appear in blue.* | 7/7 — all teachers agreed when T4 stated this | Teachers were very emphatic; only time they all agreed. | Is this associated with the materials or with teachers' ability to teach the vocabulary? |
| Group work effective teaching practice for English learners.  
*These statements appear in grey.* | 4/7 — T1, T2, T6 and T7 said they incorporated all the mentioned strategies |                                                                         | It appears that teachers don't have a large repertoire of strategies to call on. May indicate need for more PD. |
| Negative view of English language learners by some staff/low expectations. | 4/7 — T1, T2, T3, T7 | These four teachers felt strongly about this. | T4's statement contradicts this. He believes there is a high level of teacher expertise and high expectations for learners. |

**Insights:**
T5 made a comment about teachers feeling "caught" between making kids produce and a perceived lack of support for kids from parents/home. This was made during the talk about attitudes toward students. *Wonder what this means? Need to look into this more.*

T6 said K teachers have no English Language Development materials. *Need to check into this.*
# Messy Data Analysis Table

<table>
<thead>
<tr>
<th>Theme</th>
<th>Frequency</th>
<th>Emphasis</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insights:</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>