

CONCEPT PAPER 17

Figuring Out What's Going On For People: Simple Inductive Thematic Qualitative Analysis

David Wedaman

March 10, 2018

davidwedaman.com

Synopsis:

In this concept paper I argue that it is crucial to try to find out what your staff are thinking and feeling, and I propose a simple way to achieve that end: adding open-ended questions on staff surveys. Such questions allow people to speak in their own language, which frees them to say things you may not expect, but which also makes analysis somewhat more complicated than using more common, quantifiable responses. However, analyzing qualitative data is not that hard, and it is worth the time investment; this paper goes on to explain how to conduct a relatively sophisticated, but quite simple, inductive thematic qualitative analysis, using a spreadsheet.

Do I Care About What's Going On For People?

Basically, yes, you should. It is a commonplace that the people you work with will tend not to want to tell you what they are thinking or feeling. Especially if there is a potential conflict related to any information they might share: if it might anger you, or undermine agreements, or suggest change, or even just create some kind of unpredictable outcome. This goes double if you are in a supervisory relationship with them. This reticence seems to be a fundamental characteristic of all human organizations from families to major corporations. I won't go into why this is the case except to say that many thinkers suspect it has something to do with largely unconscious anxieties about the assumed consequences--to everyone--of being open.

But I will say that it causes no end of problems to the organization, and it appears on every level. On the level of trust: if you don't

communicate, trust is undermined. Or on the level of learning: if you don't give feedback, both individual and team learning are short-circuited. Or decision-making: if you don't share your data, reasoning, or claim, it will be hard to make good decisions with you. Or justice: if we can't point out places people are being treated unfairly, we can't help the moral arc of the universe bend in the direction it should. In short, honest, open, forthright communication--knowing what is going on for people--is crucial to organizational success at every level.

However, this is not easy. Sometimes you just happen into a context where a team is at ease, trusts each other, is comfortable working out problems, likes to learn, isn't afraid of looking bad, isn't too awed by power relations, and so talking and sharing comes easily and naturally. That, in my experience, is as rare as spun gold. If you don't happen to have fallen right into the perfect workplace dynamic, you'll need to employ other means.

With reason you will wish to work towards a long term goal of creating just such a culture as would make people feel comfortable speaking their truths. But in the meantime, you still need to get as much of a sense of what people are thinking and feeling as you can, because you need that data in every way and context. So you will likely need to find hybrid, interim, experimental, or creative techniques to collect this information--techniques that make it easy on people to share--even though the culture around sharing may not quite be there. These techniques can take a lot of forms, from paying close attention to tone and body language, to anonymous suggestion boxes, to classroom response systems.

A very common and practical approach is to use a survey. Surveys, especially the online versions, have the advantage of being easy to create, easy to administer, easy to fill out, anonymous, familiar, and something people can do at their convenience wherever and whenever they like. Surveys also have some built-in limits, though, that can work against your purposes, and actually block thoughts or feelings from getting to you: so this paper addresses this. I share a simple way to use a survey, and without a lot of work, turn it into a thoughts and feelings gathering machine of the most productive power, using a particular kind of analysis to make sense of the data you gather.

What's Wrong with Surveys?

Nothing is wrong with surveys. Any attempt to talk to people about their experience is a good one. And the recent proliferation of online surveys are easy to use. The trick is that most surveys tend to collect quantitative information, or ask people to choose from preselected answers. Right? They ask you to rate things on

a Likert scale of 1 to 5, or they ask you to choose from a list of possibilities.

This is not horrible; the problem, though, is that when you use pre-selected answers, the survey design itself limits the ways the respondent can answer. I have to choose from a list of possible ideas: I can't say what I think in my own way. The imagination of the person who designs the survey becomes a kind of fence around my self-expression. What if I have something to say that doesn't fit into the answers? And isn't it especially important to know what people think, when what they think doesn't match what you, the designer, expect?

There is definitely a convenience to analyzing the results of the kinds of surveys that limit your answers: you can do it almost instantaneously, by glancing over the data: 5 out of 6 respondents were "highly favorable," etc. This might not be a bit deal if the survey isn't that important--if the data you want are fairly simple, or obvious, or you aren't trying to dip a little more deeply into what is really going on for people. But as I note above, you need to be dipping more deeply whenever you can. Fortunately, to dip a little more deeply isn't that hard. You can give people room to say whatever they think by using open ended questions where they can actually type whatever words or phrases come to them. You do have to think about your questions a little bit: you'll want to give some direction, but not too much, if you really want to invite people to share their thoughts. A common question that works: "How Can We Improve?" "What are We Doing Well?" "What is one thing we can do that will make your work-life easier . . . or more meaningful . . . or less frenzied?" Etc.

But Chaos Will Ensue!

I think there is a kind of knee-jerk fear of opening up meetings or surveys or

conversations or any communication format to the free participation of all parties: one fear is that you may hear things you won't like; the other is that you won't be able to handle or sort through the complexity of the variety of responses you might get.

I don't really address the first fear here: except to say, that you *probably* will hear things that you don't like, and that you *still* need to hear them, especially if you don't like them. I will speak to this particular point in a future paper, but here I will just draw on my years of developmental coaching to say that you needn't worry: I guarantee that you will be able to handle hearing things you don't like.

This paper specifically addresses the second concern. I have a simple way for you to draw on proven methodology to handle all the different sorts of answers you will hear. It ends up that people in the Social Sciences have been thinking about this particular problem--making sense of things people say--for a while. They have techniques, and we can use one or two. In particular, we can apply what is known as Inductive Thematic Qualitative Analysis to help us make sense of the data we gather from short-answer questions on surveys, and it's not that hard. It's even fun.

Inductive Thematic Qualitative Analysis?

Yes, I have been advised not to use phrases like "Inductive Thematic Qualitative Analysis," because it stresses people out, and they seek to put distance between themselves and me as quickly as they can. This is, however, the official name of what we will be doing: I think it's good to be precise. On the other hand, it's not that complicated, even if it sounds fancy, so please don't distance yourself yet. What "Inductive Thematic Qualitative Analysis" means is basically this: it's a way to make sense (that's the analysis part) of language (that's the

qualitative part). "Thematic" comes from the fact that it looks for themes, or patterns. And it starts from patterns *in* the language it looks at, rather than starting from patterns that come from somewhere else. That's the inductive part.

You Already Have the Tools!

The good news is that, although it sounds abstract and theoretical to say you'll be "looking for patterns in language," it isn't really either. In fact, you already know how to do it; we are all sophisticated language processors. Basically we go around finding patterns in language all day long. I will give you an example. Take the sentence "I wish we had more opportunities to speak directly to our CEO." That is the kind of thing people tell you when you start to invite them to tell you what is going on for them, isn't it? Let's make sense of that sentence. It basically has one idea: who ever said it wants to be able to talk to the CEO. If you collected, say, 10 people's feelings about their work, and two said something like this, you would have a theme that came from your data, and you would have found it there because you let people say things in their own way, and then you paid attention. You would also be able to say in your report something like this: "using inductive thematic qualitative analysis, we determined that 20% of our respondents were interested in being able to talk to the CEO."

So Let's Do It

This is the point of the paper where we've done enough talking and should start doing things, so let's do that. Let's assume you are going to use a simple online survey to gather short-answer, free-from answers like the one above. One line responses. I'm going to assume you can design the survey and get the results into a

spreadsheet of some kind, like Google Docs, and that you have some familiarity with spreadsheets, so that I can give some generic instructions. With those assumptions in mind, here is a how-to guide to using simple inductive thematic qualitative analysis. For the record, although I'm making this process accessible, I'm not veering far from the main road of qualitative analysis; and to that end I'm using a published and respected guide as a methodological support. I think you'll find that these instructions align well with the guidelines in the following excellent article on this subject: Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3 (2). pp. 77-101. ISSN 1478-0887.

Step 1: Gather the Data

First step, ask a question that generates some open-ended responses. Cut and paste these responses in Column A of a new spreadsheet, one per row.

Step 2: Create a Header Row and Name your Data Column

Expand the width of the column so that you easily read the lines. Insert a header row and use it to name the first column "Data." For our purposes, we're assuming a general individual response length not much longer than this very sentence. And that ought to be able to fit in one line in one column that isn't above half of your viewable spreadsheet's width.

Step 3: Add Code Columns

Now create four columns to the right of the data column, in columns B - E, and name them "Code 1, Code 2, Code 3, Code 4" in the header row.

Step 4: Add Code Book, Definitions and Tabulation Columns

Add a blank column to the right of your code columns, in column F, simply for visual clarity, then add an additional column called "Code Book," in Column G. After it add a column called "Definitions," in Column H. Give this column a little width, because you'll be defining terms here. And finally, add "Tabulation," in Column I. You are now effectively prepared to start your inductive thematic qualitative analysis!

Step 5: Start Coding

We peeked a little bit at how you make sense of language in qualitative analysis a few paragraphs back. Now we'll do it for real. The process is not too intimidating: basically, you are reading each response, and looking for its key ideas. These ideas we will give a one-word name that makes sense to you. And that name we call a "code." As the response will only be a sentence, it won't have more than one or two ideas. Each idea you identify goes into a code column. Hopefully you won't find more than four codes in any one response, because if you do, you'll need to insert a fifth code column to capture it.

Let's go back to our example sentence above: "I wish we had more opportunities to speak directly to our CEO." The main idea here

is that the respondent wants to speak to management; so we need a kind of general name for that idea. I like “feedback.” You may also want to capture the fact that the CEO is mentioned by name, so a second code could be “CEO.”

Step 6: Keep Coding

So you coded your first sentence: now keep going. Go through a lot of them. If it's a small set, say, under 25, you might be able to code them all quickly. If it's more, you may want to break after, say, 50. As you're coding your sentences, reuse codes where you can: those are your patterns starting to emerge. If someone else mentions the CEO, for instance, reuse the code you invented for that: you're starting to see a CEO pattern. You may find that you have a lot of different codes and no immediate pattern to any of them. Don't worry! It's a first pass. This is an iterative process.

One rule of thumb: don't use the same code in a sentence more than once. It's not inherently bad to do this: clearly an idea can be expressed twice by a person in a sentence. But one of the methodological decisions we've used to frame this simple analysis is that our “context” for each code is essentially the full, brief sentence it came in. The code marks up the full answer text; and that's as fine a point as we'll use. That lets us do interesting things in our reporting step, which I will explain in due time.

Step 7: Start Thinking About Patterns

So now you've coded at least a good chunk of sentences. It's time to step back and think about the codes as codes. Do you see

patterns? Are particular codes used a lot (that's good; it means you've hit on something)? Are some one-offs? Are some “mergable?” You may find that some of your codes are very similar: think about whether you might create a kind of uber-code that contains one or more of the similar codes, and so lets you see a larger pattern across all the multiple data points. For example, in our hypothetical data set, we might find that another manager is named explicitly in another response, say, the CFO. At that point we can decide whether we would like a more general code, say, “Manager,” that replaced the codes “CEO” and “CFO,” and lets us represent a more general pattern.

Step 8: Iterate

The process of coding is iterative. That is, you can expect to do a few passes, each time coming to know your data set better, and thinking more about what your codes measure, how they form patterns, and how you want to frame those patterns. As we showed above, whether you want to talk about individual managers or managers as a group determines whether you would code “CEO” or “Manager,” or even both. (Because you can code one idea in two different ways.) You might also go the other direction, and break a code into smaller, more detailed codes, because as you got to know the data set, you realized there was ambiguity there that you want to reduce.

Step 9: Think about Manageability of Codes

You've probably realized that there is a sweet spot to the number of codes you use, and how much they represent. You could mark up each sentence with 6 or 7 codes, using such

specificity in your coding that few codes ever repeated. That would be an extreme in which you wouldn't be able to see much of a pattern across responses. The other extreme is just as bad: that would be to use so few codes so often that each code comes to represent too much. As you are coding and looking back over your codes, you are going to naturally feel a kind of sweet spot and find your way to a manageable set of codes that also means something important.

Step 10: Develop the Code Book

Our next step is into the realm of code definition. The "code book" is like a little dictionary of your codes. Here you list the code and describe what it means to you. You'll see that a code book quickly becomes necessary as you find yourself parsing possible meanings in different ways. Now is a good time to start your code book, as you have gone through an iteration or two of data coding cycles, and are probably starting to hone in on how you are going balance manageability and meaning in your codes.

To create your code book, list all the codes you use in your codes columns in the Code Book column. Each code you use should appear in the list. Then to the right, in the Definitions column, explain as concisely as you can what the codes mean. For our "CEO" code, we might say, for instance, in the "Definitions" column, "The CEO is mentioned."

Important point: now that you have started to define your codes, if you make any changes in the coding to the left, you will need to make sure you record that in the code book. For example, if you decide to combine two codes into one, broader category (which happens a lot), you'll need to replace the two

sub codes with the new code, and create a new definition that explains the larger category.

Step 11: Consider Another Set of Eyes

You can iterate the previous steps as much as you like to get the right feel of codes that both mean and are manageable. An important thing to also consider at some point is whether you want to invite someone else to look at your codes with you. My advice is that you should: another person validating your codes--that they make sense, that they are consistently applied, and looking for anything you overlooked--makes your analysis better. Plus it lets you be more confident that other people are likely going to understand your analysis, since your assistant did: and this is the hallmark of thoughtful work--that you can explain to people how you came to the conclusions you came to.

Step 12: Count Your Codes

Let's say you've coded everything, and you like the codes you chose--they seem to capture key patterns, and they show up in meaningful ways. You also have a manageable set: they aren't so diffuse that they defy description: in fact, you may even be getting excited about telling the story you see in the data (that's a good sign). We won't do a whole lot of elaborate mechanical analyses in this paper: but we will do one: we'll count how many times the codes appear.

Here's where using a spreadsheet comes in. If you were consistent in using the same spelling each time you used a code, including the listing in the code book, and no invisible spaces have crept in anywhere, you

can now with a fairly easy formula have the spreadsheet count the occurrences of a code.

There are probably a few ways to count the number of occurrences of codes: in a recent case, I used this equation in a Google Doc in the Tabulation column:

=countif(\$B\$2:\$E\$80,G2)

This basically works by looking to see how many times the code in the particular Code Book column cell (G2) appears in the range of codes columns (B2 - E80). This range represents the columns where you list your codes: we start on row two, because we have a header row, and I picked 80 as a random end row: yours may go further.

After you add an equation like this to your tabulation column, and you start to see the frequency with which your codes appear, you'll immediately start to see new ways to think about your codes: some will pop out because they are more frequent; some will recede.

And here's the trick I alluded to previously: because we kept to one instance of a particular code per response, we can say that the number of times the code appears represents one person expressing that idea. 25 "CEO" codes means 25 people said something about the CEO. So it's easy to start to make some sense of our patterns by seeing how many people talk about which ideas. And it's easy to talk about the data.

Step 13: Look for Other Themes!

So you've coded everything in your data column. And you've counted your codes, so you know which are more prevalent. You already have a story to tell that gets at deeper meaning than surveys normally do, because you let your respondents tell *you* what they think, and you're

using their own language to tell that story: it's not a story you're making up, or suppressing, or fencing in somehow. You're use Simple Inductive Thematic Analysis to figure out where people are coming from!

You may not be satisfied with looking at your codes through a lens of frequency. While this is usually enough for most people in relatively informal situations to feel they've done their duty to a data set, you may want to do more. If so there are all sorts of other ways you can look for meaning; the article I cite in my references is not a bad place to get a sense of what's out there, and how people *think* who make sense of how other people think. But you can actually do a lot of this yourself, trusting your intuition.

There are two easy simple ways to think about finding themes. The first way is "meta." You can look across your codes, and ask yourself, do these *codes* connect to each other in any larger patterns? And would those be interesting to think about? For example, let's say you were collecting codes on the different ways people asked to communicate to management. And let's say you had something like this: from a survey with 60 respondents, you have a code "Email," with 24 occurrences, a code "Phone," with 10, and a code "In Person," with 8. Looking at those codes individually lets you see a marked preference for email as the communication method: that's good to know already. But taken together, there is another message: they are all people asking to communicate, regardless of method: and together they add up to 42 out of 100 people asking to communicate: and that is significant in a more general sense.

There's the "micro" angle, as well: you can ask if any of your codes really are worth breaking into smaller units. Let's say we coded our example sentence about talking to the CEO with the code "Talk," which meant that the respondent mentioned talking to the CEO in

any way whatsoever. But perhaps we noticed as we were coding and tabulating that some of these “Talk” codes indicated people wanted to talk, and some of the codes indicated people *didn’t* want to talk. Both pro and con effectively meet the definition of “about talking to the CEO,” because we weren’t that particular in our code definition. In this case we may realise it makes sense to break this code into sub-codes like “Pro-Talk” and “Anti-Talk,” so that we can capture that important nuance in our report.

Step 14: Report!

We’ve made it to the fun part: having collected our free-form responses, popped them in a spreadsheet, configured our spreadsheet, coded, tabulated frequency of codes, and looked across codes for larger themes, we can probably start to tell a story about what the data are saying. This part I won’t describe in this paper, but it basically involves you explaining what you think the data mean, using your codes, code definitions, frequency numbers, and perceived themes as your evidence.

It may help you to start your report by writing about one code at a time, telling the story of that particular code; and to do this it may help you to look at all instances of it in your data, to remind yourself of what people were saying. In our spreadsheet this is easy to do, by searching for the code name, toggling through the highlighted instances of the code, and reading the related sentences.

On the other hand, it may help you to start with the major themes you are seeing, and then bring the coding evidence in as you need it. Let your composition style be your guide.

Conclusion

We’ve made it to the end. In this Concept Paper we’ve gone from some rather abstract theory--explaining why it is important to include some way to gather people’s thoughts and feelings *in their own words* into your workplace activities--to some concrete, hands-on work with a simple way to do just that: a spreadsheet designed to help us apply simple inductive thematic qualitative analysis

If this has been helpful and interesting to you, let us know by visiting davidwedaman.com.

References

Braun, V. and Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3 (2). pp. 77-101. ISSN 1478-0887.

This article is intended to be an informal, informational white paper reporting on consulting engagements, research, or thinking done by David Wedaman; it is governed by a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license. Identities are intentionally obscured. For more information on our consulting and research services see davidwedaman.com.