

Answers to end of chapter questions

Chapter 1

What are the three most important characteristics of QCA as a method of data analysis?

QCA is (1) systematic, (2) flexible, and (3) it reduces data.

What are the three phases in the development of quantitative content analysis as a research method?

First phase: early applications, lasting from approximately the eighteenth until the early twentieth century;

Second phase: content analysis coming into its own, lasting until the late 1940s;

Third phase: interdisciplinary and methodological elaboration, still ongoing.

What were Kracauer's main points of criticism of quantitative content analysis?

Meaning is often complex; it may be latent (i.e. not immediately obvious); and it may appear only once in a given text.

Name four characteristics that distinguish QCA from quantitative content analysis.

Any four from among the following are correct: QCA focuses on latent meaning; it takes context into account; reliability is handled less strictly than in quantitative content analysis; validity checks are just as important as the assessment of reliability; it is at least in part data-driven; inferences to the context, the author, or the recipients are more frequent than in quantitative content analysis; and there is more flexibility in going through the sequence of steps.

Chapter 2

Which characteristics does QCA share with other qualitative research methods?

Interpretation: It is applied to symbolic material that requires interpretation and it focuses on meaning (personal or social);

Naturalistic: It does not change your data.

Context: Context is taken into account in arriving at interpretations.

Reflexivity: The positions of others should be taken into account when creating a coding frame; and you should make the grounds for your interpretations transparent.

Inductive: Coding frames in QCA are at least in part data-driven.

Validity: This is an important criterion in evaluating your coding frame.

In what respects does QCA combine features of qualitative and quantitative research?

QCA allows you to examine many different features of your material (holistic), but you have to decide on one meaning for each feature (reductive).

It aims for a balance between the general (categories as a higher level of abstraction that allow you to compare) and the specific (trying to preserve as much of the unique meaning as possible).

It combines linear elements (sequence of steps) and cyclic elements (such as: repeatedly going over your material when developing your coding frame; modifying your coding frame; etc.).

It is typically part data-driven (describing your material) and part concept-driven (as in drawing on an interview guide when developing your coding frame).

It is variable-oriented (with each main category functioning as a variable), and at the same time allows for an in-depth description of cases by combining a large number of variables.

Reliability and validity are of equal importance as evaluation criteria.

Chapter 3

Name four or more differences between coding and QCA.

Any four of the following are correct. The first four are the most important differences.

Coding is an analytical method; its focus is on how categories relate to each other. QCA is more of a descriptive method; it focuses on how data relate to each other.

In coding, codes are typically data-driven only; in QCA, categories are typically part data-, part concept-driven.

Coding is an iterative, cyclic method, and you will continue to revise your codes throughout the analysis. QCA combines linear and cyclic elements, and you will try to finalise your coding frame as early as possible in the research process.

Coding and QCA emphasise different quality criteria. In coding, the focus is on trustworthiness and credibility, whereas in QCA consistency is also important.

In coding, creating and applying your codes are collapsed into one step. In QCA, you first develop your coding frame, try it out and modify it, and only then apply the categories to your material.

In coding, the focus of code definitions is on the conceptual level, whereas in QCA it is on how to recognise instances of the category in your data.

In coding, codes need not be mutually exclusive. In QCA, all subcategories within one main category should mutually exclude each other.

Segmentation (i.e. dividing your material into units of coding) is an important step in QCA, but not in coding.

Which research questions are suitable for QCA and coding respectively?

QCA is a suitable method with descriptive research questions, including comparative research questions, and coding is suitable for analytic research questions, especially those that involve developing a theory.

What are the most important differences between discourse analysis and QCA?

Discourse analysis is based on the constructivist notion that language shapes social reality. QCA, on the other hand, does not make any explicit assumptions about the nature of social reality, language, and their interrelations. Implicitly, it is based on realist assumptions.

Discourse analysis can be either descriptive or critical, whereas QCA is typically descriptive.

In discourse analysis, the focus is on processes of reality construction. QCA, on the other hand, is used to describe static manifestations of social reality.

In using discourse analysis, you pay as much attention to the ways in which language is not used, what is not being said, as to what is said. When using QCA, on the other hand, you focus on what is present in your material.

How can discourse analysis and QCA be combined despite these differences?

First, you can place QCA in the service of the critical-interpretative attitude underlying discourse analysis, i.e. you can use QCA as one of your methods in conducting discourse analysis.

Second, you can conduct a critical QCA, where categories refer to the ways in which something is expressed in your material.

Name three or more differences between semiotics and QCA!

Any of the following three are correct:

Semiotics is concerned with providing a detailed description of the (denotative) meaning of your material, whereas QCA provides a description of the meaning of your material in selected respects only.

Semiotics provides an in-depth description of the individual specifics of each instance in your material. When applying QCA, on the other hand, you focus less on individual specifics; instead, you compare individual instances and thereby reduce your data.

In semiotics, the plausibility of your interpretation is the decisive quality criterion; consistency between different researchers in their interpretation is not expected. In QCA, on the other hand, consistency between researchers is an important quality criterion.

When using semiotics (similar to using discourse analysis) you pay attention to what is absent in your material as well as to what is present; in QCA, you focus on what is present.

When doing a semiotic analysis, you take into account the relationship between one sign (your focus) and other signs, i.e. you look at one sign from multiple perspectives. In QCA, you can

assign only one meaning to any section of your material from any given perspective (with subcategories having to be mutually exclusive).

What kind of material is especially suitable for doing a semiotic analysis?

Semiotics is especially useful for analysing material which is intended for a wider audience.

With this kind of material, the cultural context of any one sign is more salient than in individual communication.

Chapter 4

What is the difference between main categories and subcategories in a coding frame?

Main categories (also called dimensions) are those aspects in your material that you want to know more about. Subcategories specify what is said about these aspects (relevant meanings).

In what sense does one have to make choices when building a coding frame?

You cannot include all the meanings in your material in your analysis. First, you have to decide on certain aspects about which you would like to know more. You then go on to build your coding frame around these aspects (main categories). Second, you have to make choices by focusing on certain things that are said about these aspects (what you find relevant: subcategories). In this way, you reduce the variety of meanings in your material to the distinctions specified by your main categories and your subcategories.

What does the complexity of a coding frame depend on?

In conceptual terms, the complexity of your coding frame depends on your research question. In practical terms, it depends on the number of dimensions (main categories) you include in your coding frame and on the number of hierarchical levels in the frame.

Describe the structure of a coding frame of medium complexity.

Coding frames of medium complexity can either consist of one dimension that reaches down through several (more than two) hierarchical levels, or else it can consist of several dimensions that comprise only two levels, i.e. several simple coding frames 'strung together'.

What are the requirements that you should keep in mind when building your coding frame?

Unidimensionality: Each dimension in your coding frame should capture only one aspect of your material.

Mutual exclusiveness: Subcategories within one dimension should be defined in such a way that any unit of coding can be assigned to one of these subcategories only.

Exhaustiveness: A coding frame is exhaustive if all units of coding in your material can be assigned to one of the categories in the frame.

Saturation: A coding frame is saturated if each category is used at least once during the process of analysis. This criterion only applies to data-driven coding frames (where it is met by definition).

Other important criteria in evaluating coding frames are reliability and validity. But these cannot be assessed as you are building your coding frame; they can only be assessed once you have applied the frame to parts of your material.

Chapter 5

Which strategies are available for distinguishing between relevant and irrelevant parts of your material?

A first strategy is to use a coding frame to differentiate between relevant and irrelevant parts of your material. This requires that you define what you understand 'relevant' and 'irrelevant' to mean.

A second strategy is to make decisions about what you consider irrelevant in a research team.

What is meant by using concept-driven and data-driven strategies for building a coding frame?

Using a concept-driven strategy for building your coding frame means to base your frame on previous knowledge. You can do this by drawing on theory, on the results of prior research, by drawing on everyday experience or on logic. If you have used a topic guide to collect your data, you can also use this guide as a deductive framework.

Using a data-driven strategy means to create your categories from your data, depending on what you find in your material.

Name three strategies that are useful when building a data-driven coding frame.

Any three out of the following are correct:

Aggregation (also called: progressively summarising your material)

Subsumption

Contrasting

Adapting coding from Grounded Theory

What are the four parts of a category definition? Which of these are mandatory, and which are optional?

The four parts of a category definition are: a name, describing what you mean by a category, examples, decision rules. A name, description, and examples are mandatory; examples can be

hypothetical if necessary. Decision rules are only needed where there is overlap between categories.

What is a decision rule, and what information should it contain?

A decision rule is needed when two categories (potentially) overlap. A decision rule tells the coders which out of two overlapping categories to use. It should specify what is not to be included in a category and what category to use instead. Sometimes negative examples are also included, i.e. examples of units of coding that are not covered by the category in question.

Why is it necessary to revise your coding frame? How often do you have to do this?

To revise your coding frame means to go over your main categories and subcategories in structural terms (checking for overlaps between categories, for example), to 'tidy' it and remove any loose ends. This is necessary because the more complex a coding frame grows, the easier it is to lose sight of the overall structure as you keep on adding more categories. You have to revise your frame as many times as you expand on it, i.e. add to it a substantial number of new categories.

Chapter 6

What are the steps in generating data-driven categories through progressive summarising?

Step one: paraphrasing all parts of your material that strike you as relevant;

Step two: reducing each paraphrase to the main statement;

Step three: summarising similar paraphrases; to be repeated until you reach the desired level of abstraction;

Step four: generating a category name.

Which steps in open coding can be adapted to creating a data-based coding frame? What does each step involve?

Conceptualising: In conceptualising, you look at your material from up close, pinpointing relevant concepts and gradually coming to identify similarities and differences. In QCA, this corresponds to identifying categories in your material.

Defining categories: This involves identifying similar concepts, grouping them together into categories, and defining categories in terms of what the categories have in common.

Developing categories: Developing categories in Grounded Theory refers to identifying the ways in which these categories vary. In QCA, this corresponds to arranging your categories in a structure involving main categories, subcategories, and hierarchical levels.

What are the steps in generating categories by subsumption?

Preceding subsumption: deciding on a main category, i.e. a perspective under which to examine your material;

Step one: Scanning your material for relevant concepts and creating a subcategory to cover the concept;

Step two: Scanning your material for relevant concepts until you reach a relevant passage;

Step three: Checking whether this is similar to or different from other passages that you have already looked at;

- If it is similar: Mentally subsuming the concept under an already existing subcategory;

- If it is different: Creating a new subcategory;

Repeat steps two and three until you reach the end of the material that you are using for building your coding frame.

What are the steps in generating categories by contrasting?

Step one: Identifying similarities with your first source and creating a subcategory for each similarity that you identify;

Step two: Identifying similarities with your second source and creating a subcategory for each similarity that you identify;

Step three: Identifying differences between the two sources and modifying your categories so that they capture these differences.

Which of the four strategies are especially suitable for creating an entire coding frame the data-driven way, including both main categories and subcategories?

Progressively summarising your material;

Adapting coding from Grounded Theory.

Chapter 7

What is segmentation?

Segmenting means dividing your material into units of coding such that each segment/unit fits into exactly one (sub)category of your coding frame.

What is the difference between units of analysis, units of coding, and context units?

Unit of analysis: This is the unit you have selected for QCA, each unit yielding one text. The unit of analysis will often be identical with a 'case'. Typically, each unit of analysis will contain several units of coding.

Unit of coding: Units of coding are those parts of your units of analysis that can be interpreted in a meaningful way with respect to your categories and that fit within one subcategory of your coding frame. Only units of coding are marked in your material.

Context units: The context unit is that portion of the surrounding material that you need to understand the meaning of a given unit of coding.

Give some examples of formal criteria for segmentation!

Chapters, sentences, clauses, paragraphs, articles, 'turns'...

When would you use a thematic criterion for segmenting your material?

It is best to use a thematic criterion for segmentation if your material has no clear, inherent structure.

What are the three steps involved in segmentation?

Step one: Marking the relevant parts of your material;

Step two: Deciding on your criterion of segmentation;

Step three: Marking your units of coding.

Describe one strategy for segmentation when using a thematic criterion!

Any of the following would be correct:

You as the researcher identify and mark all your units of coding.

You start out by identifying and marking the first units of coding together with any other coders.

Once you agree on what a theme is and where one theme begins and ends, you then go on to identify and mark the remaining units by yourself, getting back to the other coders only occasionally to make sure that you are still applying the group consensus.

All coders (including you) identify and mark all units of coding independently from each other. They then compare and resolve any differences through discussion.

All coders (including you) go through the material step by step: First you identify the next unit of coding independently from each other, then compare and reach an agreement. You then go on and independently from each other immediately code this unit, then get together again and discuss your codes. You then turn to the next unit of coding – and so on.

All coders (including you) start out by identifying and marking the first units of coding. Once you agree on what a theme is and where one theme begins and ends, you then divide the material among each other. Each coder goes on to identify and mark the remaining units in this part of the material independently. At regular intervals you get back together and decide jointly on a few units of coding to make sure that you are all still applying the group consensus.

Chapter 8

What does the pilot phase involve?

The pilot phase has three parts: (1) the trial coding, i.e. applying your coding frame to a part of your material, either with two coders or at two different points in time; (2) the consistency check, i.e. comparing the two rounds of coding; (3) adjusting your coding frame.

How much of your material should you include in the trial coding?

The decision of how much material to include involves a trade-off between the criteria of variability and practicability. On the one hand you should select your material so as to include as much variability as possible and to be able to apply as many categories from your coding frame as possible. On the other hand you will probably have to re-code a substantial part of the material that you include in the trial coding, at least to the extent that you later adjust your coding frame. Between 10% and 20% of your material will often constitute a reasonable trade-off between the two criteria. But ultimately the decision depends on the size of your coding frame and the degree of variability in your material: The larger your coding frame and the more variable your material, the more material you should include in your trial coding.

What would be good ways of subdividing the coding if your coding frame contains more than around 40 (sub-)categories?

You can either subdivide the coding by dimensions only or according to units of coding and dimensions simultaneously. If you subdivide the coding by dimensions only, you code all your material first on dimension 1, next on dimension 2, and so on. If you subdivide the coding by units of coding and dimensions simultaneously, you start out by dividing your material into smaller thematic parts (for instance your participants' responses to one of your interview questions). You then code this part of your material according to one relevant dimension after another. Next, you continue with the second smaller part of your material which you again code on one dimension after another.

What would be a good way of preparing for the discussion between the coders after the trial coding?

It would be useful to create and fill in a comparative coding sheet (with units of coding as rows and coders or points of time as columns), noting down the category to which each unit of coding was assigned by each of the coders or at each of the two points of time. This is a good basis for identifying differences between the two rounds of coding, i.e. instances where the unit of coding was assigned to different categories, and highlighting these. This helps to focus the following discussion.

Name two of the reasons why coders may have assigned a unit of coding to different categories.

Any two of the following are correct:

One of the coders has made a mistake.

The coders agree on their understanding of a given unit of coding, but differ in their understanding of the potentially relevant categories.

The coders differ in their interpretation of a given unit of coding.

What do you do if two coders differ in their interpretation of a unit of coding?

The coders should adjust their understanding of the relevant categories, if necessary, and you should modify the coding frame, clarifying definitions and adding decision rules to deal with potential overlap between categories. During the pilot phase, no further action is needed, and there is no need for coders to reach agreement.

Chapter 9

What is meant by reliability?

An instrument is said to be reliable to the extent that it yields data that is free from error.

How do you assess the reliability of a coding frame?

To assess the reliability of your coding frame, you look at the consistency of coding between different coders or at different points in time. You can quantify the degree of consistency by calculating a coefficient of agreement, or else you can discuss any differences in the research team.

What are the two considerations you should keep in mind when interpreting a coefficient of agreement?

First, the degree of consistency to be expected depends on how 'standardised' your material is. In qualitative research, standardisation is typically low, a fair degree of interpretation is needed, and therefore lower consistency is to be expected. This is even more so with large coding frames.

Second, you should take into account the distribution of instances of inconsistency across your material in interpreting a coefficient of disagreement. Revising your coding frame is most necessary where inconsistencies accumulate.

What is meant by validity?

An instrument is considered valid to the extent that it captures what it set out to capture. A coding frame is valid to the extent that the categories adequately represent the concepts under study.

Which type of validity should be used in evaluating data-based coding frames, and what are the telltale signs of low validity?

To evaluate the validity of data-based coding frames, the concept of face validity is most useful. Telltale signs of low validity are: high coding frequencies for residual categories; high coding frequencies for one out of several subcategories within a single dimension; underdifferentiated abstract categories.

Which type of validity is most suitable for assessing the validity of concept-based coding frames?

Content validity is most useful in assessing the validity of concept-driven coding frames.

Chapter 10

Imagine that you are working on your own, but a friend of yours is willing to help you with the coding. What part of your material would you ask your friend to code, to be able to compare your codes with hers? What would your decision depend on?

More important than the amount of material coded by your friend is the variability of that material – this is the first criterion for deciding. Try to include material from all the different sources you are using. The second criterion is how much time your friend is willing to invest. In sum: Include as much material as possible from as many different sources as possible, without overburdening your friend.

Imagine that you realise during the main coding that your coding frame does not allow you to describe important parts of what your participants are saying. Because of this, you add new categories to your coding frame. What do you do now? Can you continue with the main coding? Why (not)?

If your frame does not allow you to describe important parts of your material, it does not have high validity. If you simply continue with the main coding at this point, your results will only be of very limited value. It would therefore be best to turn the main coding into a second trial coding, followed by another round of revising your coding frame. Only then would you start again with the main coding.

What strategies are there for handling cases where two coders cannot agree on the meaning of a unit of coding?

You disregard this unit of coding altogether.

You take turns between the interpretations of the two coders, alternately entering the meaning assigned by the one and the other coder into your coding sheet in cases of disagreement.

You bring in a third person with expertise on the research topic to decide between the two interpretations.

You value the interpretation of one of the coders over the other's interpretation.

Under what conditions is it necessary that you transform your results onto the level of your units of analysis?

This transformation is necessary whenever your units of analysis are larger than your units of coding.

Imagine that you have been working with a coding frame that consists of three dimensions. The first dimension contains 5, the second dimension 8, the third dimension 4 subcategories. For each dimension, more than one subcategory can be coded for each unit of analysis. How many columns are needed in your transformed matrix to represent the codes for these dimensions?
17, with one column for each of your subcategories.

Imagine that you also want to integrate repetitions of subcategories into the above matrix for the three dimensions. How do you represent information about code repetitions in this kind of matrix?

The number of columns remains the same. In each column you enter the coding frequency for the respective subcategory.

Chapter 11

What are the main strategies for presenting your findings in qualitative style?

Describing and illustrating your findings using continuous text;

Describing and illustrating your findings using text matrices;

Doing additional data exploration and analysis and presenting these results.

What is a text matrix, and how can you make use of text matrices in presenting your findings?

Text matrices are tables that contain text instead of numbers that are used alongside continuous text. They allow you to summarise and at the same time illustrate selected parts of your findings. They can be used as follows:

To illustrate categories;

To describe a small number of cases in depth (or for illustrating a specific case);

To compare cases.

What does looking for patterns and co-occurrences involve?

In looking for patterns and co-occurrences, you focus on how your categories are related, asking questions such as: Do specific categories occur together or near each other? Do selected categories stand in a specific relationship to each other, for instance one category serving as an indicator or a consequence of another?

What are the main strategies for presenting your findings in quantitative style?

Providing absolute frequencies;
Doing descriptive group comparisons;
Using inferential statistics.

Name three ways of presenting absolute frequencies in your results section.

Presenting absolute frequencies as part of a continuous text;
Integrating absolute frequencies into text matrices;
Reporting absolute frequencies in charts.

In what respect do you have to be careful not to overinterpret your findings when comparing descriptive coding frequencies across groups?

First, chances are that your sample for each of the groups is not in fact a representative one. Second, descriptive coding frequencies do not tell you whether any differences do in fact hold beyond the sample in the population. To determine whether any differences are 'true', 'significant' differences, inferential statistics are needed.

Chapter 12

What types of software are there for doing textual analysis?

There are three types of software for analysing text:
Software for qualitative data analysis (CAQDAS software);
Software for quantitative content analysis;
Software that integrates features of qualitative and quantitative data analysis.

Which of these can support you in doing QCA?

Software that integrates features of qualitative and quantitative data analysis best supports QCA, although sometimes such packages lack flexibility. Software for qualitative data analysis also supports many of the steps in QCA.

Which steps of QCA are especially well supported by software?

Building your coding frame and trying out different ways of structuring the frame in this process; selecting cases for and doing the trial coding; assessing the validity of your coding frame; evaluating your coding frame for mutual exclusiveness of your categories, exhaustiveness, and saturation; revising your coding frame; analysing your results for patterns by looking for co-occurrences; presenting your results in either qualitative or quantitative style.

How do you go about segmenting your material into units of coding using software?

If you are using software for qualitative data analysis, you have to use the coding facility to segment your material. You do this by marking the segments on the screen and assigning 'unit codes' that indicate on which dimension a unit is to be coded. All units that are to be coded on the first dimension, for example, are assigned the code 'unit dimension 1'; all units to be coded on the second dimension are assigned the code 'unit dimension 2', and so on. If you are using software that integrates qualitative and quantitative features, you will have to prepare your file so that segments are clearly indicated, for example by dividing segments from each other using '/'.
/.

Which strategy for interrogating your dataset is especially helpful in doing QCA?

Looking for co-occurrences within the same unit of analysis is especially helpful.