

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/254735438>

Analyzing Internet Forums: A Practical Guide

Article in *Journal of Media Psychology Theories Methods and Applications* · January 2012

DOI: 10.1027/1864-1105/a000062

CITATIONS

20

READS

881

3 authors:



[Peter Holtz](#)

Leibniz-Institut für Wissensmedien

31 PUBLICATIONS 316 CITATIONS

[SEE PROFILE](#)



[Nicole Kronberger](#)

Johannes Kepler University Linz

48 PUBLICATIONS 1,476 CITATIONS

[SEE PROFILE](#)



[Wolfgang Wagner](#)

Johannes Kepler University Linz

117 PUBLICATIONS 3,002 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Analytics for Everyday Learning <http://afel-project.eu/> [View project](#)



psychologie interculturelle [View project](#)

All content following this page was uploaded by [Peter Holtz](#) on 18 September 2017.

The user has requested enhancement of the downloaded file.

Published as:

Holtz, P., Kronberger, N., & Wagner, W. (2012). Analyzing internet forums: A practical guide. *Journal of Media Psychology*, 24(2), 55-66. doi: 10.1027/1864-1105/a000062

Analyzing Internet Forums: A Practical Guide

Peter Holtz^{1,2}

Nicole Kronberger²

Wolfgang Wagner^{2,3}

RUNNING HEAD: Analyzing Internet Forums

1) Friedrich Schiller University Jena, Germany

2) Johannes Kepler University, Linz, Austria

3) University of the Basque Country, San Sebastián, Spain

Corresponding Author: Dr Peter Holtz, Friedrich Schiller University Jena, Am Planetarium 4, D-07743 Jena, Germany. E-mail: peter.holtz@uni-jena.de

Submitted at the Journal of Media Psychology (JMP) Sept. 1st 2011

Accepted for publication April 4th 2012

This article does not exactly replicate the final version published in the journal "Journal of Media Psychology (JMP)". It is not a copy of the original published article and is not suitable for citation.

Analyzing Internet Forums: A Practical Guide

Over the last two decades, the internet has profoundly changed the way people communicate, be it for private communication or discussions about political and societal issues (cf. Bargh & McKenna, 2004). An increasing number of studies in the field of psychology have used data from internet forums as analysis material (e.g. Abdulla, 2007; Copes & Williams, 2007; de Vries & Valadez, 2008; Galasińska, 2010; Holtz & Wagner, 2009; Sneijder & Te Molder, 2004; Tanner, 2001; Williams & Copes, 2005). However, the use of online material for psychological research can still be considered an emerging field (cf. Skitka & Sargis, 2006). Based on two studies of our own we will show how internet forums can yield useful data for social psychological research. In particular we focus on studies about ideologically ‘sensitive’ communities, which are difficult to access by other means.

First, we will discuss methodological issues concerning the use of internet forums for social scientific research in general such as the selection of appropriate forums and material from these forums, the necessary refinement of this material for computer assisted qualitative data analysis, and possible analysis strategies. Afterwards, we will show how internet forums can yield useful data for social psychological research by discussing two studies of our own. In the first example, discourses within a German Neo-Nazi forum are analyzed. The second example deals with discourses in internet forums for young German Muslims.

Methodological Issues

What is an Internet Forum?

An internet forum or message board is an online discussion site. Internet forums have a tree-like structure: usually, different topics are discussed within different thematic sections and sometimes sub-sections. Within the sections or sub-sections, users can start a discussion – a so-called ‘thread’ – with a ‘starter posting’. Other users can reply to the starter posting or to other users’ comments. These messages are called a post or a posting. In many forums,

threads and postings can be read by every internet user, but in order to achieve the right to post or to start a thread, users will have to register and log in. There are other forums where postings can only be read by registered users.

Why Analyze Internet Forums?

Abundant material: Besides commercially oriented discussion boards, there are also many forums run by religious, political, or other societal interest groups. Members of these groups may know each other in ‘real life’, and the groups may also be linked to an ‘official’ non-governmental organization (NGO), church, or party. Alternatively, they may just be virtual communities (cf. Rheingold, 1993), without the members knowing each other in person. Usually, these forums are used almost exclusively by members and supporters of the organization or community for discussing matters of concern of the respective interest group. Hence, such forums allow for an analysis of typical discourses taking place within such communities.

An obvious advantage of internet forums is the almost unlimited amount of material for analysis. Some forums, with thousands of users, feature millions of postings in hundreds of thousands of threads. Even small forums usually contain more than enough text material for any kind of social scientific analysis. Because the material exists already in digital format, labor-intensive procedures like the transcription of audio material are not necessary.

Natural data with few social constraints: Internet forums are hierarchically structured. Hence, it is comparatively easy to find, select, and sample segments of the forum relevant to the research question. In a sense, forums constitute a kind of not moderated ‘virtual focus group’ (Moloney, Dietrich, Strickland, & Myerburg, 2003), in which members of a community discuss topics without a researcher interfering and possibly influencing the expression of thoughts. Hence, material from internet forums can be considered as relatively authentic ‘natural data’.

Being natural data in a virtual social setting also means that the total set of opinions in the forum, very much like focus groups, is more than the posters would have produced individually. Contributions by other forum posters provoke new and often more detailed responses, which may clarify the thinking about an issue in groups more effectively than in individual interviews as Stephenson, Kniveton, and Wagner (1991) have shown for group testimony.

In contrast to face-to-face situations, the relative anonymity of the internet motivates contributors to more openness. In forums for radical, extremist or other ideologically sensitive communities, users will express their opinions more freely and may be less concerned about social desirability than in focus groups or other interview settings (Glaser, Dixit, & Green, 2002).

Public data: The public accessibility of the analysis material makes it comparatively easy for other researchers to retrace the analysis process, from the sampling of the analysis material to the final analysis, and to reappraise the original researchers’ conclusions. This adds more transparency to analyses than is usually achieved in qualitative research.

Potential Issues with Analyzing Internet Forums

Anonymity: Whereas the relative anonymity of the internet is an advantage as it reduces social constraints, it also complicates analysis insofar as there is normally only little socio-demographic information available about the users. Most often, users participate under a

fictitious 'nickname'. Even though the age and sex of the posters is often shown as part of their profile, there seems to be no way to verify this information. However, while there is relatively little information on the individual users, it is possible to characterize the social group organizing and using the forum by taking into account background information such as mission statements or introductory pages. Sometimes, forums are directly linked to NGOs and other interest groups. Even if this is not the case, the operators of the website will usually take an interest in explaining the purpose of their forum themselves. Hence, when the research focuses on discourses within social communities, the lack of information on the individual users may not pose too much of a problem.

Deindividuation: Another issue can be the users' tendency to make more extreme and more offensive statements on the internet (Williams et al., 2002) than they would in face-to-face situations, due to deindividuation effects (Lea & Spears, 1991; Reicher, Spears, & Postmes, 1995). Deindividuation is facilitated by the lack of identity-related information that characterizes the forums' relative anonymity. A tendency towards more aggressiveness in the virtual realm may be a problem for certain research questions, but there is some empirical evidence that, most of the time, users of internet forums are indeed giving their real opinions on certain topics, although at times in a relatively aggressive and offensive tone (Glaser, Dixit, and Green, 2002).

Privacy: From an ethical point of view, one could argue that communication in internet forums is private and should not be used for scientific analysis without the informed consent of the users. This issue was discussed in the fields of sociology, psychology, and medicine (e.g. Eysenbach & Till, 2001; Kraut et al., 2004; Skitka & Sargis, 2006). The general consensus seems to be that there are some forums that are designed to create a kind of private or semi-private communication sphere. This is the case, for example, with forums and chat rooms that serve as self-help groups for people suffering from certain diseases or for victims of crimes (Eysenbach & Till, 2001; Finn & Lavitt, 1994; Mo, Malik, & Coulson, 2008). In

such cases, and without the users' consent, research can indeed be seen as a massive intrusion into the users' private sphere.

The issue will be different for forums run by political, religious, and other social interest groups. One of the main functions of such 'public' forums is to inform non-members about the organizations' agenda and goals and to attract new members. Usually, the postings in these forums can be read by everybody. In our view, this justifies the consideration of communication within such forums as 'public behavior' (Herring, 1996), which is intended to be regarded by as many other people as possible. Analyzing such boards without consent may be less of an ethical concern.

Nevertheless, it should not be forgotten that users of such 'public' forums may at least be annoyed by the possibility of their postings being analyzed in scientific research (King, 1996). It is necessary, for each single case, to deliberate whether the public interest in the discourse within this respective group outweighs the group members' potential wish for privacy. In any case, the privacy of the users should not be compromised more than necessary. For example, there is usually no need to publish the users' nicknames in scientific publications, making them potentially identifiable. Furthermore, the researcher may try to avoid literal quotes from the forums because the respective postings could be found very easily with internet search engines. All in all, a publication should contain as little potentially identifying information about the individual users as absolutely necessary. Nevertheless, there is indeed a trade-off issue here between protecting privacy on the one hand and transparency and accessibility of data on the other. The researcher needs to consider how much information to provide on a case-by-case basis.

Representativeness: It must be acknowledged that in most cases it will not be possible to make claims regarding representativeness for a certain population. Not every member of a given social group may have access to the internet and only a few of those who have will engage in discussions within such forums.

If possible, the researcher can use triangulation strategies and compare the results of the analysis of internet forums with other data sources like focus groups, individual interviews, and quantitative or qualitative survey data. If triangulation is not feasible, the researcher will have to take these limitations into account in the interpretation. Apart from this, an analysis of material from forums in conjunction with background information about the website and its operators can, in any case, be used to formulate hypotheses regarding conditions under which such views as expressed in the forum are more or less likely to be found.

Corpus Construction

Selecting appropriate forums.

First of all, relevant forums must be defined and selected. When considering using material from internet forums as a data source, the researcher will most likely be interested in discourses within certain (online) communities. This implies, first, collecting potentially relevant websites and then checking whether the forums are amenable to analysis in terms of quantity.

The scope of the research project determines the criteria of group and content specificity. If researchers are interested in a small radical or extremist community, they will look for the forum which is most 'typical' for the discourse within this group and which contains the most material relevant to the topic in question (see exemple 1). Experts or 'insiders' can be helpful in judging the typicality of forums or in characterizing the community running and using the forum.

If researchers are interested in a wide range of discourses within society, they will either select a very large forum, e.g. associated with major newspapers or internet portals, or will compare different forums so as to maximize perspectives on a given topic (see example 2). The researcher can also proceed in a stepwise manner and start out with one or a few forums and then extend the corpus until saturation of perspectives is reached and no additional variety can be detected (Bauer & Aarts, 2000).

Selecting appropriate sections and threads.

After the selection of one or more relevant forums, the researcher will define relevant sections and threads according to the research questions. In general, we consider it preferable to sample more material than seems necessary, as it is always possible to skip unused material. Software packages for qualitative analysis allow for finding relevant text passages within a large data set quickly once the data are in a standard word processing format.

It may happen with large forums or when comparing a number of different forums that not all potentially relevant material can be included in the analysis. This is where sampling comes in. Besides random sampling procedures, it may be interesting to include postings within a certain time period. Then, only threads that transpire within a certain time span will be included in the analysis. Particularly when juxtaposing different forums, researchers are well advised to clearly define, justify and document their sampling strategies. A mix of incompatible sampling methods will obfuscate differences and similarities between the forums.

Some small forums may pose a problem if postings are not taken up by others and responses are scarce. In most cases, threads with more responses will be preferable, as they allow insights in the development of the discourse. For the same reasons, the researcher will also prefer threads featuring postings of a large number of different users. Under certain conditions it will be of interest to see which topics do *not* provoke much debate and do *not*

arouse interest within a community. Then, of course, focusing on threads with comparatively few responses is the method of choice (cf. figure 1).

< Figure 1 >

Refining the Material

Depending on the amount of textual material, software packages for computer-assisted qualitative data analysis can be helpful or even necessary for project and data management. For example, in our two exemplary studies, text corpora comprising more than 500,000 (study 1) and more than 800,000 words (study 2) were analyzed. Without software support, it is almost impossible to conduct such an analysis. The CAQDAS project provides a continuously updated overview of available software [endnote 1].

Even though material derived from forums is digitally available, it requires refinement for effective analysis. In our experience, the following steps have proven useful to ensure a smooth procedure. Note that none of the following steps should be done on the original material, but always on a copy, just in case you must go back to the original layout at some point: (a) determining the text format, (b) removing irrelevant elements of the website, (c) a uniform layout should be applied to the material.

(a) In a first step, the content of the website is exported into text processing software and saved on a local storage and backup medium. One should not forget that forums may suddenly go offline for various reasons.

(b) Second, it is to be decided which elements of the website are relevant for further analysis and which are not. For example, pictures embedded in postings, user avatars, emoticons/smiley, user information, and usernames can sometimes be of relevance for the analysis, but more often they are not and need to be removed along with automatically generated content of the web page such as headers, footers, forum rules, and advertisements.

Additionally, researchers have to decide how to deal with ‘quotes’ from earlier postings. If the goal of the study is, for example, to compare expressed opinions on the level

of the individual users, keeping quotes can bias the word count. In this case, quotes must be deleted automatically using search and retrieve functions of the text processing software or by hand. The more researchers are interested in ‘informal group opinions’ (Mangold, 1960) expressed within the forum, the less of a problem are quotes from earlier postings. An exception is where the researcher is interested in interaction patterns. Here it is necessary to explicitly take into account to which ‘parent’ message a poster is referring to (cf. Jeong, 2005). In this case, quotes should be replaced by markers that point back to the original post from where the quote was taken.

Finally, because in many cases the website material will appear as a table, the table should be converted to plain text for further analysis.

(c) To facilitate the identification of structural elements like postings and threads by computer programs, a uniform format must be applied to the corpus. We recommend first to replace all hard returns with soft returns. Afterwards, we recommend inserting a single hard, instead of a soft return where the original poster ends a paragraph and two hard returns to mark the end of the posting. This format allows software packages to code single postings containing certain keywords automatically.

According to our experience, it is helpful to assign every single thread as a primary document in itself in such software packages, also to account for postings belonging to certain threads and to increase the software packages’ processing speed.

Choosing Methods of Analysis

In principle, data from internet forums can be used for almost all kinds of qualitative analysis. We will focus in this section on the methods used in our studies and discuss the use of text-reducing methods like content analysis as well as text-enhancing ‘in-depth analysis’ methods, and will look at how to quantify results, if required. For each case, we will also discuss the efficient use of software packages.

Text-reducing methods.

The general idea of text-reducing methods like content analysis is to capture central themes and contents within a corpus of texts or to compare different text corpora with regard to their themes. As a means for reducing the complexity of the analysis material, the researcher or the research team will develop a set of relevant content categories (Krippendorff, 2004; Weber, 1990; Mayring, 2000).

One of the key methodological questions within a content analysis concerns the definition of the unit of analysis. We recommend using the single posting as unit of analysis. For us, this is the most natural transition (e.g. Krippendorff, 2004) in a communicative exchange within forums. Hence, whenever a posting contains elements that are defined as belonging to a content category, the posting will be coded as featuring this content category, no matter how often the respective category is referenced within the posting. This unit of analysis allows capturing conversation elements, where the sequential structure of the posts mirrors ‘turn-taking’ similar to natural conversations (Sacks, 1995). However, if the researcher is interested, for instance, in speech act analysis (Searle, 1975), different units of analysis may be required; for example sentences or propositions within postings.

Using postings as units of analysis in content analytic approaches is also useful when looking at different levels of analysis: These may be opinions expressed by individual users within different threads and sections and discourses about different topics within e.g. different discussion boards. All major software packages for qualitative analyses allow for such a matrix-approach combining different hierarchical levels of analysis.

Depending on the research question, different amounts of textual material from forums can be subjected content analysis. Within a more qualitative content analytic approach (Stemler, 2001; Mayring, 2000), researchers will select a comparatively smaller amount of text. Here, the scientist will develop relevant categories on the basis of the textual material itself. If a more ‘quantitative’ or automatized content analysis is chosen, the researcher will

establish categories prior to the analysis based on theoretical considerations and pre-studies ('a priori categories'; cf. Weber, 1990). In this case, a larger amount of material will be analyzed.

Our example 1 represents a more qualitative content analysis, while in example 2 we illustrate automatic coding of theoretically derived categories.

Text-enhancing methods.

If researchers are interested in in-depth qualitative analyses, it is necessary to look more closely at the contextualized discourses and to use text-enhancing qualitative analysis methods as well. To gain a better understanding of certain postings, the researcher will consider contextual and other background information on the forum and the respective social group, such as the introductory pages of the forum and forum rules, other statements from the social groups running the forums, or earlier research on the social groups in question.

In research within the varieties of 'discursive psychology' approaches (e.g. Potter & Wetherell, 1987; Wagner & Hayes, 2005), the researcher will put an emphasis on the style of communication about certain topics and the effects the styles have on participants' responses. Besides, the use of metaphors provides insights into participants' world views (Lakoff & Johnson, 1980). This task is greatly facilitated by modern software packages which assist the researcher in code-based theory building (cf. Lewins & Silver, 2007), for example through 'attach memo' functions, mapping tools, and tools for graphical visualization that highlight relationships between theoretical and discursive elements, issues, and themes.

Taken together, the hierarchical structure of forums informed by a previous content analytic structuring of the corpus makes internet forums a rich source for in-depth qualitative research.

'Quantifying' qualitative data.

Even when using primarily 'qualitative methods', some research will require 'quantifying' textual material as well. This may be the case if there is an interest in clearly determining

differences between discursive objects within a forum (case study 1) or if the focus is on differences between different groups and their dedicated forums (case study 2). An existing file with forum discourses makes it easy to analyze a large enough text corpus with statistical methods such as multidimensional scaling, clustering, or Correspondence Analysis.

Texts represent categorical data—words, topics, themes or objects—that can be arranged in cross-tabulations and subjected to Correspondence Analysis. The rows and columns of a cross-tabulation reflect row profiles, column profiles as well as the strength of association between row and column categories expressed by the cell count. A Correspondence Analysis first calculates the margin sums of columns and rows that are the average profile across rows and across columns. Second, each row profile is compared to the margin profile and a Chi-square distance or dissimilarity between row and margin profiles is calculated. The same is applied to column profiles. Both result in distance matrices that undergo ‘eigenvalue’ vector decomposition. The final result is a dimensional representation of each, row and column categories, as well as of their interrelatedness. The maximum number of dimensions in the result is $n-1$, where n is either the number of row or of column categories, whichever category number is smaller. Hence, a 4 by 6 cross-tabulation yields a maximum of 3 dimensions: The number of row categories (4) is less than the number of column categories (6) and ‘4’ minus ‘1’ is ‘3’. Each dimension is has a “percentage of explained inertia”, which can be roughly understood as the degree of explained ‘variance’ in the data set. Usually, only the graph of the first two dimensions is used to visualize the data. Note that mere proximity of a column and a row category is not enough to establish a relationship. Instead cosine distances are used (for details see, e.g., Greenacre, 1993). The practical use of this method will be shown in the next sections.

Alternative Research Strategies.

Data from internet forums also allow for longitudinal studies. The creation date of the postings is usually displayed below or above the postings and can be coded automatically

with software packages. The only practical problem can arise if other dates (e.g. registration dates) are also displayed. In most cases, however, these problems can be solved using sophisticated pattern match functions like, for example, the GREP search function of Atlas.ti (cf. Muhr & Friese, 2004). Alternatively, the researcher must delete such entries by hand.

Research strategies similar to participant observations, quasi-experiments, and natural experiments can be conducted as well. Online ‘participant observation’ allows attaining a close and intimate familiarity with forum members over a period of time and to observe their practices in a ‘natural’ environment. Nancy Baym’s (2000) study on a soap opera fan community falls under this category.

For quasi-experiments, researchers will participate in a discussion and post certain messages to evoke responses, e.g. in different forums. Subsequently, one can analyse the reactions of the users towards these messages. Glaser, Dixit, and Lee (2002) used such an approach within US-American White racist internet chat rooms.

For research designs à la natural experiments, the researchers will systematically compare, for example, simultaneous reactions to relevant societal and political events in different forums. Recently, Holtz, Wagner, and Sartawi (forthcoming) analyzed reactions to the ‘Swiss minaret ban’ in secular/moderate as well as fundamentalist/Islamist German-language forums that let the authors discern the fundamentally different styles of discourse in Salafist and more secular online communities. Additionally, internet forum data also lend themselves to be used with other approaches, such as social network analysis as well.

Example 1: Racist Discourse within a German Neo-Nazi Forum [endnote 2]

Background

In the course of a research project on essentializing discourse and its consequences for racism, prejudice, and stereotypes (cf. Wagner, Holtz, & Kashima, 2009; Wagner, Kronberger, Nagata, Sen, Holtz, & Flores-Palacios, 2010), we analyzed the tendency to essentialize and naturalize outgroups in an openly racist extreme-right online community: the ‘Nationales

Forum Deutschland', NFD (Holtz & Wagner, 2009). Projecting essence onto a social category means to think, talk, and act as if the respective category were a discrete 'natural kind' and as if its members were all endowed with the same immutable attributes determined by the category's essence (Wagner, Holtz, & Kashima, 2009). Within this study, we first looked at which topics and themes are connected to which social groups and then analyzed the role of essentializing or naturalizing discursive strategies.

The analysis material

The forum is closely connected to the NPD (National Democratic Party of Germany), which can be classified as an extreme-right party (cf. Rydgren, 2007). As of July 13th, 2010, the NFD was being used by 1,996 registered members. It featured 127,459 postings in 14,924 threads [endnote 3]. The forum consists of four thematic sections (general topics, news, politics, and regional issues), which are further divided into 34 thematic subsections.

We analyzed all of the 4,997 postings (as of August/September 2006) in the 545 threads from the sub-section 'society' (Gesellschaft) from January 1st, 2004 through June 25th, 2006. User information, avatars, pictures, and smileys/emoticons were included as well. The 4,997 messages were authored by 317 different users.

Privacy and ethical aspects

As far as the privacy of the users is concerned, the NFD is described in the introductory page as a tool serving the networking between NPD and NPD sympathizers. It is intended to inform about NPD positions and party activities and to attract new members to the organization. Hence, the messages in this specific forum can be seen as 'public behavior'. There is also a strong societal and social scientific interest in extreme right-wing discourses and it would be difficult to access this community by other means.

Qualitative content analysis

After technically preparing the analysis material for a computer-aided analysis with Atlas.ti, we first coded every posting that referred to a particular social group that we determined

beforehand. The theoretically identified target groups had to be mentioned in more than 50 postings. The groups which met our criteria and which were ultimately included in the analysis were Africans/Blacks, Jews, Muslims, Poles, and Turks.

Following this, we searched for relevant and frequently occurring content categories, or themes, in an iterative process according to Mayring (2000). These were: exploitation, mind control/harassment, conspiracy, German identity, procreation, sex and rape, criminal behavior, and foreign infiltration. Two independent raters coded a sample of 52 messages with an inter-rater reliability of Cohen's Kappas ranging from .70 (foreign infiltration) to 1 (sex & rape).

Groups and themes were then cross-tabulated and subjected to Correspondence Analysis. The two-dimensional graph (figure 2) shows the extremist discussions connect social groups to specific themes. Muslims, Turks and Poles are clustered together with close proximity to the themes foreign infiltration, sex and rape, and criminality. Blacks/Africans as well as Jews took separate positions and were related to different themes: German identity and procreation for the former group and exploitation, conspiracy and mind control/harassment for the latter.

< figure 2 >

Discursive analysis of essentializing strategies

The relation between the three group clusters and the themes was analyzed in-depth. For example, the analysis reveals that extreme right-wing discourse heavily essentializes the target groups of Jews and Africans/Blacks and ascribes them immutable group-specific attributes that effectively liken them to different 'natural kinds'. Thereby, the group of Jews appears as a kind of their own, with super-human powers and influence. Africans and Blacks are despised, firstly because the supposedly 'essential' characteristics prohibit them to be categorically mixed with Germans (i.e. to become German by nationality) due to their incompatible essence, and secondly, when they procreate with Whites. Such procreation

produces ‘bastards’ without identity that are met with disgust. Such ‘hybrids’ are unanimously regarded as ‘miserable creatures’, which are not able to lead fulfilling lives (for details see Holtz & Wagner, 2009).

Discussion example 1

This study of a German Neo-Nazi forum allowed us to analyze discourses within an extreme right-wing community that would have been almost impossible by personal interviews. Neo-Nazis may not have participated in interview or focus group studies (not speaking of the discomforts for the researchers); if they had, they would have avoided certain issues and talked about other groups very differently than when they feel being among like-minded people in the forum, thereby superficially obeying the relatively strict laws in Germany against sedition (cf. Frindte, Wettig, & Wammetsberger, 2005).

Example 2: Lifeworlds of Young Muslims in Germany

Background

In this section, we present results from a study on lifeworlds of young Muslims in Germany. One part of this project encompassed the analysis of data from nine different forums used first and foremost by young German Muslims. Apart from this, the project also featured a longitudinal questionnaire survey and focus groups in three different German cities.

The analysis material

In a first step, we explored the range of German-language forums for young Muslims with the help of two research assistants with Turkish as a paternal language. About 50 of these forums were considered for analysis. Based on the size of the forum and the group and content specificity, we selected eight of these forums for closer analysis. Another forum (Delikanet) was included at a later stage to have more material on Turkish-German communities available. The forums’ orientations ranged from secular and moderate to very religious and Islamist (Moghadam, 2005) (see table 1).

In a next step, we identified four fields of interest, which further guided our selection of postings and threads from the forum sections and sub-sections: religion, society and politics, youth-related issues, and family, marriage, and partnership. We balanced the proportion of material for each of the themes across boards and sampled subsections for further analysis. Except for one forum [endnote 4], we only used threads with more than ten postings responding to the initial. Whenever possible, we sampled 700 to 1,000 contributions from each forum posted between early 2008 and mid-2009. Our sampling significantly reduced the material from larger forums, while we used all analyzable material from the smaller Ansarnet and Muslimaboard forums. Altogether, 6,725 postings in 377 threads were analyzed (see table 1 for details). Due to the abundance of postings, we decided not to include pictures, avatars, etc. in the analysis.

< Table 1 >

Privacy and ethical aspects

All forums state that they intend to inform others about the respective community's aims and goals and to connect similar minded individuals in their site description. Hence, postings in all of these forums can be considered public behavior. We decided to use these data without informed consent, but we did not analyze any personal information that might have been mentioned.

Analysis Strategy

For further procedures we used both text-reducing and text-enhancing methods. In a first step, we used an automatized content analysis and quantified the results using Correspondence Analysis. In a second step, we compared discourses on the relevant topics between the different forums.

Content analysis and correspondence analysis.

For the content analysis, we started out with 90 categories, which resulted from either theoretical considerations or from a first run through the material. For each content category,

keywords were defined, taking into account all possible spellings and synonyms. Next, we used Atlas.ti to automatically code every posting that contained at least one of the keywords or synonyms of interest. Afterwards, one of the Turkish speaking researchers checked a control sample of postings for category and keyword consistency.

The resulting categories were prepared for statistical analysis: Very rare categories were dropped and very similar or overlapping categories were merged. For example, the category ‘religious expressions’ was not included in the following analysis, because some forums used religious expressions like ‘Insha’Allah’ (if it is God’s will) in virtually every posting. Finally, the cross-tabulation of the nine forums by the remaining 40 content categories was subjected Correspondence Analysis (figure 3).

< figure 3 >

Correspondence Analysis guides further qualitative inquiry.

In the following, we interpret the forum clusters and their relation to themes to illustrate the ‘typical’ discourses (figure 3). The Vaybee and the Delikanet forums are positioned in cosine proximity to themes of integration and everyday life as a foreigner in Germany. Identity-related topics are also typical for these forums. The Vaybee and the Delikanet forums unite secular immigrant communities (mostly second- or third-generation immigrants from Turkey) who are motivated to integrate in the German host culture. They display hyphenated identities (cf. Sirin & Fine, 2007) by labeling themselves as German-Muslims, Turkish-Germans, or German-Turks.

The Way 2 Allah, the Misawa, and the Ammar 114 forums are positioned close to religious topics like prayers, commandments, and religious dialogue with Christians where the concept of ‘duty’ is important as well. They can be described as religious or – at least in the case of the Way 2 Allah forum – fundamentalist (Almond, Appleby & Sivan, 2003; Herriot, 2007). However, politics do not play an important role for them.

The Ahlu Sunnah and the Ansarnet forums are positioned close to topics that are typical for Islamist discourses, such as jihad and sharia. There are also frequent references to political issues such as the Near East and the Middle East conflicts. Revealingly, the term ‘kuffar’ (infidel) is used widely in these forums. Discursive analysis shows, on the one hand, that kuffar is used to emphasize the essential differences between Muslim and Western culture. On the other hand, it is used as a derogatory term for Muslims attempting to integrate into Western culture. These two communities can be described as Islamist.

The Dima Dima forum is positioned right between the Islamist and the secular forums with political topics being typical. It can be described as a forum where religion does not play a primary role but in which users sometimes express radical political views and in which heated political debates frequently take place.

Finally, the position of the Muslimaboard forum – a forum only for Muslim women – shows that in this forum, all thematic units are mentioned and discussed in an almost equal proportion.

Discussion example 2

The quantified information from Correspondence Analysis was used as a guide for the more in-depth qualitative inquiry. Determining the thematic orientation and prevalence of each forum would have been much more difficult by qualitative ‘eye-balling’ only. Hence, with a large text corpus, a first Correspondence Analysis is the method of choice.

Its result also allowed us to verify our sampling procedure to include forums encompassing moderate/secular, religious/fundamentalist, and Islamist/radical groups. Another interesting finding is that besides groups which are not very religious and politically radical (‘moderates’) and groups which are very religious and politically radical (‘Islamists’), there are also groups which are very religious but in no way politically radical, and groups which are politically radical but not very religious. Hence, when looking at radicalization processes among young German Muslims, one’s theoretical ‘space’ must be multi-

dimensional rather than one-dimensional to capture the various shades of life worlds existing in an immigrant community (Howarth, Wagner, Magnusson, & Sammut, forthcoming).

General Discussion

Data from internet forums can be used for a wide range of psychological research. In our own studies, we focused on social psychological research on ideologically sensitive communities, which are difficult to assess by means of interviews or survey studies.

The main advantages of internet forums are the amount of easily accessible data, the content and group specificity, and the naturalness and authenticity of the data. These advantages can also turn into weaknesses: there is relatively little information on the individual users and the researcher has to find methodological ways to effectively make use of the huge amount of data. Consequently, fitting sampling strategies must be developed for each case. Automatized or partly automatized analysis strategies can be used as a means for getting a first overview over the data and for facilitating further in-depth qualitative analyses. In our two exemplary studies, we used a combination of content and discourse analytic approaches. Whereas study 1 shows an in-depth analysis of discourses within one single online forum, in study 2, nine online forums, mainly used by young German Muslims, were compared.

In these examples, we used internet forums primarily for analyzing ‘informal group opinions’ (Mangold, 1960) within these communities. Such group opinions should not be understood in the narrow sense of consensually shared arguments or evaluations. More than concrete positions on matters of opinion, group opinions define *what is at the centre of attention* for a social group (p. 49). Whereas other lines of research focus on the particularities of online communication, here internet forums are used as a proxy for ‘everyday discourses’ within informal networks. Like in focus groups, the interdependence of statements by different persons is in this sense not a bane, but a boon, as analyzing reactions towards statements by other community members allows for insights into typical sense-making processes using unobstructed ‘naturalistic materials’ (Potter & Hepburn, 2005).

In every single case, the intrusion into the users' private sphere must be weighed against the scientific benefit of the study. In our opinion, forums that were created with the aim of informing others about the respective communities' goals and attracting new members are among the least problematic as far as privacy is concerned.

Every step of the research process, from the identification of relevant websites, via the sampling and preparation of the data, to the final analysis, is to be guided by the dominant research interest and should be thoroughly documented. Given the novelty of this line of research, detailed documentation of the researcher's decisions and their justifications are to be included in reports and papers in much more detail than with research using widely established qualitative methods. For the future these developments promise exciting new insights in social behavior.

References

- Abdulla, R.A. (2007). Islam, jihad, and terrorism in post-9/11 Arabic forums. *Journal of Computer-Mediated Communication, 12*, 1063-1081.
- Almond, G.A., Appleby, R.S., & Sivan, E. (2003). *Strong religion: the rise of fundamentalisms around the world*. Chicago: University of Chicago Press.
- Bargh, J.A. & McKenna, K.Y.A. (2004). The Internet and Social Life. *Annual Review of Psychology, 55*, 573-590.
- Bauer, M.W. & Aarts, B. (2000). Corpus construction: A principle for qualitative data collection. In Bauer, M.W. and Gaskell, G, editors, *Qualitative researching with text, image and sound: a practical handbook* (pp. 19-37). London: Sage.
- Baym, N. (2000). *Tune in, log on: Soaps, fandom and online community*. Thousand Oaks, CA: Sage.
- Copes, H. & Williams, J.P. (2007). Techniques of Affirmation: Deviant Behavior, Moral Commitment, and Subcultural Identity. *Deviant Behavior, 28*, 247-272.
- de Vries, S.R. & Valadez, A.A. (2008). Let Our Voices Be Heard: Qualitative Analysis of an Internet Forum. *Journal of Creativity in Mental Health, 3*, 383-400.
- Eysenbach, G. & Till, J.E. (2001). Ethical issues in qualitative research on internet communities. *British Medical Journal, 323*, 1103-1105.
- Finn, J. & Lavitt, M. (1994). Computer-based self-help groups for sexual abuse survivors. *Social Work with Groups: A Journal of Community and Clinical Practice, 17*, 21-46.
- Frindte, W., Wettig, S., & Wammetsberger, D. (2005). Old and new anti-Semitic attitudes in the context of authoritarianism and social dominance orientation—two studies in Germany. *Peace & Conflict, 11*, 239-266.
- Galasińska, A. (2010). Leavers and stayers discuss returning home: Internet discourses on migration in the context of the post-communist transformation. *Social Identities: Journal for the Study of Race, Nation and Culture, 16*, 309-324.

- Glaser, J., Dixit, J., & Green, D.P. (2002). Studying hate crime with the internet: what makes racists advocate racial violence? *Journal of Social Issues*, 58, 177-193.
- Greenacre, M.J. (1993). *Correspondence Analysis in Practice*. London: Academic Press.
- Herring, S. (1996). Linguistic and critical analysis of computer-mediated communication: Some ethical and scholarly considerations. *Information Society*, 12, 153-168.
- Herriot, P. (2007). *Religious Fundamentalism and Social Identity*. New York: Routledge.
- Holtz, P. & Wagner, W. (2009). Essentialism and attribution of monstrosity in racist discourse: Right-wing internet postings about Africans and Jews. *Journal of Community and Applied Social Psychology*, 19, 411-425.
- Holtz, P., Wagner, W., & Sartawi, M. (forthcoming). *Discrimination and Immigrant Identity Work: Fundamentalist and Secular Muslims Facing the Swiss Minaret Ban*.
- Howarth, C., Wagner, W., Magnusson, N., & Sammut, G. (forthcoming). 'It's only other people who make me feel black': Acculturation, identity and agency in a multicultural community. *Political Psychology*.
- Jeong, A. (2005). A guide to analyzing message-response sequences and group interaction patterns in computer-mediated communication. *Distance Education*, 26, 367-383.
- King, S.A. (1996). Researching Internet communities: Proposed ethical guidelines for the reporting of results. *The Information Society*, 12, 119-128.
- Koster, W. & Houtman, D. (2008). 'Stormfront Is Like a Second Home to Me': On Virtual Community Formation by Right-Wing Extremists'. *Information Communication and Society*, 11, 1-38.
- Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online. *American Psychologist*, 59, 105-117.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd edition). Thousand Oaks, CA: Sage.

- Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Lea, M. & Spears, R. (1991). Computer-mediated communication, de-individuation and group decision-making. *International Journal of Man-Machine Studies*, 34, 283-301.
- Lewins, A. & Silver, C. (2007). *Using software in qualitative research: a step-by-step guide*. London: Sage.
- Mangold, W. (1960). *Gegenstand und Methode des Gruppendiskussionsverfahrens [Subject and method of group discussions]*. Frankfurt a. M.: Europäische Verlagsanstalt.
- Mayring, P. (2000). *Qualitative Inhaltsanalyse: Grundlagen und Techniken [Qualitative content analysis: basics and techniques]*. Weinheim: DSV.
- Mo, P.K.H., Malik, S.H., & Coulson, N.S. (2008). Gender differences in computer-mediated communication: A systematic literature review of online health-related support groups. *Patient Education and Counseling*, 75, 16-24.
- Moloney, M.F., Dietrich, A.S., Strickland, O., & Myerburg, S. (2003) Using Internet forums as virtual focus groups. *Advances in Nursing Science*, 26, 274-286.
- Muhr, T. & Friese, S. (2004). *User's Manual for ATLAS.ti 5.0*. Berlin: Scientific Software Development.
- Potter, J. & Wetherell, M. (1987). *Discourse and social psychology: Beyond attitudes and behaviour*. London: Sage.
- Potter, J. & Hepburn, A. (2005). Qualitative interviews in psychology: Problems and possibilities. *Qualitative Research in Psychology*, 2, 1-27.
- Reicher, S.D., Spears, R., & Postmes, T. (1995). A social identity model of deindividuation phenomena. *European review of social psychology*, 6, 161-198.
- Rheingold, H. (1993). *The virtual community: Homesteading on the electronic frontier*. Boston, MA: The MIT Press.

- Rydgren, J. (2007). The sociology of the radical right. *Annual Review of Sociology*, 33, 241-262.
- Sacks, H. (1995). *Lectures on conversation* (ed. G. Jefferson). Oxford: Wiley-Blackwell.
- Searle, J. R. (1975). Indirect speech acts. *Syntax and semantics*, 3, 59-82.
- Sirin, S.R. & Fine, M. (2007). Hyphenated selves: Muslim American youth negotiating identities on the fault lines of global conflict. *Applied Developmental Science*, 11, 151-163.
- Skitka, L.J. & Sargis, E.G. (2006). The Internet as psychological laboratory. *Annual Review of Psychology*, 57, 529-555.
- Sneijder, P. & Te Molder, H.F.M. (2004). 'Health Should Not Have to Be a Problem': Talking Health and Accountability in an Internet Forum on Veganism. *Journal of Health Psychology*, 9, 599-616.
- Stemler, S. (2001). An overview of content analysis. *Practical Assessment, Research and Evaluation*, 7, 137-146.
- Stephenson, G.M., Kniveton, B.H., & Wagner, W. (1991). Social influences on remembering: Intellectual, interpersonal and intergroup components. *European Journal of Social Psychology*, 21, 463-475.
- Tanner, E. (2001). Chilean conversations: Internet forum participants debate Augusto Pinochet's detention. *Journal of Communication*, 51, 383.
- Wagner, W. & Hayes, N. (2005). *Everyday discourse and common sense: The theory of social representations*. Houndsmills: Palgrave Macmillan.
- Wagner, W., Holtz, P., & Kashima, Y. (2009). Construction and Deconstruction of Essence in Representing Social Groups: Identity Projects, Stereotyping, and Racism. *Journal for the Theory of Social Behaviour*, 39, 363-383.

- Wagner, W., Kronberger, N., Nagata, M., Sen, R., Holtz, P., & Palacios, F.F. (2010). An Essentialist Theory of “Hybrids”: From Animal Kinds to Ethnic Categories and Race. *Asian Journal of Social Psychology, 13*, 232-246.
- Weber, R.P. (1990). *Basic Content Analysis*. Newbury Park, CA: Sage.
- Williams, J.P. & Copes, H. (2005). 'How Edge Are You?' Constructing Authentic Identities and Subcultural Boundaries in a Straightedge Internet Forum. *Symbolic Interaction, 28*, 67-89.
- Williams, K.D., Govan, C.L., Croker, V., Tynan, D., Cruickshank, M., & Lam, A. (2002). Investigations into differences between social-and cyberostracism. *Group Dynamics: Theory, Research, and Practice, 6*, 65-77.

Endnotes

Endnote 1:

<http://caqdas.soc.surrey.ac.uk/Support/Choosingsoftware/softwareoptions.html>; retrieved December 13th, 2010.

Endnote 2:

This sections draws upon results published in Holtz & Wagner, 2009

Endnote 3:

<http://www.nationales-forum-deutschland.de>, retrieved July 13th, 2010 ; retrieved July 13th, 2010; the NFD went offline in late 2010.

Endnote 4:

The extremist Ansarnet forum was only used by a handful of users. Nevertheless, we considered this forum to be important as we also wanted to include data from a very radical Muslim community. Hence, we decided to use every possible thread from the Ansarnet forum regardless of the number of response postings. The Ansarnet forum went offline in early 2010.

Acknowledgements

Exemplary study 2 was supported by a research grant from the German Federal Ministry of the Interior to Wolfgang Frindte (Jena University, Germany), Klaus Boehnke (Jacobs University, Bremen, Germany), Wolfgang Wagner (Linz University, Austria), & Henry Kreikenbom (aproxima - market research agency, Weimar, Germany).

Tables

Table 1: Number of analyzed postings, threads, and sampling rules for all nine forums selected for example 2.

Forum name	Link to homepage	Number of analyzed postings	Number of analyzed threads	Percentage of eligible threads selected for further analysis
Ansarnet	de.ansarnet.info	274	44	all threads
Ahlu Sunnah	www.ahlu-sunnah.com	966	45	~7%
Ammar 114	www.team114-forum.de	1036	72	~8%
Delikanet	www.delikanforum.net	817	51	~10%
Dima Dima	www.dimadima.de	558	18	~30%
Misawa	www.misawa.de	741	27	~20%
Muslimaboard	www.muslimaboard.de	365	17	all threads
Vaybee	www.vaybee.de	974	40	~5%
Way 2 Allah	www.way2allah.de	994	63	~10%
Total:		6725	377	

Note: Among all forums, only threads with at least ten response postings were sampled; in the case of very lengthy discussions, only the first 50 postings were included; in the case of the Ansarnet forums, also threads with fewer than 10 response postings were included (see endnote2).

Figure Captions

Figure 1:

Sampling strategies in view of the hierarchical structure of internet forums

Figure 2:

Correspondence analysis of matrix cross-tabulating ethnic groups and themes. Squares indicate the position of groups (in italics); diamonds stand for themes; circles and ellipses visualize clusters of groups and themes discussed below (Holtz & Wagner, 2009, p. 418).

Figure 3:

Correspondence analysis of matrix cross-tabulating forums and related themes. Squares indicate the position of forums (in italics); diamonds stand for themes; ellipses visualize clusters of forums and themes discussed below.

Figure 1

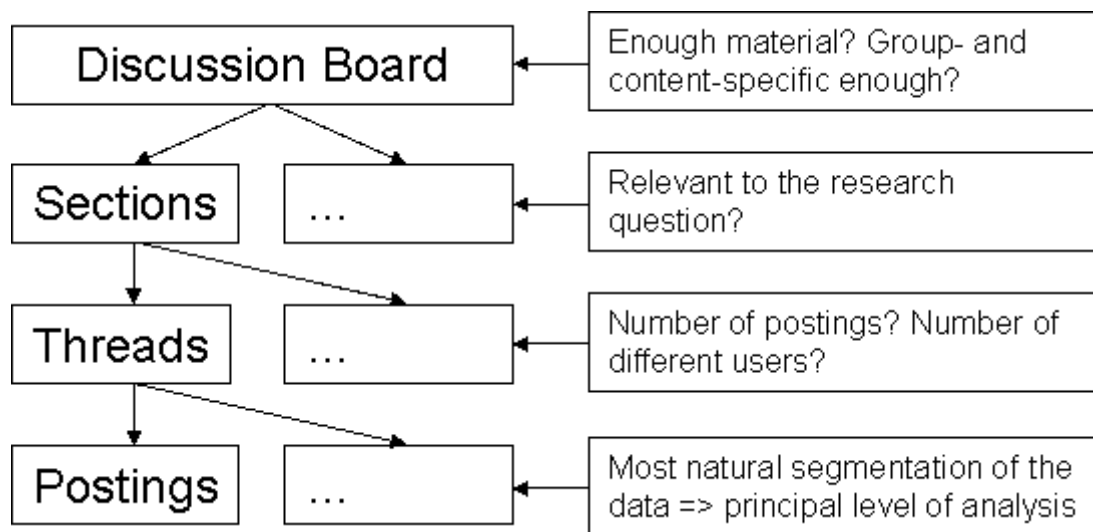


Figure 2

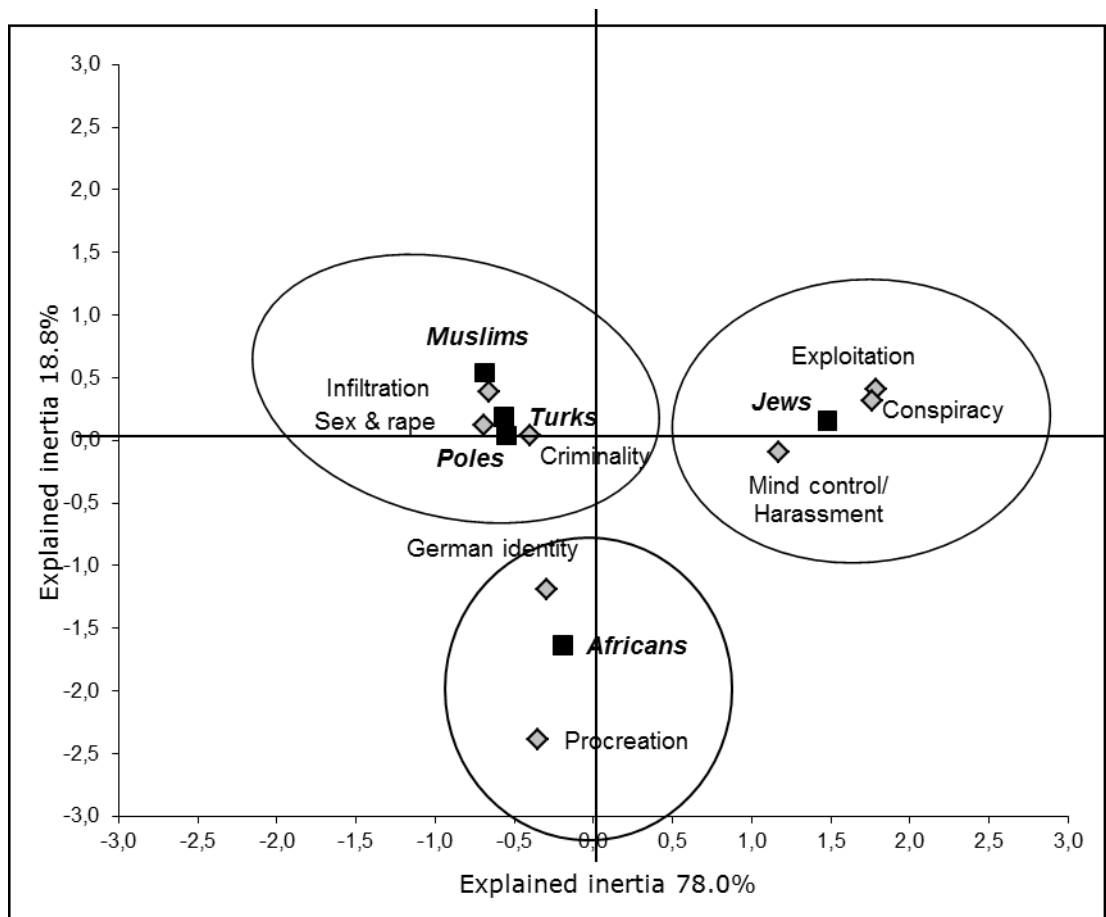


Figure 3

