Content analysis

Content analysis is a very broad category of approaches to the analysis of qualitative data where the purpose is to describe written, spoken or visual communication. The simplest example of content analysis is by word count: how often does a particular word appear in a particular type of text? For example, this can be used to explore racism or sexism, by counting how many times racist or sexist terms appear in a given type of document. This approach is different to discourse analysis, which focuses on how meaning is made in social and linguistic contexts.

This can also be applied to interviews or spoken communication as well, so you could analyse broadcast interviews with politicians, for example. It can apply to topics as well as words, so you could identify how often a given topic, such as motivation or pay, appears in internal company management memos, no matter what words are used to describe it. In this example, you can see that content analysis is being used to gather quantitative data.

However, it can also be used to generate qualitative data. Where the analysis of content focuses on understanding the meanings underneath the surface of the text, implications, references, allusions to other texts and ideas that are not spelled out, or any other type of hidden meaning, this is a qualitative approach to content analysis. The rest of this guide focuses on the explicit approach to content analysis. If you are more interested in hidden, implicit meanings, then refer to thematic analysis.

Applications

Any method of communication, whether verbal, visual or textual, can be analysed on the basis of its content. So that could include (but is not limited to):

• photographs • product labels • advertisements	• poems • letters • novels	• newspaper articles • classroom materials • song lyrics
• phone calls • conversations • meetings • interviews	• press releases • videos • plays • TV programmes or films	• websites • social media • brochures

Content analysis can be used to see how things change over time. For example, you could go into an archive of material such as comic books from the 1960s to the present and see how the artists' representations of violence change. You would have to set a fixed criteria for this, in order to make this a quantitative analysis. For example, you could count how often in a given issue of X-men blood is shown, or a

limb is broken. You could also look at how vivid the colours are, or how much of a given page or frame is occupied by the injury.

Content can also include themes, such as achievement or conflict. For example, you could analyse TV adverts and see how often conflict between couples is shown.

Often content analysis is done on existing materials, like children's stories, curriculum and lesson plans, music videos, song lyrics, and so on. However, it can also be done on data that the researcher collects, through interviews or other means. For example, if you were an education student on a placement and as part of your lesson you encouraged the children in a class to write a story about themselves, you could (subject to permission from the school and parents) conduct a content analysis on the children's stories. You might be interested in their use of language (number of grammar errors, for instance, or use of particular vocabulary), or you might be interested in the narrative of their story (who appears in the story, how many children mention animals, is there a villain, and so on). Refer to the Qualitative Methods guide under Research Methods - any data collection technique mentioned there can be analysed using content analysis.

Setting the criteria

This is the most important element of a content analysis, because clear definitions of your criteria and categories mean that other researchers could repeat your study and get similar results. This is a key measure of quality known as 'replicability'. Your categories will be determined by your research questions and your theoretical approach.

If you are doing, for example, an analysis of representations of women's bodies in magazines targeted at a female audience, you might define the following categories:

Category	Definition
Health	Mention of physical health or well-being, other than weight-loss. e.g. vitamins, nutrition, etc.
Weight-loss	Tips, images or discussion of weight loss
Strength	Discussion of or images of physical strength e.g. performing a physically challenging act
Attractiveness	Images or discussion of aspects of beauty related to sexual attraction
etc.	

You might develop these categories yourself on the basis of an initial reading of your sample (this is an inductive approach - one that puts the data first). Or you might find

when conducting your literature review that other people have done relevant studies, which give you some themes or clues about what to look for. Alternatively, you might have a theory in mind which provides you with a clear framework (this is a deductive approach - one that puts theory first).

For example, you might be doing research on how managers motivate employees and whether this still reflects Maslow's theory, the hierarchy of needs. You can use Maslow's hierarchy as your categories, and look for examples in what managers do that reflects these categories.

Category	Definition
Self-actualisation	Reaching one's full potential. E.g. manager encourages self-reflection, improvement, leads continuing professional development.
Esteem	Respect, confidence, independence. E.g. manager praises team member and encourages workers to be independent.
Love/belonging	A sense of community, interpersonal relationships. E.g. manager encourages new member of staff to socialise, introduces them around the office, leads team bonding exercises.
Safety	Personal, financial, health, care against accidents. E.g. manager reassures employee about job security or takes health and safety actions.
Physiological	Relating to survival e.g. food, air, water. E.g. manager buys pizza when team works late.

Maslow, A.H. (1943) A theory of human motivation. *Psychological Review,* **50** (4), pp. 370–96. Retrieved from http://psychclassics.yorku.ca/Maslow/motivation.htm

NB. Although a classic, please note that Maslow's theory has been widely criticised and the use of this as an example should not be taken as an endorsement.

The clearer your definitions are to start with, the better. However, if they need to change to reflect the reality of what you are seeing, then do so. Make a careful note of how the categories have changed so that you can refer to this in your methodology. If you are using a theoretical approach, this is actually a very important finding because it shows that the theory may not be adequate to explain the situation.

Selection

What is your **unit of analysis**? For example, in the table above looking at women's magazines, both images and text are referred to. This suggests that the unit of analysis is each image or sentence. A broader analysis might take the entire article as the unit of analysis, if it was only interested in the main topic of the article, not the details of ideas used. Similarly, in the Maslow example above, the unit of analysis seems to be individual actions taken by the manager. So the act of buying pizza is a single unit of analysis. You need to decide this before you start, and identify it in your methodology.

What is your **sample size and approach**? in other words, consider how many of each unit you need to make a reliable statement, and what other factors might influence the outcomes. In the women's magazine example, you might think that the season would influence how they talk about women's bodies, and so might the individual publication. So you would need to build a sample around those two variables (and you might already be thinking of even more). You can compose a table like this:

Magazine	Summer	Winter
Grazia	3	3
Vogue	3	3
Marie Claire	3	3
Cosmopolitan	3	3
	Total	24

These numbers might seem small, but remember that in this case, our unit of analysis is every single sentence and image included in the entire magazine, potentially even including advertisements, which takes the page number total up very high. If you were looking at the articles as the unit of analysis, this might be considered too small a sample.

Similarly, for the Maslow-manager example, you need to consider how long or how many times you would need to observe a manager in order to catch a range of these actions. In some businesses, this might mean that you would need to get them at the right time of year - for example, accountants are always busiest in April at the end of the tax year, so if you want to see them under stress, that's the best time to go. Alternatively, you might consider that you want to observe several different managers in different types of business to get a wider overall picture. If this project is starting to sound unmanageably big, then think about other ways you could get this information. You could actually do interviews with managers and/or staff, and design your

questions to elicit the information that you want e.g. 'When type of action do you take to motivate your staff?' Then their responses can be grouped into your categories.

Procedure

Content analysis can be applied to any type of qualitative data. You can use it to analyse your interviews, to identify patterns in blogs, tweets, photos on social media, selfies, Tumblr tags, adverts, music videos, song lyrics, phone calls, historical documents - the list of potential data sources is endless.

However, remember that just because it is on social media doesn't necessarily make it public. If you want to analyse people's Facebook updates, you must get their permission. Ethical standards still apply, refer to the University's Ethics code and procedures.

A rough procedure for content analysis would be:

- establish permission and access to the data
- establish your categories (if using a deductive approach i.e. you already know what they will be)
- work out your sample size and approach
- collect and capture your data using notes, photos, copies, scans, voice recordings - whatever is appropriate
- establish your categories (if using an inductive approach i.e. you will work them out based on the data)
- code the data - find instances of your categories and note them down
- analyse your findings
- interpret your findings in relation to theory and previous research - what do they mean?

For example, here is how you might code and record examples of managers' actions from interviews. The numbers relate to the time on the recording of the interview. A blank space means no information.

	Employee 1	**Manager 1**
Esteem	2:56 manager praised work on report 5:34 manager sent email to say well done on presentation	11:02 emails regularly sent to employees 12:29 weekly team meeting, round the table thanks
Belonging	20:23 company retreat, manager led exercise in team bonding	

You might be thinking that this is a time-consuming way to record your findings. If so, you might consider using NVivo which is a software programme designed to help with

qualitative data analysis. It allows you attach labels like these categories to elements of text, images or recordings, and lets you import web pages, Tweets, images and many other document formats directly. For more information, see the guide on NVivo.

Analysing your results

Remember that describing what you found is not enough. This is what you will need to do in the results section, but in the discussion section, you need to say what it means. So, for example, if you find that Grazia talks about weight-loss significantly more than it talks about women's strength, whereas Marie-Claire mentions health more than weight-loss, what does that mean?

Also, note that if you are doing a quantitative form of content analysis (i.e. counting instances), you will be expected to do more than show bar charts describing the data. You also need to consider this when you plan and design your data collection. Refer to our support guides on quantitative data analysis for more information. You will need to compare your results to previous research, explain whether and why your results are similar or different, and make connections to relevant theories.

So in conclusion:

Content analysis is where you analyse communication by defining categories for what you are looking to find out and take a systematic approach to identify where these categories appear.

Please note that you should NOT rely entirely on this guide to write your methodology. You will be expected to do further reading. Here are some suggestions:

Further reading

Grbich, C. (2012) *Qualitative data analysis: An introduction*. London: Sage. Chapters 16 & 17 refer to content analysis.

Krippendorff, K. (2012) *Content analysis: An introduction to its methodology*. London: Sage.

Robson, C. (2011). *Real world research: A resource for social scientists and practitioner-researchers*. 3rd ed. Chichester: Wiley.

Silverman, D. (2011) *Interpreting qualitative data: A guide to the principles of qualitative research*. London: Sage.

Other useful guides:

Click here for more Research Methods guides.