QCAmap Step by Step – a Software Handbook

Philipp Mayring, Klagenfurt/Austria 2020

Introduction: Foundations of Qualitative Content Analysis

(from Mayring, Philipp (2019) Qualitative Content Analysis: Demarcation, Varieties, Developments [30 paragraphs]. Forum: Qualitative Social Research, 20(3), Art. 16, https://dx.doi.org/10.17169/fqs-20.3.3343.

I want to outline the most salient points in the characterization of the type of content analysis that we have developed (MAYRING, 2015; MAYRING & FENZL, 2019).

- The qualitative content analysis' approach (as with the quantitative content analysis) is category-based, that is its distinguishing feature. Categories refer to aspects within the text, which put the meaning of those aspects in a nutshell. Text evaluation is, therefore, restricted to the selected category system. Text contents that are not addressed by the categories or holistic impressions are not taken into account or would have to be addressed with other text analysis methods.
- The qualitative content analysis procedure is research question oriented. Text analytical questions (possibly several) are derived from the main aims of the research project. These questions should be answered at the end of the analysis. This clearly distinguishes the qualitative content analysis from other completely open, explorative methods such as grounded theory.
- Qualitative content analysis is characterized by strict rule management and systematic. Process models enable the procedure to be described step-by-step, and this has proven itself in countless research processes. The specific rules for each technique are reviewed in a pilot study and should not be changed after that.
- I have described several specific evaluation options as part of the qualitative content analysis. Recently, I proposed and developed eight techniques (MAYRING, 2015):
  1. Summarizing
  2. Inductive category formation
  3. Narrow context analysis
  4. Broad context analysis
  5. Formal structuring
  6. Content structuring
  7. Type-building content analysis
  8. Scaling structuring.

  Besides the techniques above, there are also mixed variants. In other places (MAYRING, 2014), I have referred to structuring as an ordinal or categorical deductive category application. Further, I have made reference to type building and content structuring among others as mixed techniques. The decision for a specific content-analytical technique depends on the formulation of the research question. It is possible to use individual techniques alone, but also several techniques can be used simultaneously during one of the iterative steps in the content analysis.

- The content-analytical rules for the individual techniques are not arbitrary but have a solid theoretical foundation in the processes of everyday text analysis. In particular, regarding how these processes have been examined in cognitive psychology and psycholinguistics. For summary and inductive category formation, these are reductive operators (omission, generalization, construction, integration, selection, and bundling; MANDL, 1981), on which the gradual reduction of text segments is based. For explications, it is rather context theories from linguistics. Whereas for deductive category applications, reference is made to the categorization theories from general psychology and language development research (MAYRING, 2014). The result of this was that the exact wording for a human-readable general category requires an explicit definition (definition theory), a cognitive anchoring in typical examples for the category (prototype theory), and rules to demarcate the categories from one another (decision bound theory, MURPHY, 2002). These coding guidelines are the basis for the three-part coding—definitions, anchor examples, and coding rules—applied in the procedure that I recommended (MAYRING, 2015). Thus, when trying to determine content-analytic rules, I try to use strategies that draw on the everyday handling of texts, a method that is common in qualitative research, for example, when linguistic approaches to storytelling in everyday life are employed in the rules for narrative interviews.

Software tools for qualitative text analysis

Within the last 30 years many approaches of computer assisted qualitative text analysis had been developed (see e.g., KUCKARTZ, GUNENBERG & DRESING, 2007; SILVER & LEWINS, [2]2014). None of those programs is specially adapted for Qualitative Content Analysis. This was because the separate processing steps involved in qualitative content analysis are difficult to implement in conventional and commercial programs. For example, it is not easy to keep the central content-analytical rules (category definitions, levels of abstraction, coding guidelines) constantly visible alongside the analysis. This only works to a limited extent using the memo function in MAXQDA, and that actually belongs to the grounded theory methodology. A table notation, central for summary and coding guidelines, can only be achieved partially. For these reasons, we have developed our own software program QCAmap (FENZL & MAYRING, 2017; MAYRING, 2014), which offers the following advantages:

- free use;
- interactively guiding users through the steps of content analysis;
- templates for the individual techniques such as summary, inductive category development, and deductive category application
- templates for the individual analysis units that should be defined as well as content analysis rules;
- ongoing maintenance and further development of the program as a web application;
- interactive possibilities for raters, also for intercoder comparisons;
- a manual (MAYRING, 2014) that can be downloaded free of charge.

The interactive nature of the program ensures that the essential steps of the qualitative content analysis are actually carried out. The program has been used in more than twenty

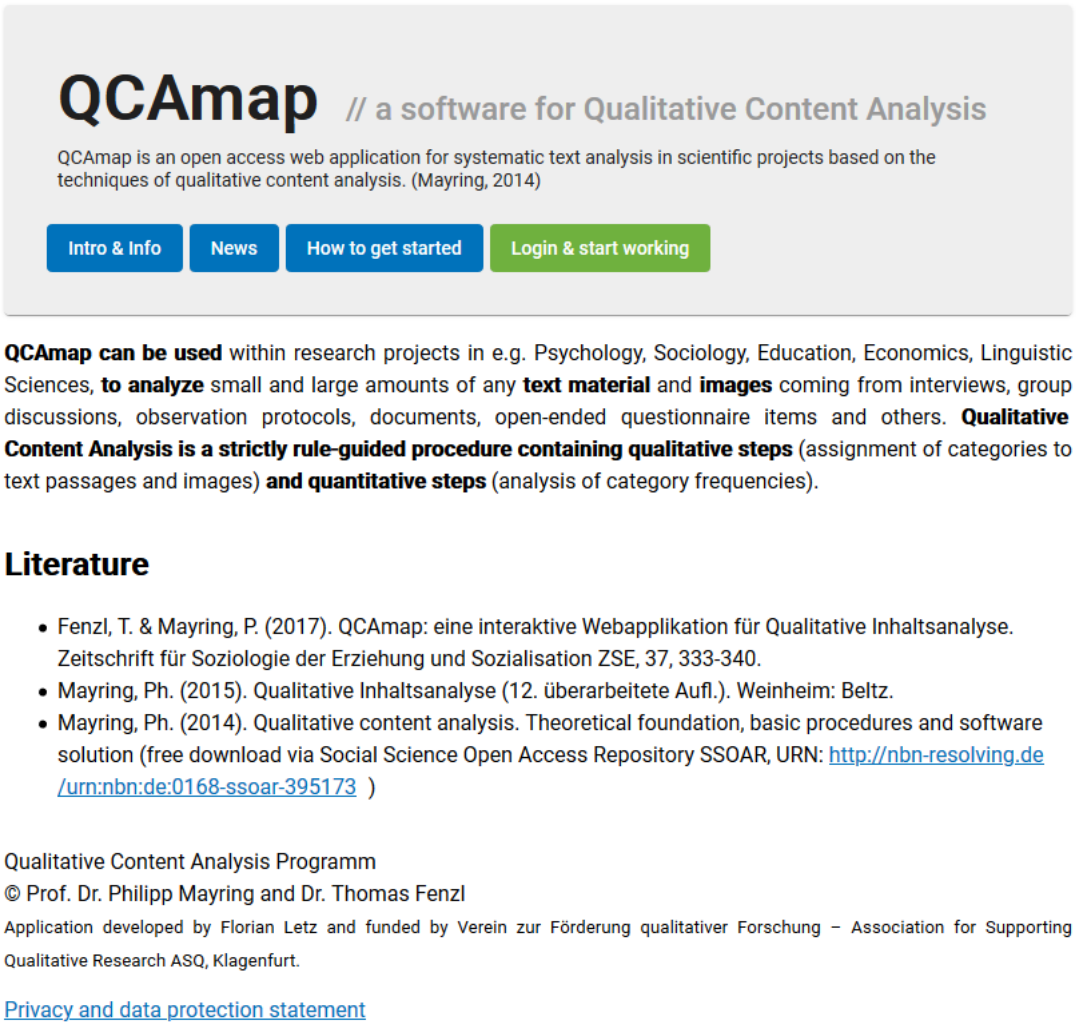
thousand projects since 2013. A new version with brand new features is planned for 2020. In future versions, we plan to integrate video analysis into the program.

Precisely because of its intermediate position between qualitative and quantitative analysis, qualitative content analysis appears to be an important methodical starting point. On the one hand, it is used for including measurements and numerical data from standardized interviews, questionnaires, observation systems, or test instruments in research projects. On the other hand, it also takes into account data from open surveys and observations in such a way that the resulting texts are analyzed as systematically as possible in an analysis that is guided by the research question as well as being both theory-based and rule-based in its approach.

For the references and for further remarks to the logic of qualitative Content Analysis, as well a discussion of the differences to other concepts of content analysis see the whole article, in English and in German, free access at:
www.qualitative-research.net/index.php/fqs/article/view/3343

QCAmap2020 – Step by Step

When searching for QCAmap you will immediately come across the link https://www.qcamap.org and its associated homepage:



The first blue button (Intro & Info) will take you to our homepage for Qualitative Content Analysis (www.qualitative-content-analysis.org). There, we strive to inform you (in English and German) about the following:

- Latest news,
- The Association for Supporting Qualitative Research ASQ (https://qualitative-content-analysis.org/en/non-profit-asq/), a non-profit organisation located in Klagenfurt (Austria) behind the know-how and development of QCAmap,
- Our software QCAmap, including the new features in the latest version QCAmap2020 (https://qualitative-content-analysis.org/en/software-2/qcamap-

2020-2/) and a series of slides (https://qualitative-content-analysis.org/wp-content/uploads/QCAmap2020Intro.pdf),
- Publications of our team on Qualitative Content Analysis,
- Possibilities to support our (free open access) work (sponsorship),
- The team (Philipp Mayring, Thomas Fenzl, Stella Lemke)
- Workshops on QCAmap (the annual Austrian summer workshop and external workshops organized by the team).

The blue news button leads you to the first bullet point in the list above.

“How to get started” provides a quick introduction about how to start the program.

The green button (“Login & start working”) takes you to the free open access program QCAmap.

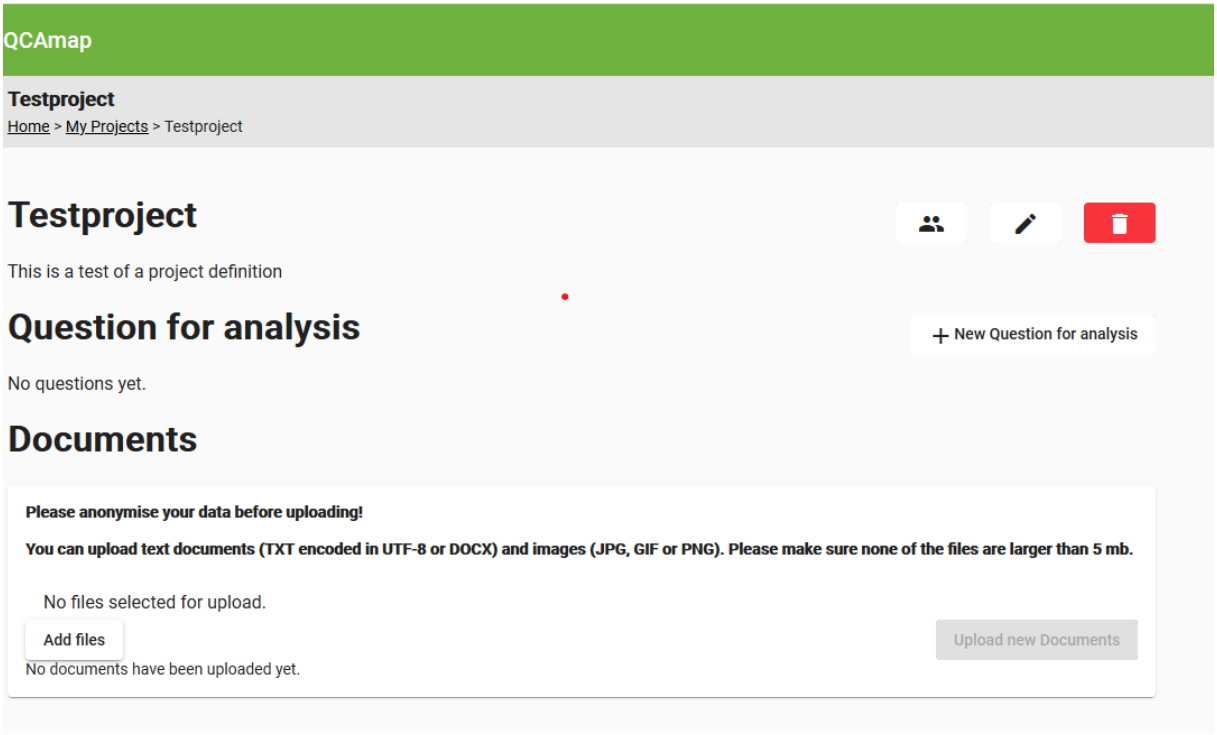
To create an account, the first step is to fill in a valid email address and a password. You will receive an email (please check your spam folder as well) that contains a link for you to confirm your email address. Once that is done, you have access (email and password) from anywhere that has the necessary hardware, browser, and internet connection.

This is the first screen:

You can define a project (“Create new project”) on this screen. A project is a series of documents (texts, pictures) and a main research question that maybe has several sub-questions for analysis. Perhaps you have been invited for an intercoder agreement test from another QCAmap user; this would be displayed in “Coder Agreements”. The green band, is shown on every page in the program, it gives you the possibility to check your profile, to read the data protection statement and to send an email to the support hotline (“Help”). Please do not forget to logout when you have finished your work.

Creating a project

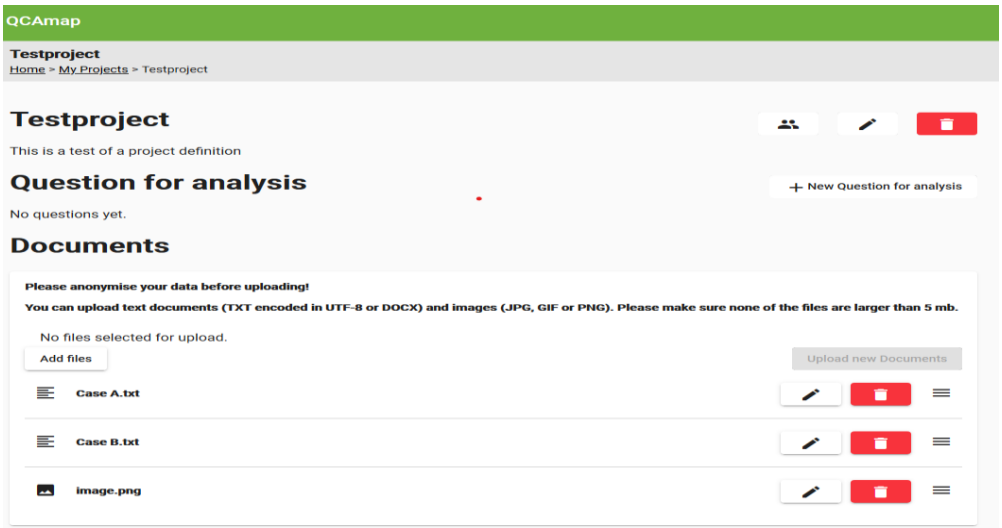
If you create a new project, a new screen will appear in which you can type a title for the project and then a description. You can change those settings afterwards. Once the projected is created, the following screen is shown:



The title (Testproject) and description (This is an example project definition) always appear when you open this project. The first line contains three symbols, and the first symbol allows you to share the project with other persons (they must have a QCAmap login, and then you can assign them editing or read-only rights). Project information can be edited (the pen symbol) and the project can be deleted (red bin symbol).

A question for analysis has not been defined yet; however, we can do it afterwards. If you do it straightaway the system would inform you that you have to upload documents before coding. Before uploading documents, please make sure that you have anonymised the data. That means inserting black stipes over people's eyes in pictures; or using abbreviations for personal information (e.g., names, places) in text. The format of text files is indicated by the file extension, for example, as txt or docx files, or JPG, GIF, and PNG files for pictures. Press "Add files", select the files, and press "Upload new Documents".

Below you can see that we have uploaded two texts and one picture:



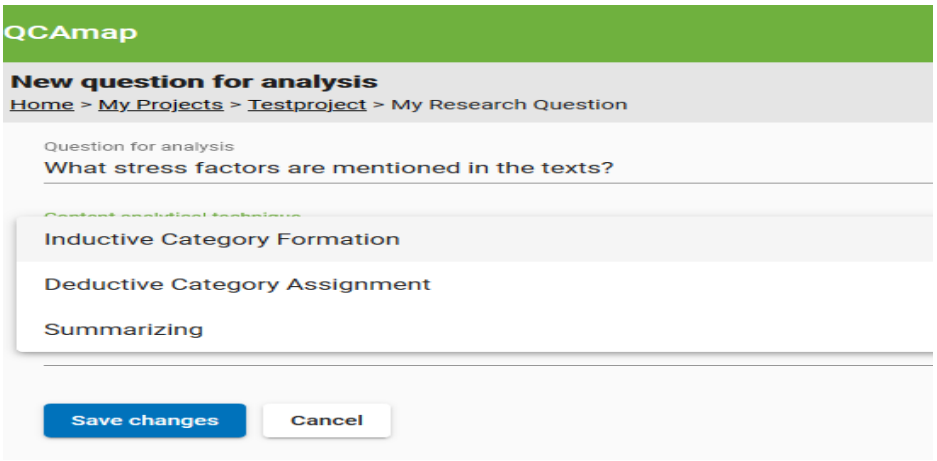
Doing text analysis by Qualitative Content Analysis, the definition of a research question is fundamental. We do not carry out with a free impressionistic interpretation of the material (which could be appropriate in explorative projects). Instead we will answer research questions connected to the theoretical background of the project. The elaborated formulation of the research question (as a sentence with an interrogation mark at the end!) is crucial because it leads you to the most appropriate content analytical technique:

- What characterizes the field? What is important in …?
  ⟶ Technique of summarizing
- What forms of … can be found? What do they say about …?
  ⟶ Inductive category formation
- Can I find … in the material?
  ⟶ Deductive category application

We can formulate a general research question and several sub-questions about the same material. In the program, those questions are denominated “Question for analysis”. After starting it, you have to decide on the most suitable technique. A pull-down menu gives you the three possibilities that have been implemented so far.

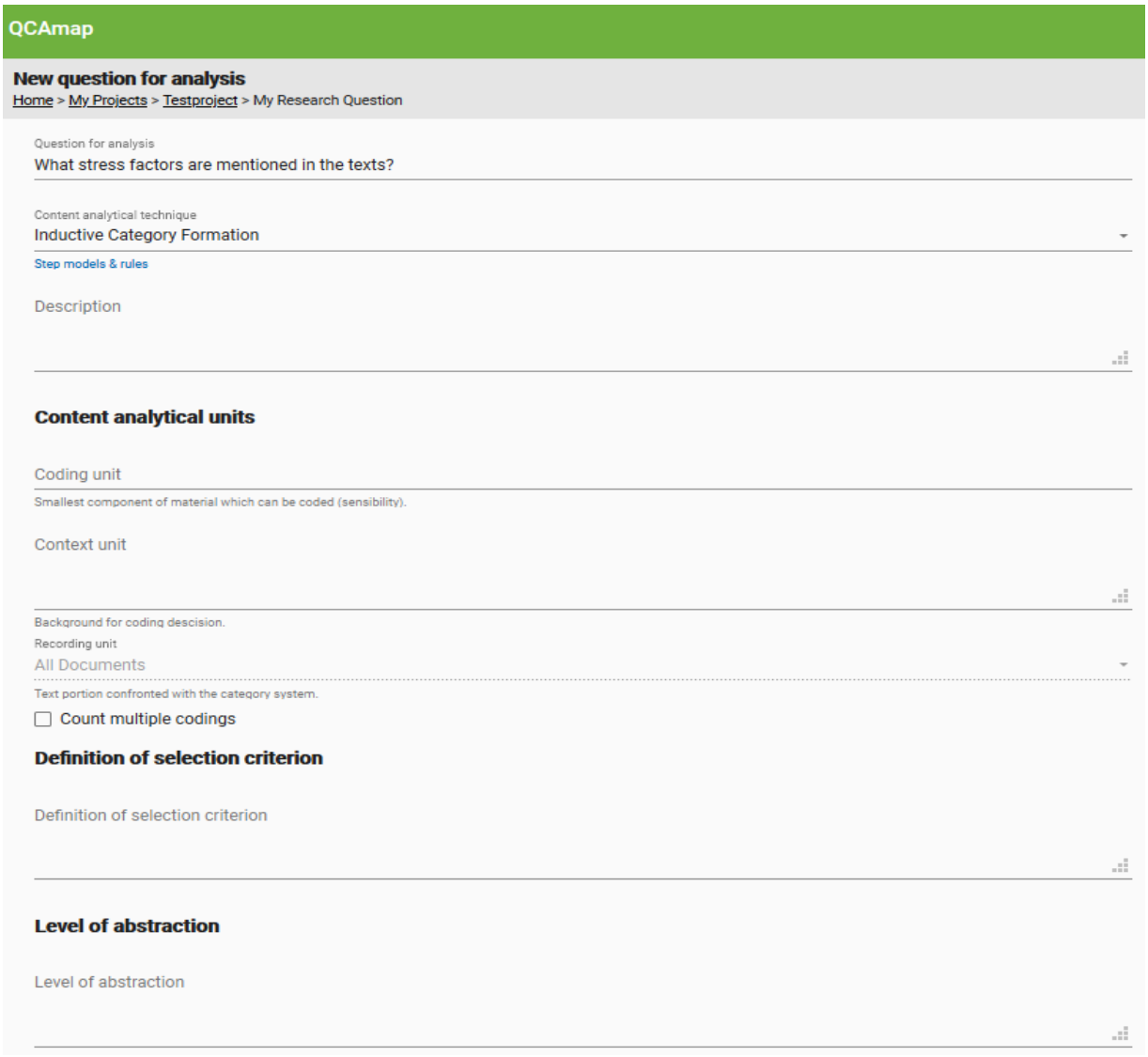
Inductive category development

Roughly speaking a large part of all questions for analysis within Qualitative Content Analysis are inductive category formations.



Those choices and inputs have to be saved manually. Afterwards (in coding mode), all actions are saved automatically. Of course, you can change it at the question for analysis level.

If you select one of the three techniques, the corresponding procedure and rules are displayed: for summarizing, the levels of paraphrasing, generalization, and reduction, for inductive category formation, the category definition and level of abstraction, and for deductive category application the categories and coding guideline. In our example, we have formulated the following question for analysis: “Which stress factors are mentioned in the texts?”, which clearly needs an inductive category formation.



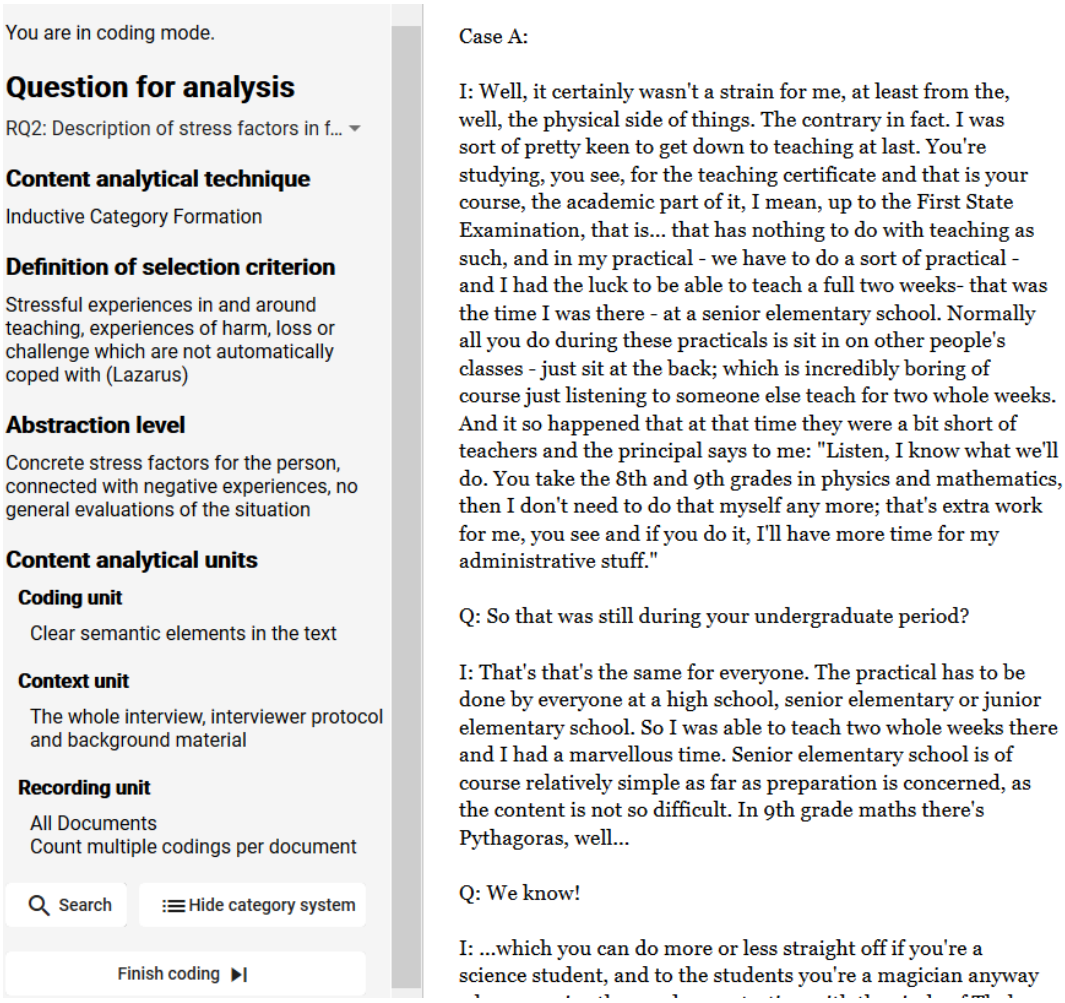
In all techniques, the content analytical units (explained on the screen) have to be specified. Furthermore, a step model and corresponding rules can be displayed (the respective link is written in blue) for all techniques. Click the link, and it will opena separate screen.

Three content analytical rules have to be defined (if you forget this, the system warns you and does not allow coding). In inductive category formation the recording unit is fixed as “all documents”. Recording unit means the amount of material that is to be confronted using the category system. New categories can be added from document to document during inductive category formation. The whole list of categories represents all documents at the end of coding, thus this has to be the recording unit (you cannot change it).

A decision has to be made with respect to counting multiple codings. At the end, a list of all categories is presented (as an Excel file) with the frequencies of occurrences of categories. If a category was coded several times within one document, it could make sense to count this, especially when the abstraction level is high: A person describes problems to pay the rent and difficulties to afford a holiday trip as stress factors. Both

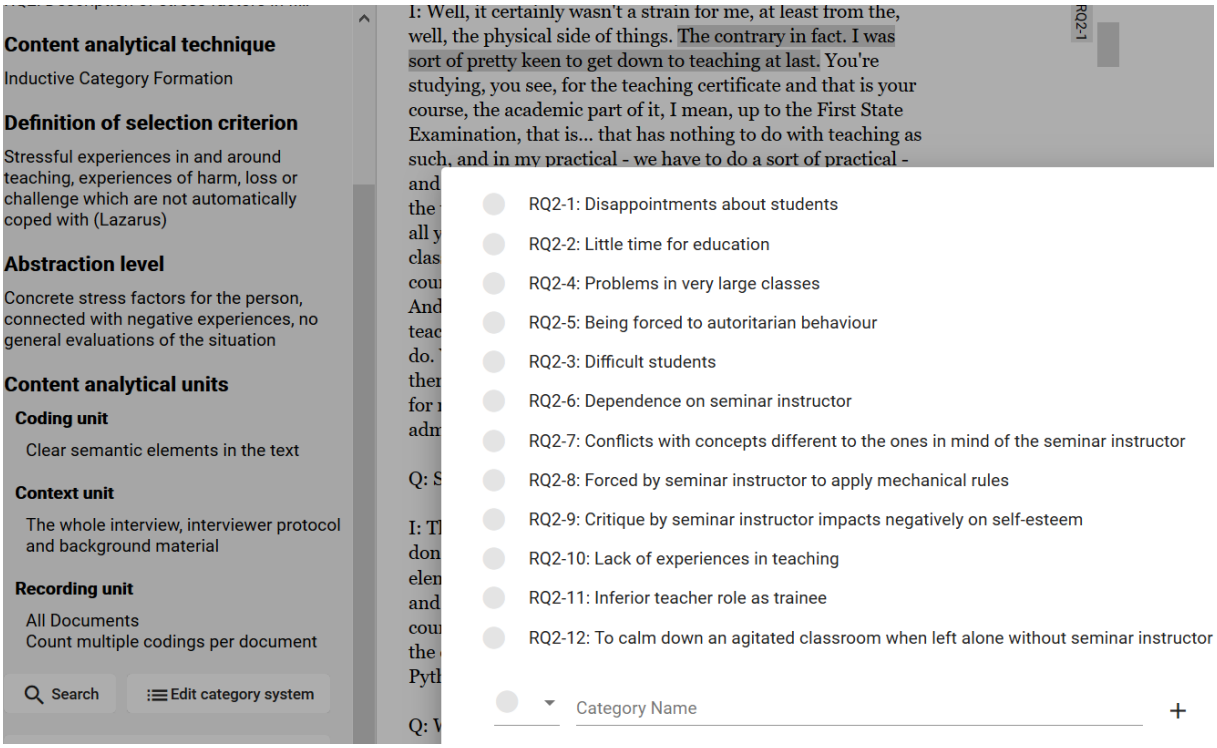
of them are coded as financial stress and therefore should be counted twice. If, however, you decided for a very concrete level of abstraction, and the same person repeatedly talks about his problems to pay his rent, this should be counted only one time. In our example, the “Count multiple codings” checkbox is not selected.

Please don’t forget to save your inputs and choices, once that is done, you can switch to the coding mode:

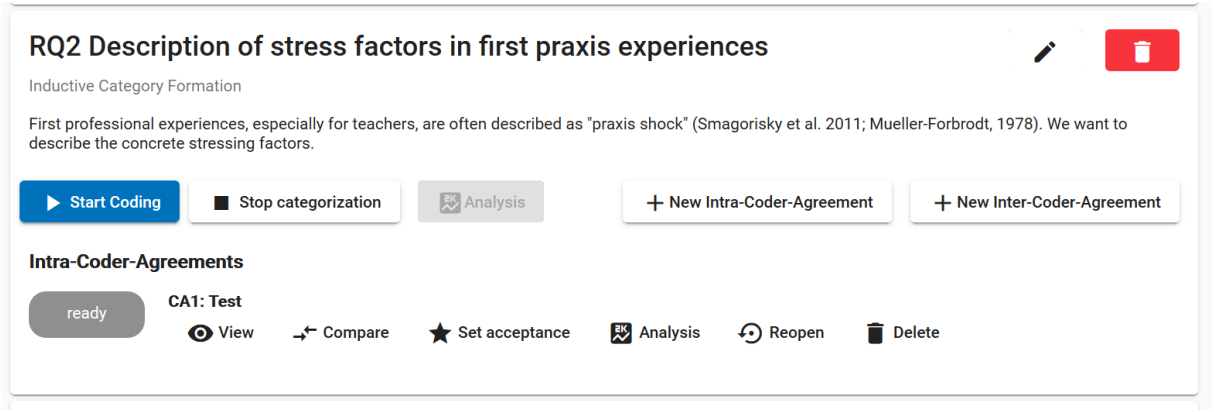


All definitions and selections are displayed on the left side to guide you through the coding process. A search function displays all text passages that were coded with a category. You can look at all categories developed up to this point on the right-hand side of the page or hide this. If you have worked through all documents, you can finish coding and enter the analysis tool. Nevertheless, you can reopen coding at any point if necessary.

Marking a text passage (in the example below, line two and three) opens a window containing all categories formulated up to now. You can subsume the text passage to one of those categories or formulate a new category (the last line in the category window). All categories are indexed with the number of the research question (RQ) and a continuous number.

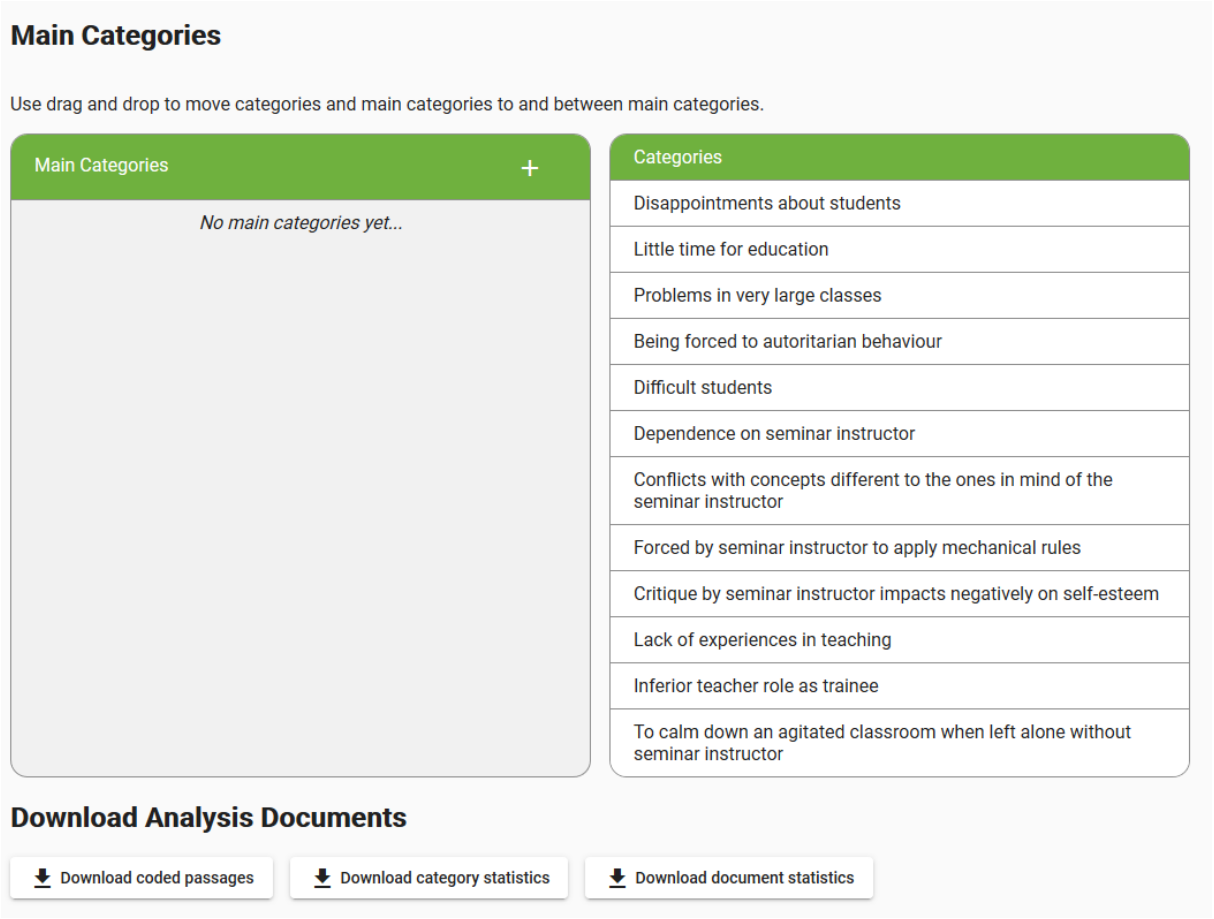


If you have finished coding you are presented with the following screen:



If you press “Stop categorization”, then the “Analysis” is activated. It’s possible to go back and do further categorizations (“Continue categorization”) in the material, and as you would expect different output files would be created, if any categorizations were changed. You can also check for coder agreements (I will describe this later).

The “Analysis” button takes you to the following screen:



First, you can formulate main categories, integrating the inductive categories to groups. This is a process of inductive (looking for similarities between the categories) or deductive theory bound argumentation. If you have defined main categories, the output files will contain frequencies for them as well.

Further analysis is done by creating three output files in Excel-format:

- “Download coded passages” creates a list of all categories and their corresponding text passages.
- “Download category statistics” creates a table of categories and their frequencies, including percentages of all categories and percentages of documents, in which the categories occurred:

A	B	C	D	E	F
Category ID	Category Name	Absolute Count	% of SUM	N of Documents	% of Documents
RQ2-1	Disappointments about students	1	0.05	1	0.25
RQ2-2	Little time for education	1	0.05	1	0.25
RQ2-4	Problems in very large classes	1	0.05	1	0.25
RQ2-5	Being forced to autoritarian behaviour	1	0.05	1	0.25
RQ2-3	Difficult students	2	0.11	1	0.25
RQ2-6	Dependence on seminar instructor	4	0.21	2	0.5
RQ2-7	Conflicts with concepts different to the ones in mind of the seminar in	1	0.05	1	0.25
RQ2-8	Forced by seminar instructor to apply mechanical rules	2	0.11	1	0.25
RQ2-9	Critique by seminar instructor impacts negatively on self-esteem	3	0.16	2	0.5
RQ2-10	Lack of experiences in teaching	1	0.05	1	0.25
RQ2-11	Inferior teacher role as trainee	1	0.05	1	0.25
RQ2-12	To calm down an agitated classroom when left alone without seminar	1	0.05	1	0.25
		19			

Naturally, within Excel you can edit the table. The category ID indicates the number of the research question (RQ2 in this case) and the category number associated with that question.

- “Download document statistics” creates an Excel-file, which can be further analysed with statistical procedures. It is a table of categories and occurrences; it contains four documents and eight inductive categories in our simple example:

Document	RQ2-1	RQ2-2	RQ2-4	RQ2-5	RQ2-3	RQ2-6	RQ2-7	RQ2-8
Case A.txt	1	1	0	0	0	0	0	0
Case B.txt	0	0	1	1	2	0	0	0
Case C.txt	0	0	0	0	0	3	1	2
Case D.txt	0	0	0	0	0	1	0	0

Deductive Category Assignment

You should choose deductive category assignment, if the theoretical background of the research question enables you to formulate categories before working with your texts. In addition, if you are interested in the occurrence and/or frequency of those categories within the texts. The program will guide you through different steps of analysis.

There are two forms of deductive category assignments: a nominal list of deductive categories or an ordinal category system. We will examine an example with an ordinal category system, and the following the research question: Has "practice shock" affected the self-confidence of the individual? (categories high self-confidence – middle self-confidence – low self-confidence).

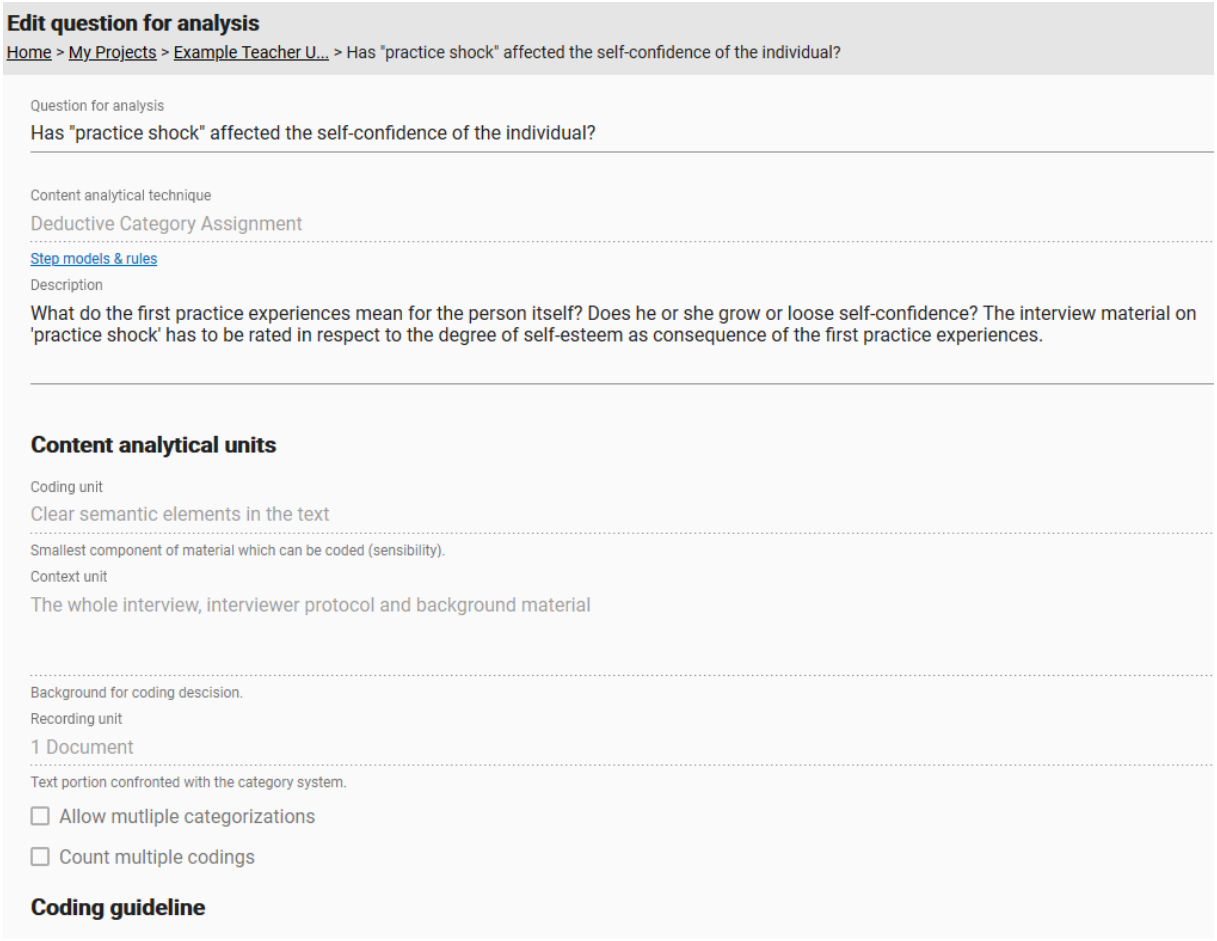
As before ,the blue link (picture next page) would take you to information about the specific step model and the rules for analysis in a separate window.

You have to define the content analytical units. They might be different from the ones given above for the inductive question for analysis.

In deductive category assignment the recording unit is predefined as “document” because every document is coded separately with the common deductive categories.

Usually, every document is coded by one of the categories. For working with ordinal categories, this is essential. I want to know if the document (the interview partner) shows high, middle, **or** low self-confidence. I would not “Allow multiple categorizations” because at the end of the analysis I want to know how many

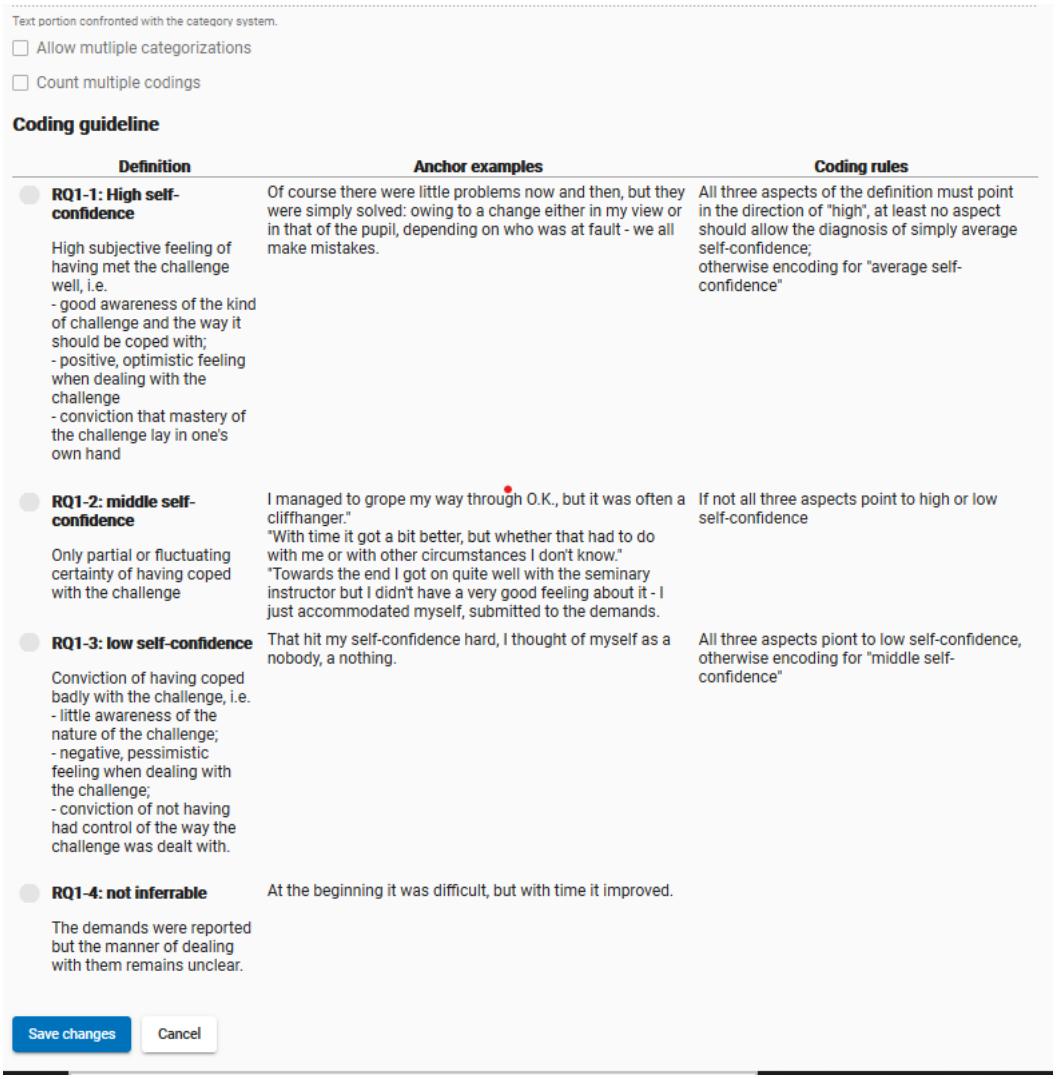
persons (documents) show high, middle, or low self-confidence (for this reason the respective checkbox is not selected at the bottom of the picture below).



If the documents (e.g., interviews) are very extensive, covering different areas, it might be interesting when a given category occurs several times in one document. In that case, you should click “Count multiple codings”. Again, for ordinal category systems it is not recommended because it is an act of interpretation, taking all relevant phrases together, to decide for one of the ordinal categories, and typically not a case for counting. For the same reasons, we do not recommend coding all text passages relevant to the ordinal variable on one of the categories, and then counting which category occurs most often within one document.

The next step would be to write a coding guideline (click on “Coding guideline – add new category”). This is the central instrument in deductive category assignment. Please formulate an explicit definition for each deductive category based on the theoretical background of the project. Within the pilot phase you may find a good example for one of the categories. If so, you can put it in coding guideline as “anchor example”. If you come to a text passage for which you are not sure about the right category, then it would be necessary to decide based on the theoretical background, formulating an additional coding rule for further codings. In doing so, the coding

guideline will be adapted as you go along, as displayed in the example. Please do not forget to save changes in the coding guideline. After the pilot phase, the coding guideline cannot be changed any longer; it is crucial that you code all material with the same coding guideline.



Text portion confronted with the category system.

☐ Allow mutliple categorizations

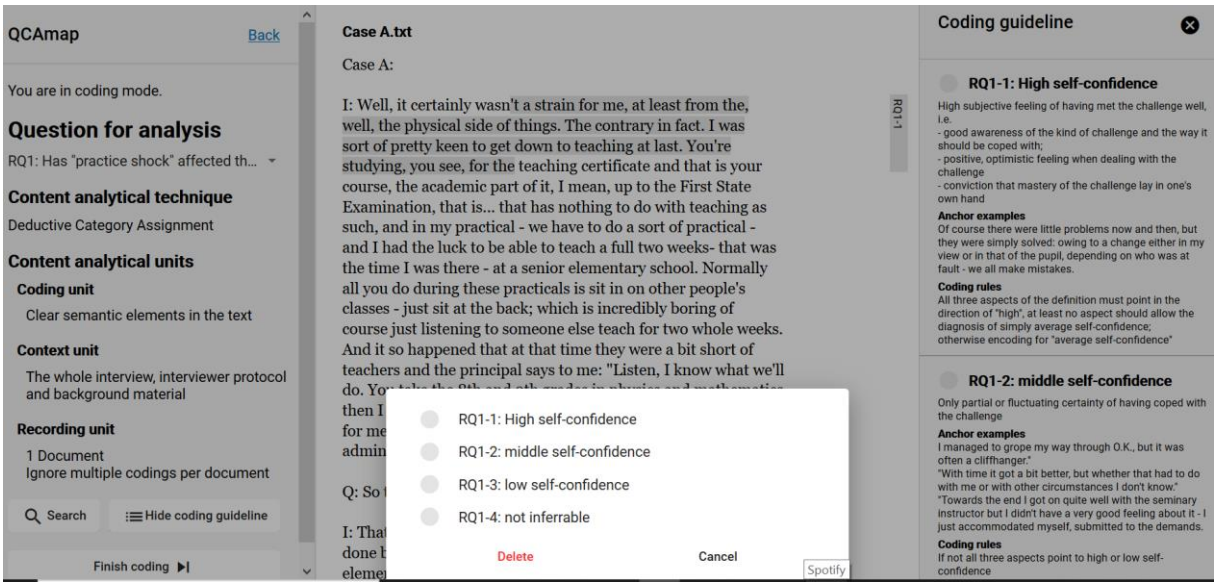
☐ Count multiple codings

Coding guideline

Definition	Anchor examples	Coding rules
RQ1-1: High self-confidence High subjective feeling of having met the challenge well, i.e. - good awareness of the kind of challenge and the way it should be coped with; - positive, optimistic feeling when dealing with the challenge - conviction that mastery of the challenge lay in one's own hand	Of course there were little problems now and then, but they were simply solved: owing to a change either in my view or in that of the pupil, depending on who was at fault - we all make mistakes.	All three aspects of the definition must point in the direction of "high", at least no aspect should allow the diagnosis of simply average self-confidence; otherwise encoding for "average self-confidence"
RQ1-2: middle self-confidence Only partial or fluctuating certainty of having coped with the challenge	I managed to grope my way through O.K., but it was often a cliffhanger." "With time it got a bit better, but whether that had to do with me or with other circumstances I don't know." "Towards the end I got on quite well with the seminary instructor but I didn't have a very good feeling about it - I just accommodated myself, submitted to the demands.	If not all three aspects point to high or low self-confidence
RQ1-3: low self-confidence Conviction of having coped badly with the challenge, i.e. - little awareness of the nature of the challenge; - negative, pessimistic feeling when dealing with the challenge; - conviction of not having had control of the way the challenge was dealt with.	That hit my self-confidence hard, I thought of myself as a nobody, a nothing.	All three aspects piont to low self-confidence, otherwise encoding for "middle self-confidence"
RQ1-4: not inferrable The demands were reported but the manner of dealing with them remains unclear.	At the beginning it was difficult, but with time it improved.	

Save changes | Cancel

Now you can code the material:



When marking the material, you can, of course, only assign one of the preformulated categories.

Having coded all documents, please press “Finish coding” (left-hand side at the bottom) and go to “Analysis”. QCAmap creates three Excel-files for you as results: the coded passages together with the code, the coding guideline, and most important for deductive category assignment, a table with documents and categories:

	A	B	C	D	E
1	Document	RQ1-1	RQ1-2	RQ1-3	RQ1-4
2	Case A.txt	1	0	0	0
3	Case B.txt	1	0	0	0
4	Case C.txt	0	0	1	0
5	Case D.txt	0	1	0	0

This table can be further analysed using statistical procedures.

Summarizing

Summarizing is a content analytical technique used for

- Smaller amounts of material because it is very extensive,
- Diffuse material, in which it’s difficult to ascertain the central points,
- Explorative research questions, where categories cannot be formulated in advance.

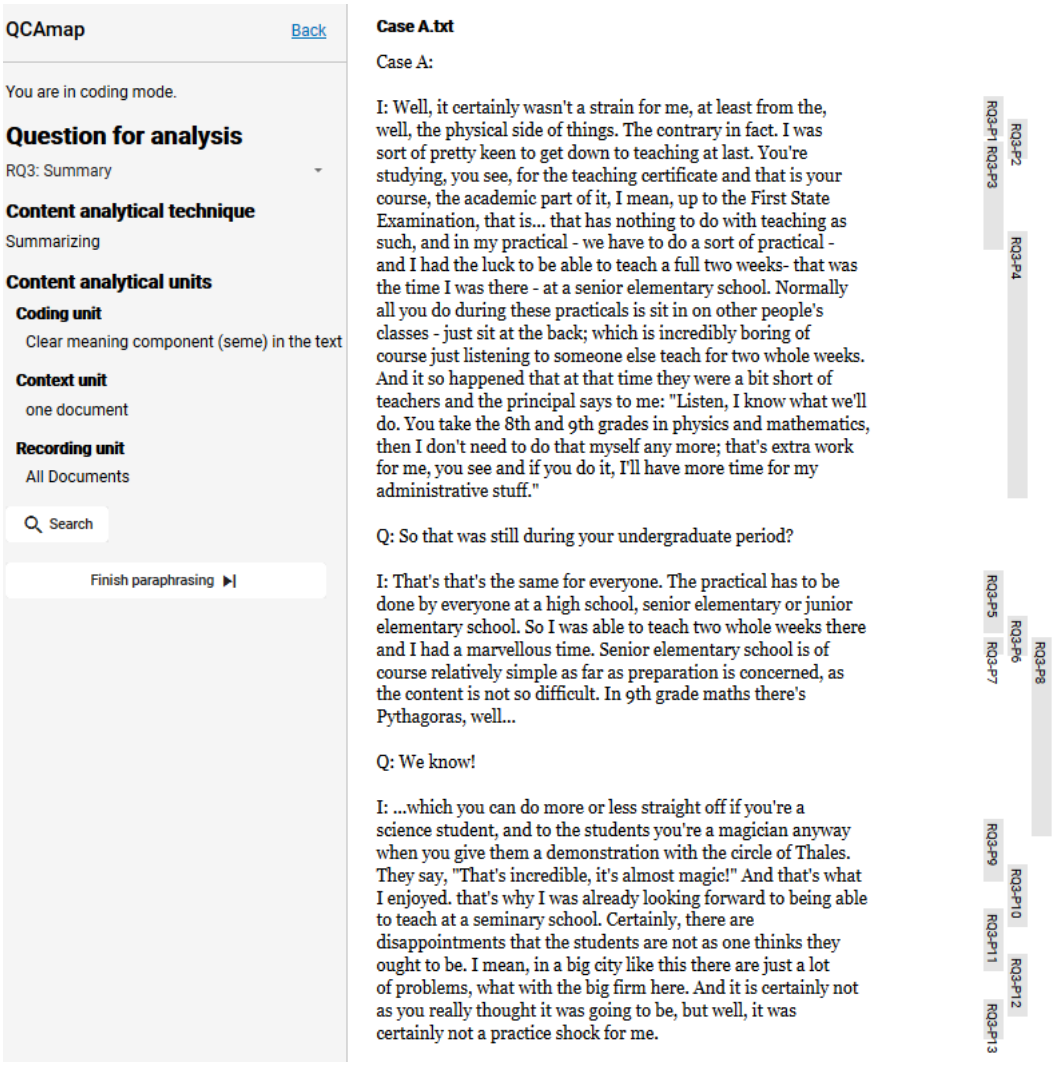
Summarizing content analytical procedures convert the whole text into single propositions and generalizes those propositions step by step to its core categories.

The result is a table with all those steps, almost as extensive as the original text. Inductive category formation is based on this procedure, but much more economical:

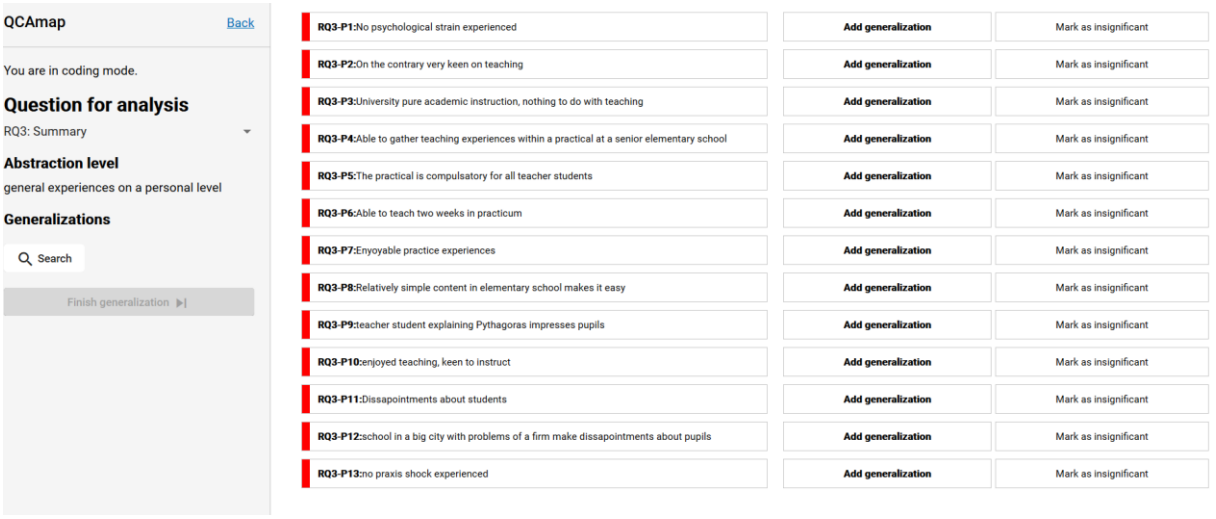
- There is no paraphrasing.
- Only text passages relevant to the research question were taken into consideration.
- The categories are directly formulated on a predefined level of abstraction and not generalized step by step.

Hence, think carefully about whether inductive category formation might be more appropriate before deciding for summarizing.

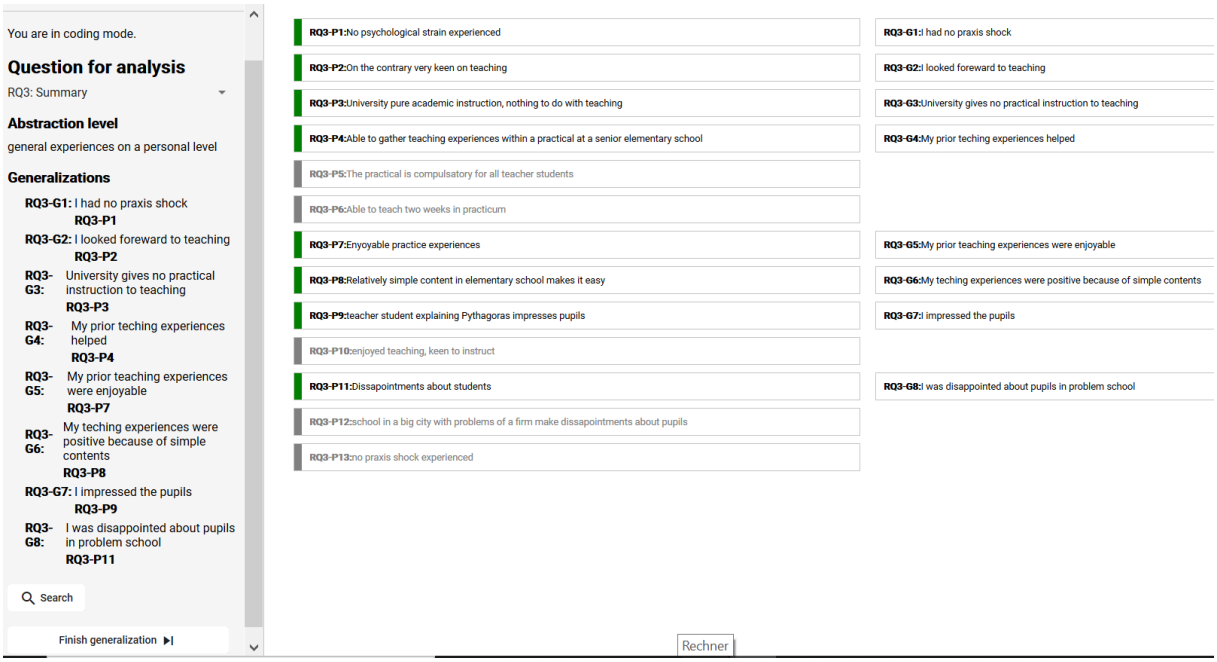
The procedure starts with the formulation of propositions. Every sentence has to be reduced to its content.



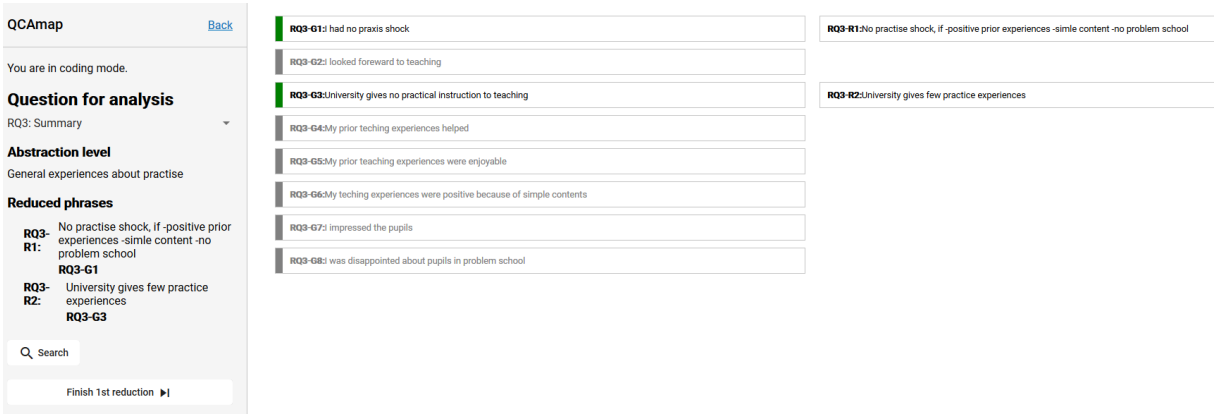
“Finish paraphrasing” at the left bottom gives you the possibility to work with those paraphrases. Immediately you are asked to define the level of abstraction for this step. Afterwards you can generalize the paraphrases or leave them out if they are duplicated or included in other generalized paraphrases.



The next step would be to summarize further. Press “Stop generalization”.



Every paraphrase has to be generalized or left out. At the end, the “Finish generalization” button takes you to the next step. It gives you the possibility to reduce the generalized paraphrases.

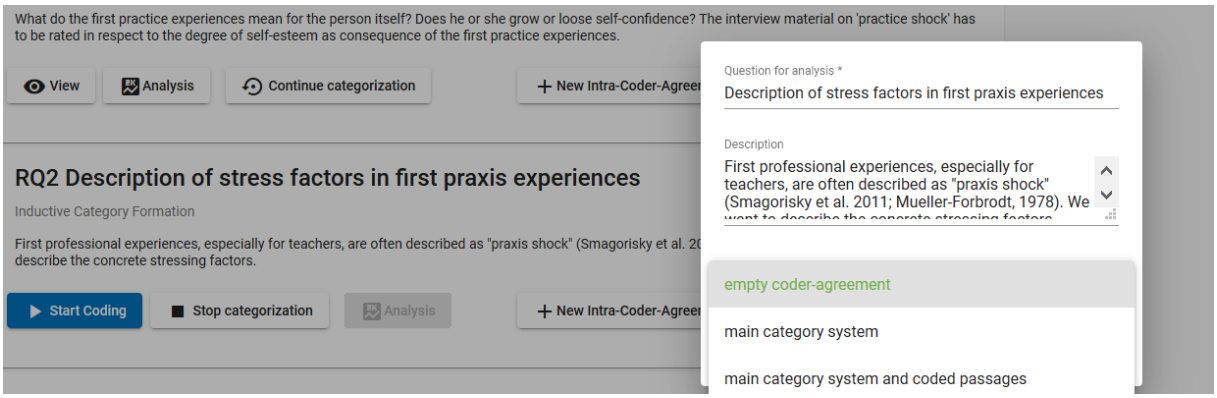


“Finish 1st reduction” leads either to the next round of generalizations and reductions or it stops the summarizing process. If we would stop here, we would have developed two summarizing categories from one page of the transcript. The two categories are on a different level, which is typical for summarizing because we did not have a category definition. The summary addresses all material, and this only makes sense for very explorative research questions. The output files contain the paraphrases, or the whole process of generalization and reduction.

Coder Agreements

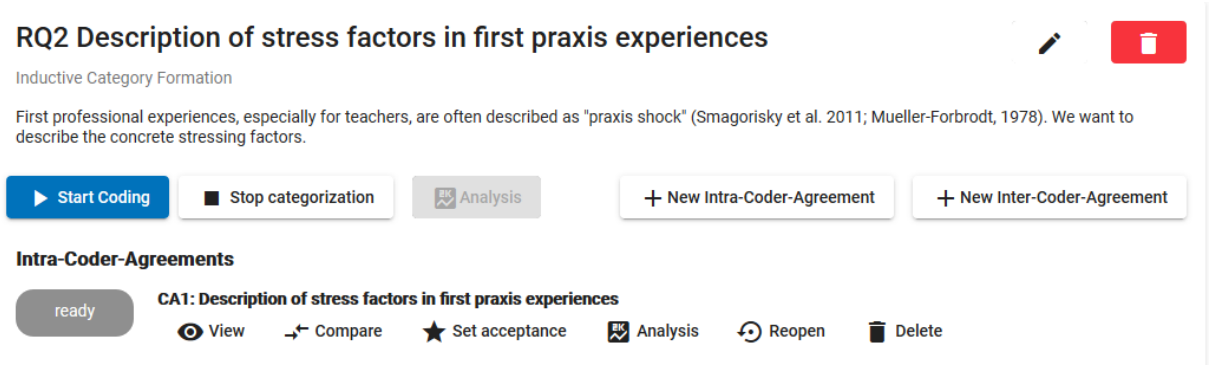
QCAmap offers the possibility to rate the intra-coder agreement as a measure for reliability and the inter-coder agreement as a measurement for objectivity as central quality criteria in qualitative content analysis.

On the project page, for every question for analysis the two checks can be opened. For inter-coder agreement tests, please fill in the email address of the second coder (who has to be registered within QCAmap). In both cases, a pull-down menu offers three possibilities within inductive category formation.



- The strongest test would be to give myself (intra-coder) or the other person (inter-coder) only the documents and research questions. She or he has to formulate the categories and assign them to the material.
- A medium test consists in a remake of the codings. The main category system is given to the second coder.
- The weakest test shares all categories and codings with the second coder. This is a form of supervision.

For deductive categories, only the second test makes sense. The second coder is given the (preformulated) categories and tries to assign them to the texts. For summarizing, no agreement checks are offered.



In this example, the intra-coder agreement with shared categories was finished, and by pressing the Compare link, I can see the test results – with original codes on the right side and second codes on the left side (thus in the excerpt below there is no agreement for the two categories shown):

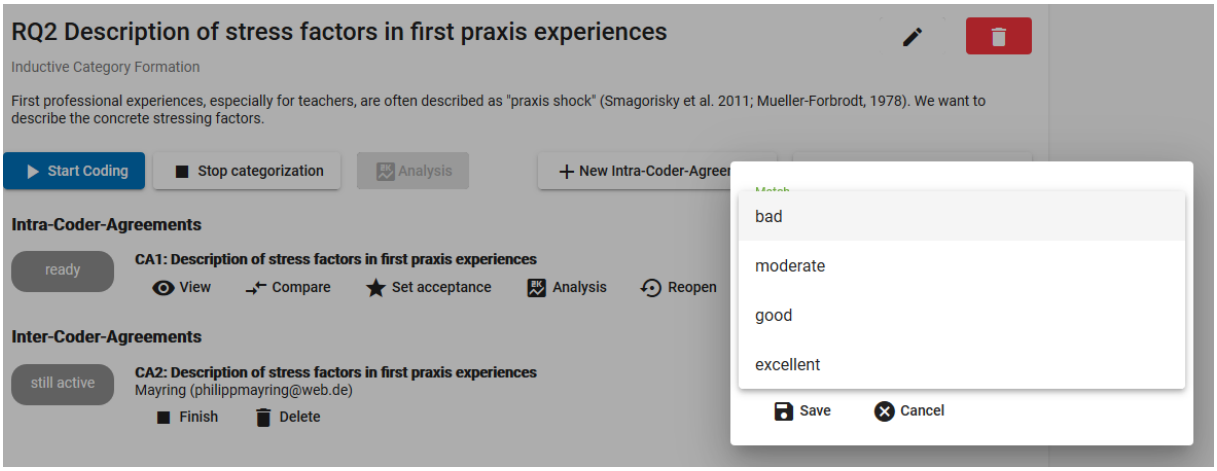
Q: So that was still during your undergraduate period?

I: That's that's the same for everyone. The practical has to be done by everyone at a high school, senior elementary or junior elementary school. So I was able to teach two whole weeks there and I had a marvellous time. Senior elementary school is of course relatively simple as far as preparation is concerned, as the content is not so difficult. In 9th grade maths there's Pythagoras, well... RQ1-11

Q: We know!

I: ...which you can do more or less straight off if you're a science student, and to the students you're a magician anyway when you give them a demonstration with the circle of Thales. They say, "That's incredible, it's almost magic!" And that's what I enjoyed. that's why I was already looking forward to being able to teach at a seminary school. Certainly, there are disappointments that the students are not as one thinks they ought to be. I mean, in a big city like this there are just a lot of problems, what with the big firm here. And it is certainly not as you really thought it was going to be, but well, it was certainly not a practice shock for me. RQ1-1

Only for deductive category assignment does a statistical inter-rater agreement calculation (e.g., Cohens Kappa) makes sense. In inductive category formation the degree of agreement has to be rated by the first project author. By choosing the button “set acceptance”, she or he can assign an agreement of bad, moderate, good or excellent.



That’s it; enjoy the program!