Request for Proposals

University of Amsterdam MOOR tool – Crowded Theory

Text with green background are optional things, which are probably not necessary for this prototype.

Introduction and history

In this document we shed light on the background of the UvA MOOR project, give a methodological account, and express the technical specifications that will serve as a basis for making the online tool.

The project started with the observation by Christian Bröer that the interpretation of qualitative data in the meetings of his programme group enabled great leaps forward, but remained restricted in time and place. This collaboration might be extended and enhanced digitally. In January 2014, Christian gave a small course on this subject, in which students were instructed to evaluate existing software that could be put to this use, and to think about what alternative new software could look like. This resulted in a prototype of a new tool that was tested with a group of social scientists during a conference in August 2014. On the basis of these lessons, Christian worked together with several others to continue thinking about a next version of the tool, and that's where we are now.

In the current phase of the research we want to offer a course on the everyday expressions of depression, and the medicalization of sadness. Within the frame of this course we ask students to make two observations of depression and/or sadness, and to subsequently analyse these in a collaborative manner within the online tool. This analysis takes place in several steps, and will eventually result in a collectively written conclusion in which all findings are

presented. Even though the students are instructed to use the tool from Coursera, the tool we envision should be useable apart from Coursera.

The wishes we have for the tool are based on extensive discussions on the underlying methodology. In the next section we will further elaborate on this.

Academic-methodological background

The type of research that we try to facilitate within the tool is qualitative research. That means that the data (in our case text) are interpreted and analysed without using numbers. The interpretation, however, is informed by an established methodology. Because we are aiming to work together on a scale and in a manner that has not been done before, we also need a new methodology for this. Below are some of the central points of our methodology, copied from an article that we are in the process of publishing.

Perspectivism

"*Crowded Theory is based on insights from the study of scientific practice that suggest that knowledge is produced in interactions between humans, technologies and objects being studied (Stahl et al., 2006). Crowded Theory deliberately designs an "interpretive zone" (Wasser & Bresler, 1996), where "multiple viewpoints are held in dynamic tension, as a group seeks to make sense of fieldwork issues and meanings" (p.6). This makes sense because a larger number of participants is better able to draw out the range of potential interpretations of a social phenomenon, including the interpretations of non-academics (Sweeney et al., 2013). Cornish and colleagues therefore suggest basing collaborative research on perspectivism (Cornish et al., 2013). However, perspectivism comes with a challenge: imagine a team of 100 members looking at the same material. Do we end up with 100 or more interpretations? How do we achieve meaningful reductions?*

We do not think that perspectivism presents greater threats for collaborative research than individual research endeavours. From our research experience, we conclude that the number of different interpretations is often limited, and researchers can manage interpretive diversity. The number of perspectives does not increase with the number of participants because perspectives can also be group properties. Whether perspectives emerge at the individual or group level depends on the research practice. Enlarging and improving collaboration could thus arrive at a more limited number of valid interpretations compared to a range of single researcher projects. More precisely:

Criteria and good practices to assure validity are still relevant and need to be built into the collaboration. Acknowledging that valid claims need reflection and disagreement, Crowded Theory shifts attention to the analysis phase and aims to enhance validity through joint interpretation. Bringing in different perspectives may increase the validity of findings, as this adds information and reduces errors. Working collaboratively can enhance the credibility of the results through what Guba and Lincoln (1989) have termed 'progressive subjectivity': constructs developed through collaborative interpretation are continuously, collectively (instead of individually) reflected upon.

In Crowded Theory, participants are put in a position to agree or to agree to disagree. The ensuing deliberation fosters congruence: agreement on a set of

alternative competing interpretations (Sweeney et al. 2013; Fischer 1998). "Mutual adjustment" (Hall et al., 2005) reduces complexity and delivers a more limited number of valid interpretations compared to multiple single researcher projects.

Common Sense

While coding is an activity that requires some training, interpreting and discussing are closer to everyday competences and are thus better suited for crowd-based research. The potential involvement of non-academics allows for transdisciplinary research. However, there is another obstacle here: can untrained participants go beyond common sense? Is it possible for team members to bracket their presuppositions and move from "first" to "second order constructs" (Schütz & Luckmann, 1973)? Again, our reply is optimistic. Practically, as teachers and team researchers, we have observed that non-academics can be introduced to social science concepts and quickly learn to apply them productively. It is possible to establish categories for analysis in dialogue. Non-academics, moreover, are not cultural dopes and are already assessing the workings of daily social life. The example of wikipedia shows that a crowd can produce knowledge of the highest quality.

Moreover, collaboration supports learning and reflexivity (Barry et al., 1999). Discussing interpretations can help with the identification of common sense: academic and non-academic participants can ask for explications and question the assumptions of other participants. Thus, they are able to move from first order to second order constructs.

Crowd Wisdom

Finally, we want to position Crowded Theory and collaborative research in relation to collective intelligence and "wisdom of the crowd". They are similar, but "collective" stresses that intelligence is a property of the group, while "collaborative" puts action up front.

The classic wisdom of the crowd examples are numerical estimates by large groups (height of building, weight of a cow), which, after averaging, are more precise than individual estimates and are almost as accurate as measurements (Galton, 1907). Wisdom of the crowd effects are based on the fact that peoples' answers are diverse but are normally distributed around a correct answer. Interpretive research does not presuppose one correct answer or a normal distribution of interpretation. "Averaging" interpretations therefore does not lead to enhanced results, although frequencies of shared interpretations can be useful, for example for an analysis of discursive dominance.

Research into Wisdom of the Crowd seems to suggest that one needs to avoid groupthink or tunnel vision (Lorenz et al., 2011), although there is some discussion about this (Farrell, 2011). Agreement between group members does not necessarily imply correct answers. Even stronger, wisdom of the crowd and collective intelligence are debated for looming collectivism (Tapscott & Williams, 2008) or "digital maoism" (Lanier, 2006). This reminds us that collaboration must not be equated with agreement or majority rule. Instead, "minority reports" are essential for establishing a – most likely limited – number of interpretations."

As mentioned above, an important point is trying to avoid echo chambers, in which later interpretations (or interpretations of participants who get involved in the research later) are being influenced too much by earlier interpretations/participants. As described it is of great importance for our methodology that different points of view can exist next to each other and these can be analysed in interaction and comparison with each other. However, when this is taken to its extreme every participant would only be working for itself, to minimize influence of other participants. This creates problems both by the gigantic (unworkable) amount of interpretations that is created as a result, as well as by the fact that the collaborative spirit of the methodology is violated. It is therefore central to maintain a balance between these extremes in the tool: no echo chambers, but also no solitary confinement.

Also previously mentioned is that it's not about averaging the interpretations. Not only do we want to encourage differences in point of view and approach at the start to facilitate diversity (which enriches discussion and findings), but we want to allow this difference in opinion to persist, up to and including the conclusion. It should really not be a big surprise that based on a large amount of data, different conclusions can be drawn, and it is important to recognize that those can be valuable beside each other, even if they are (partly) contradictory, caused by e.g. selected data and the theoretical lens used. For our tool, it is thus important that although we want to avoid an explosion of interpretations that are all equally valid ('That's just, like, your opinion'), we want to allow properly substantiated dissent, and then make explicit in our conclusion. Up to the conclusion, it should be possible to agree to disagree.

Finally it is useful to give a methodological background of the last choice of the technical specification that follows. In the interpretation-step we chose to have participants interpret a small part of the data first, before we open up all observations for interpretation. This is related to the fear that the number of unique interpretations would climb too rapidly otherwise. After this step it is possible to reduce this number by combining existing interpretations, but we expect that this is too much work (in terms of both the time required, and the number of interpretations that need to be considered concurrently to make this process meaningful). Instead we let participants crystallize their interpretations and compose a sort of shared collection of (contradictory) interpretations, that cover the meaning of a limited number of observations sufficiently, and to which not a lot of new interpretations are added with new observations. This point if called saturation in qualitative social science. From this point on it should be easier to open up all the data, and apply the 'established' interpretations, without causing an explosion of new unique interpretations.

Technical specifications

In this section the technical demands for the tool are specified. We try to leave as much room as possible for the (graphical) implementation of this.

The different central 'units' of data in the tool are the participants (accounts), observations, interpretations, groupings (of interpretations), and a conclusion. The participants will consist of a username and an email address, with an optional short self description. The observations (500 words) and

interpretations (50 words) are pieces of plain text. The conclusion is a text document with formatting of around 7500 words. We expect several hundred participants, who all contribute two observations. We expect a maximum of 100 interpretations, and a maximum of 50 groupings.

As will be described below, the interpretations are linked to observations, the groupings to the interpretations, and the conclusion to the groupings. The participants are linked to all above documents (they are the ones making these links). It should at all times be possible from 'higher' levels of data to refer and step back to the underlying sources by means of hyperlinks.

In the next subsections the technical specifications will be explained in 4 steps that chronologically illustrate the process of the participants through the research.

Step 1 – uploading and describing

In the first part of the tool the participants are able to upload observations, and to add a description to these.

Uploading should be possible by typing out (or pasting) an observation in an empty text field. It should also be possible to upload a file from a local drive, in several standard file formats (docx, doc, txt, odf, rtf). There is a maximum size for the observations of ## kb/words. The observations and descriptions are automatically checked to see if they are empty or contain cuss words. In that case the administrators/moderators are notified so a manual check can take place. In addition to text, photos and other images can also be uploaded.

Subsequently, a description can be added to the observation, in a series of open text fields. In every field a different part of the description is asked for. It is mandatory for participants to add a description.

Administrators and moderators get a notification of new observations, and can directly view both observation and description, so that checking the contents is possible, and undesired observations and/or descriptions can be removed. It should be possible for administrators to add text fields and instructions for the description, in case one of these was missed before the research started.

Step 2 – interpreting

In this step the participants write interpretations in the tool, and connect these to observations. Participants also evaluate interpretations of other participants, and are presented with the opportunity to integrate their own interpretation into another, or make a suggestion for altering another's interpretation.

From the perspective of the observation: observation a (O_a) is assigned to participant 1 (P_1) and he/she gets the instruction to read this carefully. Then a question follows with instructions for interpreting the observation, after which P_1 can write interpretation α ($I_α$) in an open text field. Now O_a goes to P_2 with the same instruction to read carefully. Next to the question for interpretation, P_2 now also sees $I_α$ and the following three options: whether his/her reading corresponds to the existing interpretation, and wants to join P_1 in this interpretation (Agree), or the reading, with some small modifications of $I_α$, could be merged with the existing interpretation (Modify), or whether the interpretations are so different that a new interpretation $I_β$ should be added to

O_a (Add). P_1 gets a notification of this in all cases, and in the case of Modify, P_2 also writes a proposal for the merged interpretation, which is sent to P_1.

This process repeats, where every next participant is always first asked to read the observation carefully on their own accord, and is then presented with all existing interpretations of O_a (in random order), with the Agree/Modify/Add instruction. When the participant chooses Modify, he/she will be presented with all the observations that interpretation is currently being applied to, and the instruction that the altered interpretation should ideally still be applicable to all those observations. If that seems impossible it is still possible to write a new interpretation (Add). All participants that have contributed to an interpretation (Agree and Modify) get a notification if a new participant makes a contribution. The original observer O_a gets notifications about interpretations from the start, and is thereby invited to offer extra contextual information when this is necessary. That's possible by editing the description of the observation. For O_a this continues until 10 new participants do not find new interpretations.

From the perspective of the participant: P_1 is shown O_b when he/she is done with O_a. The first step is again to carefully read, but now D_1 already has interpretations on his/her mind. All the existing interpretations (that is to say, of all observations) are offered with Agree/Modify/Add as instruction. Agree in this case means that the (existing) interpretation will be added to the observation. It should be clear that there is a list of interpretations per observation, and a list of interpretations across all observations (both of which steadily grow, and the second of which is offered to fill the window of interpretations per observation). When sufficiently many new observations do not result in new interpretations, this step can be closed down. It is up to the administrators to decide this moment. We can also conduct an extra check in this stage by adding 'deviant cases/outliers' in the form of observations, or by activating possible differently-interpreting participants. This is the work of the moderators. After the check, it is possible to 'lock down' the interpretations.

As the final stage of the interpretations, all observations are opened up to interpretation. In principle this works the same as before, only we expect that this will result in (much) less new interpretations. This stage is mostly the application of existing interpretations.

Step 3 – correlating

In this step the participants make groupings based on predefined (cor)relation questions. This can take the shape of a relation between descriptors (in the descriptions of the observations) and interpretations, and between interpretations themselves. The participants see a list of all interpretations, and they can drag these into an empty field. There they can group the interpretations as they see fit. It is not mandatory to use all interpretations in the network. Next to this network, it is possible to write a description of it. From the description it is possible to refer to interpretations, and also to groups of interpretations (selected by e.g. a rectangle select tool).

Other participants can then use the same three options as before (Agree/Modify/Add) to go through the existing networks. When they choose Modify, they can move interpretations around in the existing network, as well as add or delete them. They can also give a description of their edits, which forms a discussion thread below the existing description(s). In this step all participants

that contributed to a network (Agree, Modify) also get a notification when another contribution is made. It is important that the references to (groups of) interpretations remain, even when the composition of the groups changes. That means that the collection of interpretations that was meant by the reference in the first place should remain, even if the interpretations are organized in a different way in the current network. Finally it should be possible to browse through the history of the groupings in a network, and it should be possible for administrators to revert any edits made.

As an aid to identify relations between descriptors and interpretations, and interpretations themselves a matrix can be computed that shows these juxtaposed with each other, with the elements representing how often the two co-occur in an observation.

Step 4 – conclusion

In this step the participants return to the main research question and together write a conclusion document based on the networks of step 3. In this document it is possible to edit each other's text (of which the original author receives a notification). It is also possible to search through the history of the document, and changes can be reverted.

Participants and administrators

Access to the tool is only possible for people that have been registered and are logged in. All participants have a personal account on the basis of an email address. When they are logged in they receive notifications when changes are made to their observations, interpretations, networks, and contributions to the common conclusion. When participants have not logged in for some period of time but there are notifications waiting, an email can be sent that invites the person to log back in. It is also possible for participants to send personal messages to each other.

Administrators/moderators always have access to all participants, observations, interpretations, networks and the common conclusion. They are able to make changes and remove these, of which an optional notification is sent to the concerned participant.

Analytics

On the basis of what the participants are exactly doing in the tool detailed analyses can be executed. In this way it is possible to gain insight into what observations, interpretations, networks and sections of the conclusion get the most attention, and by what (groups of) participants this is mostly done.

There is also the possibility to make small changes and present these to a randomly selected half of the participants. On the basis of analyses of the results of these A/B-tests, subsequent design choices can be made.

Tenders

In a tender we expect:

- To see the core aims of this project and an impression of how you understand this project
- Insight into the skills that you think are needed to build the tool, and to what degree you possess these
- A detailed planning and distribution of tasks
- Technical specifications like used programming language, how open values will be pursued, manuals and informative documentation at completion and aftercare
- Description of the ownership of the tool as you envision it
- A detailed budget (maximum €25.000,00-)
- The tool should be piloted in December 2015 / January 2016 with a group of 500+ participants

Tenders can be sent before the 24th of October to Nynke Kruiderink, kruiderink@uva.nl. With questions you can also call; +31 (0) 20 5254929