Griesbaum, Joachim; Mahrholz, Nadine; Löwe Kiedrowski, Kim von; Rittberger, Marc

Knowledge generation in online forums. A case study in the German educational domain

formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:
formally and content revised edition of the original source in:
Aslib journal of information management 67 (2015) 1, S. 2-26, 10.1108/AJIM-09-2014-0112

Knowledge generation in online forums: a case study in the German educational domain

1. INTRODUCTION

Social Media adaption can be assessed as a paradigm change not only with regard to potentially unlimited communication but also with respect to knowledge generation (Shirky, 2008). The open social web has widely expanded the universe of available knowledge. Apart from the value of such knowledge as an information resource, active participation provides further benefits for the users involved. Asking questions in online forums or question and answering services offers the possibility to directly satisfy personal information needs. In addition, one may build up knowledge by engaging in ongoing conversations with peers. We will discuss the knowledge value of such social media a) for information seeking and b) for knowledge building purposes.

Key aspects are types of knowledge users aim for and pragmatic intents of questions, knowledge related characteristics of posts and also success. The aim is to derive a first approximation of the usefulness of forums seen from users' perspectives which we will name as knowledge value of the investigated communities. Three forums with a topical focus on education related communication serve as a test bed and case studies for the investigation. We selected forums which are targeted at teachers and students of education related study paths. In these forums, we expect that discussions focus on topics concerned with professional aspects of teaching and studying.

The paper is structured as follows. First, we give an introduction into social media usage in education related contexts. Following that, we provide an overview of current research on structures and quality of communication in social media. Subsequently, we expose theoretical considerations to provide an analytical framework. Then, the research questions are delineated and the research approach and methods are presented and discussed. Finally, data and results are described. We conclude the paper with a discussion and an outlook.

2. SOCIAL MEDIA USAGE IN EDUCATION RELATED CONTEXTS

The social web has created new possibilities for personal, communal and public information sharing (Shirky, 2010). Nevertheless, social media usage still can be seen as a primarily leisure based phenomenon. Whereas the adaptation of social media is advanced with regard to personal communication and self-expression in social networking sites (SNS) and product or service related information behavior, the significance and role of social media in professional contexts is often unknown.

Social media usage in educational contexts can be categorized into two main facets: First, the employment of social media in formal learning, e. g. using Facebook, Twitter and other social media as e-learning tools in courses and lectures. Second, the self-determined usage of social media for (mostly informal) learning. Although the first aspect is very interesting, the focus of this paper is on the second aspect, the self-determined usage of social media for information seeking and knowledge building purposes. This is what we call social information behavior. Here, we can differentiate between a) seeking and deploying existing user-generated content to satisfy information needs (e.g. reading a Wikipedia article) and b) active participation in social media to answer current information needs, for instance by asking a specific question in an online forum, and/or to build up knowledge by ongoing conversations with peers in communities.

With regard to a) the receptive use of social media, according to an online survey of Gibs (2009), there is a segment of online users who see social media as a "core to finding new information". With regard to b) active participation, Kleimann et al. (2008) found that a substantial fraction of students use social communities to communicate with peers about study related aspects. Kim et al. (2011) surveyed 446 undergraduate students and found that different social media types are used in different information seeking contexts. Whereas social network sites are preferred for everyday life purposes, question-answering services are used for leisure as well as academic needs.

In short, we can see that in education related information contexts too, social media can be assessed as an integral part of the information environment and resources users rely on. But is it worthwhile to employ social media for such information purposes? Our research tries to serve as a starting point to generate insights into the specificities, the potential benefits and problems of social information behavior in online forums for education related contexts.

3. Related literature

The information and knowledge value of social media is inevitably linked to the sequence of development and the specificities of the content created by the users. The concept of knowledge value is difficult to grasp. We will discuss and operationalize our specific perspective in section 4. In order to classify our investigation into current research, we provide an exemplary overview of investigations and approaches concerned with a) the analysis of communication structures in social media and quality aspects of user generated content, b) the role of forums in information seeking and c) online knowledge building and discourse.

3.1 Research on communication structures and quality of user generated content

There is manifold research with regard to the characteristics of social media communication and quality of user generated content. Investigations often focus on attributes of authors and contributions.

In reference to the Wikipedia, Kane (2011) investigated 188 articles. He concludes that content shaping (cleaning, structuring, formatting etc.) and top contributor experience are positively related to the quality of Wikipedia articles and that e.g. anonymous contributors are negatively associated with the quality. The volume of contributions and contributors has no relation. Liu & Ram (2011) analyzed the revision history of 1,600 articles and identified collaboration patterns. They also state that the quality of articles depends on the authors and their role in the collaboration process.

With respect to forums and question and answering services, Agichtein et al. (2008) developed a scheme to model the quality of content in Yahoo! Answers. Based on a manual analysis of 6,665 questions and 8,366 question-answer pairs for "well-formedness, readability, utility, and interestingness", they identified relationships of content, interactions between content creators and users and usage statistics as predictors of quality. According to their analysis, answer length, the number of words in the answer with a corpus frequency larger than a threshold, unique number of words in the answer and the overlap of non-stopwords between the question and the answer are among the most important indicators of answer quality.

Savolainen (2011a) investigated information quality and credibility of posts in Internet forums. The author analyzed the content of 4,739 posts of 160 threads on a Finnish forum with two value-laden subject areas (natural products and issues of racism). The postings were analyzed to discover explicit assessments of the quality and credibility of posts from other authors. According to the results, 20.5% of all posts contained assessments of the quality and credibility of contributions from other participants. Both, positive (e.g. validity) and negative (e.g. dishonesty) criteria were used to evaluate other users' messages. Usefulness, correctness, specificity and objectivity were the most often mentioned criteria for information quality. In addition to the investigation of Agichtein et al. (2008), the paper from Savolainen

(2011a) shows that quality criteria and assessments of user generated content can also be explored explicitly in the sequence of the online discourse.

Concerning social network sites, Cvijikj & Michahelles (2011) conducted a study on the topics, categories and sentiment of posts on a Facebook brand page. The data collection period encompassed a whole year. 611 user posts were manually analyzed. The following aspects were analyzed: topics within the posts, posts intentions and sentiment of the content. Results indicate that product, sales and brand are the most important topics. Intentions were mainly suggestions, requests, affect expressions and status sharing. Complaints and critique were scarcely observed. Likewise, sentiment was predominantly positive. Topics and intentions were often correlated. Product requests and suggestions and affect expressions and products were the most frequent topic-intent combinations. Although that is just a case study, the research approach shows that by grasping and combining multiple categorizations it is possible to get deep insights into various facets of online communication.

3.2 The role of forums in information seeking

With regard to the role and significance of forums for receptive information behavior, the research focus is on pragmatics and topics of information seeking in forums.

Hasler et al. (2014) explored the usage of newsgroups and discussion groups in situations of information poverty. Searching for specific phrases which indicated critical and hidden information needs users were unwilling to reveal elsewhere (e.g. “I cannot tell anyone…”) they collected 450 posts of which 200 were subject to further manual topic analysis. The authors identified 21 different topics of which health issues, relationships and pregnancy were the most prominent. The authors conclude “that these online environments provide an outlet for the expression of critical and hidden information needs”.

Pointing in the same direction, Bickart & Schindler (2001) argue that forums are a place where personal experiences focusing on real world phenomena can be found. Thus, information in forums may be more relevant to the reader than information provided by professional providers who may have less knowledge and empathy of users’ pragmatic needs. Bickart & Schindler (2001) investigated effects of forums and corporate websites as sources of consumer information. 61 students of a marketing course were surveyed after gathering online information about one of five specific product related topics from forums and corporate websites. Data denotes that forum readers showed more interest in their topics than readers of business websites.

Results of the investigation of Savolainen (2011b) suggest that forums may serve as appropriate places for “presenting questions to potential helpers”. The author investigated a Finnish discussion forum and 10 blogs focusing on the topic depression. Information needs were categorized into the following types: *opinion or evaluation*, *factual information*, and *procedural information*. Answers were distinguished with regard to *personal knowledge* and different referential types (*expert*, *networked sources*, *printed sources*…). By analyzing information needs and answers of 1,044 blog posts, 1,727 blog readers’ comments and 1,236 forum posts of 40 threads, the author concludes that the majority of needs’ focuses is on getting opinions or evaluations on issues and that the provision of knowledge relies primarily on personal knowledge of the participants.

In contrast to that, Chuang & Yang (2014), who investigated informational support on forums, personal journals and notes in the MedHelp community, conclude that fact oriented information is the most often exchanged content type. The study comprised a manual analysis of a dataset of 493 forum posts, 423 user journal entries and 1,180 notes concerned with the topic of alcoholism. The coding scheme distinguished between *opinion*, *personal*, *advice*, and *fact oriented information types*.

In sum, these studies point out a specific role and significance of forums for information seeking, especially in pragmatic contexts, where other information resources, e.g. search engines, are not able to satisfy information needs. Connected

to this and in addition to factual information, forums offer access to personal estimations and subjective knowledge. As Chuang & Yang (2014) state, success of communication or helpfulness of participation still are not very clear, meaning that research in this field is still at the very beginning.

3.3 Online knowledge building and discourse

Concerning online knowledge building and discourse, the research focus is on features of analyzed communities and on discourse analysis. In addition to studies which are related to the field of Computer Supported Communication (CMC) and Enterprise Social Networking (ESN) it is helpful to refer to research in the field of e-learning, especially Computer Supported Collaborative Learning (CSCL).

Aschoff et al. (2011) tried to capture discourse quality with a combination of statistical and manual analysis of 34 forums. The authors chose threads as the basic unit of analysis. Discourse quality was defined as a combination of quantitative and qualitative measures. Quantitative measures encompassed discourse statistics as answer probability, volume of answers, reply delay, and the number of users who post an answer. Qualitative analysis focused on topical relationship with regard to the initial post and the forum, as well as post type attributes like *answer*, *follow up question*, *aggressive behavior*, in sum 14 categories. 469 threads with 1772 posts from 34 online travel forums were analyzed. According to the results, larger communities show a higher discourse quality because of a higher activity rate and a stronger topic alignment.

Burnett & Buerkle (2004) compared two Usenet communities with a topical focus on health information. One Usenet group discussed a serious issue of physical health and the other an emotional disorder. The authors conclude that both communities show very different characteristics with regard to communication style. Whereas the group focusing on the physical health issue discussed in a mainly functional manner, the communication in the other forum featured more than 41% of posts concerned with hostility or spam. The results reveal one important point: The relevance of emotional aspects (support and hostility) in online communication. In addition, the authors discuss and illustrate important aspects of categorization schemes which try to grasp qualitative attributes of online communication. First, such schemes usually simplify online communication to a large extend. One cannot really be sure that the relevant concepts of investigations are sufficiently measured. There is often a temptation to refine such schemes, which on the other hand leads to a reduced feasibility of analysis or at least a higher effort of coding. Second, posts are very context dependent. In an extreme case, similar posts that have to be categorized with the same attributes could be of totally different relevance or pragmatic depending on the state of the discussion.

With regard to knowledge building within a company, Riemer & Scifleet (2012) investigated the communication in a microblogging platform according to its collective purpose. 1,809 messages were categorized by one researcher in an iterative bottom-up coding approach. 18 genres were detected, which were grouped into 7 top-level categories with *Discussion, Information Sharing, Updates, Problem Solving & Advice, Social & Praise* and *Idea Generation* as the most frequent types of activities. Thus, the authors differentiate between different intents of knowledge communication (providing input, creating new knowledge, harnessing existing knowledge) and also conclude that building common ground and social relationships can be seen as a prerequisite for all knowledge work.

In the context of e-learning there is manifold research with regard to online knowledge building and discourse. Clark et al. (2007) present an overview of analytical frameworks of discourse examination in the field of computer supported collaborative learning (CSCL). The researchers emphasize research methods that need to consider the context-specific nature of argumentation. The paper structures such analytical approaches into four categories. Methods that focus a) on the nature and function of contributions within the dialog, b) on the nature of reasoning, c) on the conceptual quality, d) on the patterns and trajectories of participant interaction. An example of a) is developed by deVries, Lund, and Baker (2002) who used four main categories for analysis (explanation, argumentation, problem resolution and management) which were further divided. According to Clark et al. (2007) analytical schemes that focus on the nature and function of

contributions within the dialog are well suited to capture the epistemic course of events in online discourse. Schemes that focus on the nature of reasoning in online discussions (b) grasp the types of reasoning within an argument. For example, Duschl (2007) employed *requests for information, expert opinion, inference*, and *analogy* as categories to describe the level of reasoning in contributions. Although differentiating between the different types may be difficult in coding, such approaches allow a deeper understanding of the argumentative quality of contributions and discourse. Frameworks that capture the conceptual quality of the contributions focus on the structure of the arguments (c). For instance, Kuhn and Udell (2003) tried to measure the logical coherence of arguments with categories like *Nonjustificatory Arguments*, *Nonfunctional Arguments* and *Functional Arguments*, with only the last type connected to core aspects of the problem. Finally, to illustrate approaches that address patterns and trajectories of interaction during the course of argumentation (d), Weinberger and Fischer (2006) analyzed knowledge construction along four dimensions: participation of students, epistemic assessment of theoretical constructs, formal argumentative quality and social modes of co-construction. These kinds of approaches are the most complex and are able to describe "an entire knowledge building cycle".

In sum, the overview in sections 3.1-3.3 clearly shows that there are manifold approaches and perspectives to grasp communication structures and quality of user generated content. Many investigations into the social web combine statistical measurements with a manual analysis of the content and attributes of contributions. Units of analysis range from words and statements within single contributions, to single contributions and whole threads, in the case of Aschoff et al. (2011) even forums. Furthermore, forums can be seen as a special kind of information resource. In addition to factual information, they provide access to personal opinions. With respect to knowledge building and discourse, research suggests that both factual and emotional aspects could play an important role as determinants of the quality and success of communication in social media. With regard to this, Clark's et al. (2007) overview illustrates that, depending on the particular research goals, categorization schemes can be very different. The fact that there is a relation between quality and type of content and learning becomes very obvious. Finally, the literature review points out that there is a trade-off between depth of analysis and costs (Burnett and Buerkle, 2004).Therefore, choosing the right granularity and complexity of analysis is not trivial, especially as the role and relevance of content (posts) is dependent on the state of the discussion.

4. THEORETICAL CONSIDERATIONS AND CONCEPTUAL FRAMEWORK

As mentioned in the introduction, we aim to estimate the knowledge value of online communities for information seeking as well as knowledge building for a wider audience. This double-sided perspective is intentionally chosen as discussions on information problems in many-to-many computer mediated communication spaces may be of value on the individual level of the questioner but also on higher social levels of the community. In the following, we relate basic concepts from diverse theories in the fields of Information Behavior, Social Learning and Knowledge Management to provide a conceptual framework for this research and categorize our understanding of social information behavior. Please keep in mind, that we estimate the argumentation here are as a first plunge into discourse and not the end of the discussion.

At current, research and models in the field of Information Seeking (Case 2012) mainly focus on the information needs and practices of individual information seekers and their information seeking processes and various contextual factors. Some models also take into account information exchange between the information seeker and other people (e.g. Wilson 1999). Evans and Chi (2008) investigated users' information exchange with other users throughout the search process and concluded that social exchange plays an important role in information seeking.

Beyond that, Robson and Robinson (2013) argue that information behavior also includes communication and provision of information and communication between information users and information providers. In online communities, users

act as information consumers and providers at the same time. Therefore, in such environments social exchange and communication is not only one important facet of information seeking but rather at its core. At the same time, such human-to-human information exchange resembles a many-to-many communication context, possibly including a large number of human information providers, which in turn may also act as information consumers. Hence, social information behavior can be connected to ideas of Knowledge-building-Communities as laid out for example by Scardamilia and Bereiter (1994) and communities of practice as described by Wenger (1998). With regard to this, we see a second aspect of social exchange which we want to connect to our understanding of social information behavior: collaborative learning and knowledge building. Developing an openly available knowledge base is of potential value for the user with the initial information need and also for the other users involved and possibly even helpful for users only passively "consuming" the knowledge base. Hence, social information behavior may result in a double gain: satisfying ad hoc information needs of users and also developing knowledge which may be helpful for higher social levels of communicants and recipients. Such a connection between information seeking and learning can be argued for example in reference to the concept of exploratory search as stated by Marchionini (2006). According to this, there is a broad range of *searching to learn* activities and goals which can be connected to Bloom's taxonomy of educational objectives (Krathwohl, 2002). As White & Roth (2009: 74) state, *it may be in the searcher's interests to collaboratively explore the information space and participate in shared learning.* Information seeking can be seen as a part of learning and collaborative knowledge generation can also be assessed as social information provision. This connection between the two fields is widely unexplored in Information Science.

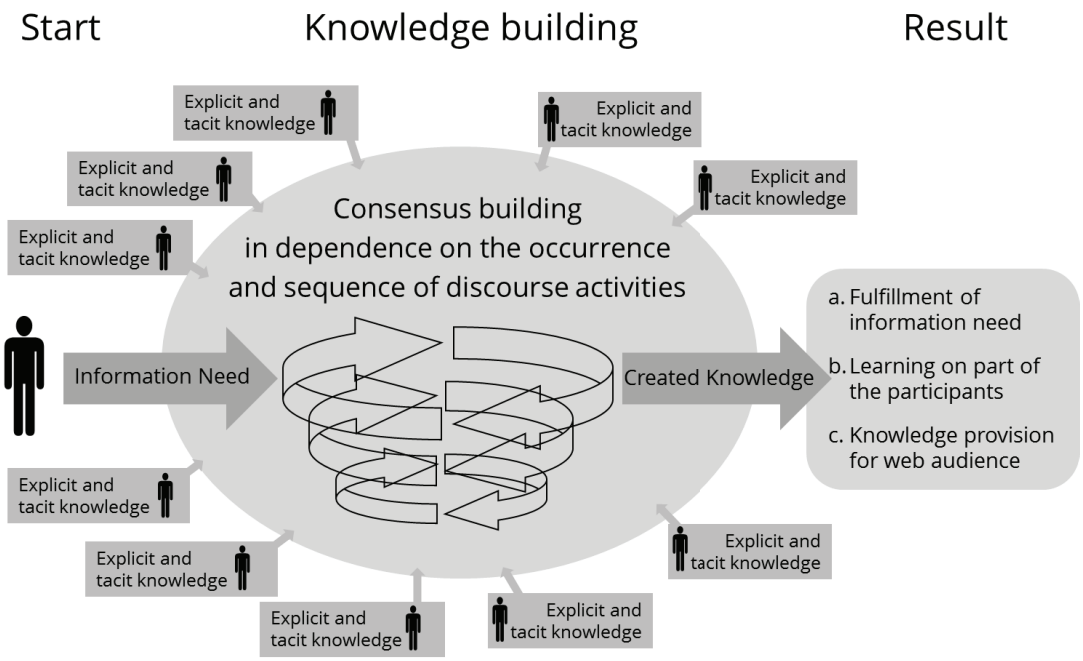
But how can we capture the "value" of such knowledge generation? On the one hand, the value for the user depends on the satisfaction of their information need. On the other hand, there are different theoretical perspectives which can be employed as basic elements to determine the quality aspects of collaborative knowledge development. According to Vygotski's perspective of learning as a discourse-dependent development to a higher level of expertise (Vygotski, 1986; Vygotski, 1979), collaborative environments promote the individuals internalization of other group membeŕs knowledge or of commonly achieved new group knowledge. Socio-genetic perspectives (Piaget, 1979) argue that communication and interaction increase the probability of cognitive conflicts which alter and enhance the internal knowledge structures of individuals through accommodation (De Lisi & Golbeck, 2005). Cognitive elaboration perspectives state that collaborative learning enhances basic information processing activities as encoding, schema activation, rehearsal (O'Donnell, 2006). While externalizing their knowledge, learners do not only provide information to others but also foster their own internal cognitive structures also trying to close knowledge gaps. If we agree on the relevance of the above mentioned activities for knowledge building, we should be able to operationalize them as quality indicators by mapping them to behaviors or contributions in the discourse. In turn, the occurrence and sequence of these discourse activities could be employed as a framework to evaluate quality aspects of information need satisfaction and knowledge generation. Weinberger (2003: 13-19) distinguishes *Quick consensus building*, *Integration-oriented consensus building*, and *Conflict-oriented consensus building* as patterns which reflect different epistemic levels of discourse development and outcome. Although there is no strict relationship between these patterns and quality, one can employ them to get an impression on the depth and extend of cognitive aspects of collaborative knowledge generation. For example, a significant number of cognitive conflicts indicate a rather socio-genetic *Conflict-oriented consensus building* pattern whereas a significant fraction of externalizations indicates *Integration-oriented consensus building.* A low number of cognitive discourse activities could be connected to *Quick consensus building* of only shallow knowledge value.

Lastly, seen from a knowledge management perspective, collaborative knowledge production can be estimated to be of special value because it has got the potential to externalize knowledge which is usually not accessible at all – tacit knowledge. Tacit knowledge is personal knowledge which is difficult to externalize. The concept was first coined by Polanyi (1958) and is widely used in the field of Knowledge Management. It is assessed as a primary success factor of knowledge development and organizational success (Nonaka & Takeuchi 1995). Tacit knowledge encompasses personal

skills, beliefs and experiences of which the individual may not even be aware of. According to Wenger (1998), online communities can be seen as fitting instruments and "places" to make such knowledge visible and available for others. Following this line of arguments, one can claim that social information behavior is of special value for information seeking as it allows access to knowledge usually not available in documents or databases.

Summing up, we argue an analytical framework of social information behavior that expands existing views on information seeking with concepts originating from Social Learning and Knowledge Management. In addition, we determine a set of basic components, processes and attributes of knowledge developments which could be used as predictors to assess specificities of collaborative knowledge development. The following figure illustrates our conceptual framework.

- Figure 1 Conceptual framework -



5. RESEARCH QUESTIONS

The goal of this study is to get insights into the structure, pragmatics, and success of knowledge creating processes within online communities with a topical focus on the German educational domain. As lined out in chapter 2, the perspective is on the self-determined usage of social media for information seeking and knowledge building purposes. If we take into account the pragmatics of both active (participative) and passive (receptive) social information behavior, success of the questioner with regard to the expressed information need and the knowledge building value for communicants and readers are of special interest (cf.chapter 4).

Our research questions are framed and structured along two basic categories

- Overview of discourse
- Factors of success and knowledge generation

First, we determine knowledge related characteristics with regard to the initiation, the course of events and the outcome of the discussions. Following that, we explore possible relations between these characteristics.

5.1 Overview of discourse

Our first research interest is on the information needs expressed in the initial posts. We want to get insights into the type of knowledge users aim for, e.g. if users rather aim for factual knowledge or/and additionally personal, subjective views on their question(s). Furthermore, we are interested in the pragmatics of communication. What are the goals of users? Do they aim for rather simple fact oriented answers or do they also seek concrete action related suggestions or even reach out for emotional support (Newman et al., 2011). We summarize these aspects in research question 1: *Knowledge types and intents of questions.*

The knowledge value of the forums will also be investigated from two perspectives: a) success for the questioner and b) knowledge building within the community. The significance of forums for information seeking can be defined as the grade of satisfaction of the questioner with regard to the information need (a). We refer to this as research question 2: *Success of communication*.

In contrast, the knowledge value of collaborative knowledge building for all participants and even passive forum recipients (b) is not that apparent. As laid out in chapter 4, we connect this to socio-cultural, socio-genetic and cognitive elaboration perspectives of communication and knowledge building. Of course, any kind of however specified knowledge value for the other participants or the wider audience of the forum discussions is dependent on the interests and contexts of the particular individuals and cannot be generalized. However, in reference to Clark et al. (2007), we try to measure the nature and function of contributions within the dialog in order to attain a conceptual overview with regard to externalization and other discourse activities and resulting characteristics of consensus building within these forums. Although not (yet) included in our analytical framework, we also want to check for *socio-emotional characteristics of communication* as the literature review indicates that emotional aspects could also be of importance. This leads to research question 3: *Attributes of the course of the discussion*.

5.2 Factors of success and knowledge generation

Research questions 1-3 give an overview of the specific attributes of the online discourse in the analyzed forums. Reaching beyond that descriptive level, we explore possible relations between these attributes and also statistical properties of discussion threads (e.g. number of authors). The research interest here is on *predictors of success* (research question 4) and on *predictors of the generation of new knowledge* (research question 5).

6. APPROACH AND METHODS

To present our research design, we first describe our procedure of sample selection. Following that, we delineate our approach of content analysis and illustrate the development of the coding scheme and coding procedure.

6.1 Sample selection

The sample selection followed a three step procedure.

First, relevant domains were identified. We used a set of 45 terms and phrases to query Google. The query set consisted of the most popular internal and external queries of a popular specialized information provider in the field, as well as of labels of education related study paths. For each query, the top 100 search results were examined in order to determine if they were suitable for further analysis. Websites were selected if the topics of these sites (or a subsection of the sites) matched topics of education related study paths and contained active discussions which related to educational topics as well as to the field of study.

Second, three selected forums (Lehrerforen.de, Paedagogik-klick.de and Referendar.de) were crawled. We chose these forums for further analysis due to their topical focuses and target groups. At the moment of data collection (February-

March 2013), Lehrerforen.de had about 15,500 members and contained 304,700 posts in 35,000 threads. The main target group of Lehrerforen.de consists of students in education related study paths as well as trainees and teachers. Discussions in the community focus on topics such as teaching degrees and examinations, as well as lesson planning, and didactic or educational counseling. The forum Referendar.de aims at a narrower target group of students in education related study paths, trainees and young teachers. Accordingly, topics in the forum mainly encompass questions, discussions and experience exchange about traineeships in the teaching environment. At the time of data collection, the forum contained of 21,000 members and 275,700 posts in 27,000 threads. Compared to Lehrerforen.de and Referendar.de, Paedagogik-klick.de is a small forum (about 2,000 members and 72,100 posts in 4,700 threads at the moment of data collection) which aims at a broad target group including teachers, kindergarten nurses, educators, students and trainees as well as parents. The topics are primarily based on education and learning, traineeship and career, as well as on expertise discussions.

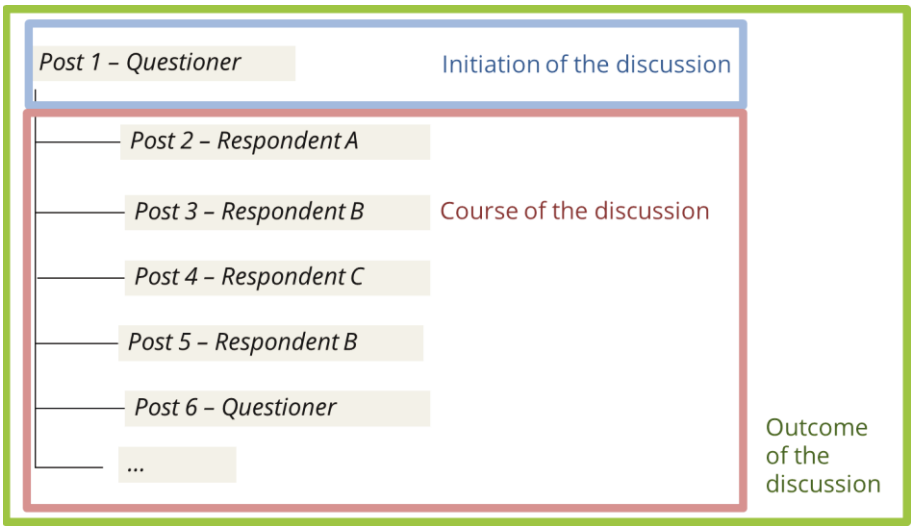
The crawler was configured with Java and Perl scripts on a Linux system using the Open Source software components Cassandra, Nutch, HTTrack, MySQL, and Nginx. Only selected sub-forums were indexed. Inclusion criteria were based on an explicit study or student-related focus of the sub-forums. [anonymized for peer review] give a more detailed description of the identification of websites and the crawler configuration and crawling procedure.

In a third step, we developed the coding scheme and performed a content analysis of a selected sub sample. We tested and refined the coding scheme on six selected threads with a total of 60 posts collected from three forums. Two threads were selected from Lehrerforen.de, one thread from Referendar.de and three threads from Studis-online.de. Studis-online.de was chosen because it deviates from the other forums. Studis-online.de is not moderated and participation does not require prior registration. For the final manual content analysis, 55 threads containing 533 posts were randomly selected from Lehrerforen.de, Paedagogik-klick.de and Referendar.de. Only threads with at least five contributions were considered.

6.2 Content Analysis

As a whole, our content analysis approach resembles a structuring procedure (deductive category application) according to Mayring (2010). Single contributions were defined as elementary units of analysis. According to our research interest, *initiation, course, and outcome of discussions* were determined as dimensions of structuring. Figure 2 illustrates this analytical approach.

- Figure 2 Dimensions of structuring -



The coding scheme was developed iteratively by two researchers in four cycles. During the cycles, the predefined categorization scheme was tested by independently coding the test sample and continuously redefined. The coding scheme relies on different theoretical considerations (chapter 4). Features are focused on cognitive activities and types of knowledge to capture the nature and function of contributions (Clark et al., 2007)

In the following, we explain our coding scheme with regard to

- knowledge types of questions and answers,
- intents of questions,
- characteristics of knowledge building,
- outcome of the discussions.

The developed categories are written in italics.

6.2.1 Knowledge types of questions and answers

With regard to knowledge types of questions and answers we oriented on the distinction or continuum of explicit and tacit knowledge. We differentiated between types of knowledge as fact oriented *objective knowledge* questions (e.g. facts or data), and questions aiming for *personal estimations* (e.g. opinions or personal experiences). The last one can be connected to tacit knowledge.

6.2.2 Intents of questions

The pragmatic aim of the questioner is the second important aspect concerning the initiation of the discussion. Regarding intents, we relied on Cvijikj's and Michahelles's (2011) classification of intents of Facebook posts and adapted the scheme to our object of investigation. Questions with a content related intent were categorized as *aim for topical information* (e.g. *"Can I interrupt my traineeship and continue at a later time?"*). Furthermore, many questions contained information on the personal context of the questioner. This indicates that communication pragmatics reach beyond the content related level of information needs. Those were categorized as *aim for suggestions* (e.g. *"I would appreciate if you had any suggestions..."*), *uncertainty reduction* (e.g. *"I'm interested to know what you think about..."*, *"I'm interested in your views on taking a break"*) and *aim for emotional support* (e.g. "*I failed my exam*."). Comparable classifications are found in studies that use concepts of social support (e.g. informational support, emotional support) as their bases for classification (e.g. Savolainen 2010).

6.2.3 Characteristics of knowledge building

With regard to the course and development of the discussion, two levels had to be considered: a cognitive, and supplementally a socio-emotional level. Again, on the cognitive level we distinguished between factual knowledge contributions *(factual answer)* and knowledge containing tacit aspects *(opinion)* to measure the type of knowledge brought into discourse. To capture knowledge building characteristics, we categorized the nature and function of contributions within the discussions according to socio-cultural and socio-genetic perspectives on knowledge building. We also included cognitive elaboration and consensus building. To analyze the knowledge value of contributions seen from a socio-cultural perspective, the attribute *new topical aspect* was introduced to measure the amount of externalization of question related concepts. A contribution was categorized to comprise new topic-related knowledge, if it contained an aspect that brought something new into the discourse. We regard this category as central to capture the knowledge development process and the knowledge value of the online discourse. Below is an example for a *new topical aspect* that is provided by a factual answer (translated from German to English):

Question: *"Is there any chance that the costs for my health certificate may be reimbursed?"*

Answer #1: *[...]*

Answer #2: *"In Bavaria you get your money back. As soon as you have begun your traineeship you can send the certificate in and you get your money back."*

In contrast to answer #2, answer #1 stated that it would not be possible to get reimbursement for a health certificate. By giving information about the reimbursement process in a specific German state, answer #2 gives a new topical aspect.

New topical aspects could also be provided by opinion-related answers as the following example shows (translated from German to English):

Question: *"I have a suspicion that one of my students self-harms. How do I raise the topic with her?"*

Answers #1-8: *[...]*

Answer #9: *"To what extent do you have to address those issues pedagogically? I think help should not be forced upon someone. My experience with those people is that they have not waited for a stranger (and that's what a teacher is) to interfere. I really doubt that interference will be successful."*

The previous answers (#1-8) were mainly concerned with *giving* advice on how to interfere. In answer #9 someone externalizes their personal experience with an insight that questions the previous focus of the discussion.

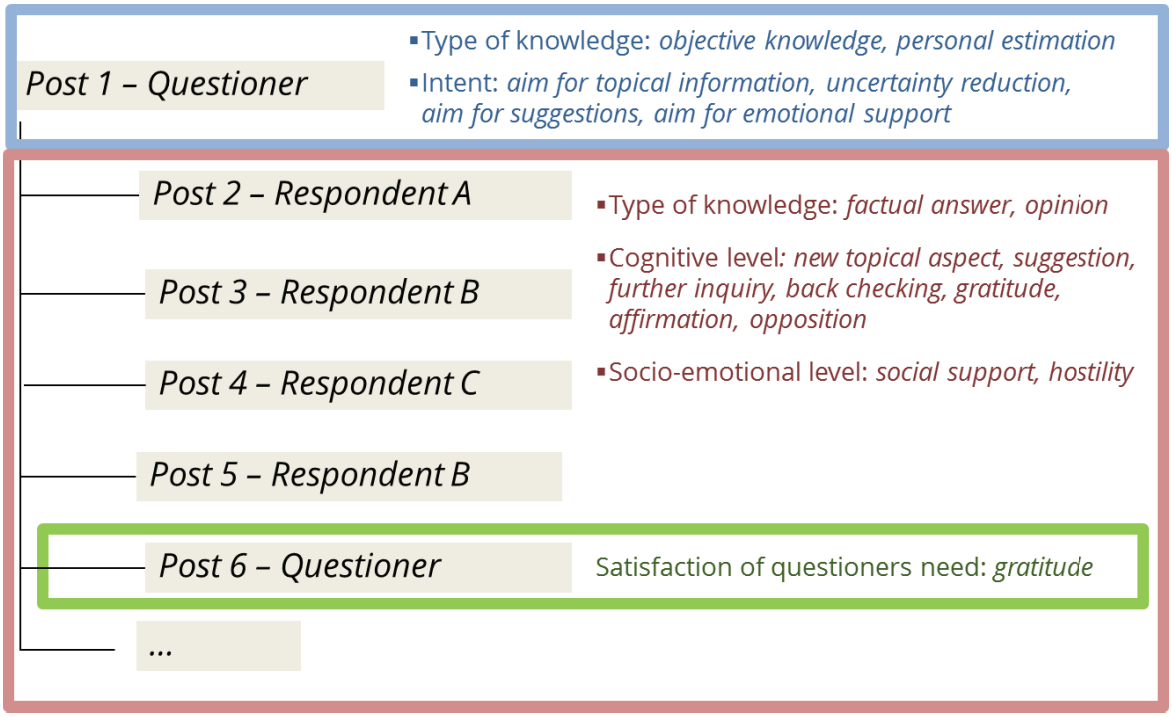
To measure cognitive support and conflicts, we employed the categories *affirmation* and *opposition.* Cognitive elaboration was captured with the categories *further inquiry* and *back checking. Suggestion*s provide content-related conclusions with regard to the pragmatical level of the questioners' intent. As such this category has a direct consensus building function within the discourse. Socio-emotional aspects were analyzed with the categories *social support* and *hostility*.

6.2.4 Outcome of the discussions

Concerning the outcome of a discussion, three possible types of knowledge values are described in figure 1 in the analytical framework: a) fulfillment of the information need, b) learning for the participants, c) knowledge provision for the wider web audience. Concerning a), the success in relation to the initial information need had to be measured. This aspect could be operationalized if the initial questioner contributed *gratitude,* mentioning that the information need was at least partly resolved. In such a case the corresponding thread was categorized as successful. As we cannot expect that questioners express their gratitude in all cases of success, we could only capture a fraction of successful fulfillment. B) and c) were both combined as community oriented knowledge building. Referring to this, our approach is even more limited. The knowledge building value of the discussion for the participants or the wider reader audience could not be measured directly and is rather object of the interpretation of the knowledge attributes and quality of the discourse as a whole. Nevertheless as argued in chapter 4, we are able to analyze the occurrence of discourse activities. In the specific context of this investigation, we are primarily concerned with the amount of externalization. In addition, we are able to relate success and amount of externalization to statistical and proportional data of our sample.

The following illustration shows the structure and features of the final coding scheme.

-Figure 3 Coding scheme-



It is important to note that the categories are feature like attributes. E.g. a question could be fact oriented and personal estimation oriented at the same time. Therefore, "Type of knowledge", "Intent", and "Cognitive level" categories are not mutually exclusive but treated as feature-like attributes.

6.3 Coding procedure

The sample set of the investigation (55 threads containing 533 posts) was coded by two researchers, at first independently and following that deviations were jointly discussed and resolved. In this way inter-subjectivity of the manual analysis was secured. The coding was executed from November 2013 till January 2014.

7. Analysis

We structure the analysis according to our research questions as laid out in section 4. First, we start with a presentation of the data set. Table 1 gives an overview of the sample.

	Lehrerforen.de	Paedagogik-Klick.de	Referendar.de	Sum/Mean
Posts (sum)	182	187	164	533
Threads (sum)	19	18	18	55
Authors (sum)	90	39	61	191
Thread length (mean in posts without initial post)	8.63	9.39	8.11	8.71
Authors per thread (mean)	5.79	4.22	4.72	4.93

Table 1 Overview of sample

The sample data reveals that the average length of threads and the number of authors per thread are basically similar in the three forums. The lower number of different authors in Paedagogik-Klick.de indicates a higher involvement per author on this domain. This is confirmed by the forum crawl data which shows the following means for posts per user in the three forums: Lehrer-foren.de 11.9, Paedagogik-Klick.de 20.3, Referendar.de 6.7. As a whole, we estimate the three forums as comparable and therefore treat the discussions in the three forums as one data set.

In the following, we first present an overview of discourse. After that, factors of success are analyzed.

7.1 Overview of discourse

7.1.1 Knowledge types and intents of questions (RQ1)

Knowledge types: With regard to type of knowledge asked for, our results show there is nearly always a demand for *objective knowledge.* Over 90% of all thread initiations ask for such explicit fact oriented information. Nevertheless, nearly half of the thread initiations also aim for *personal estimations*, meaning that opinions and personal experiences are also of importance. These results are somewhat contrasting to Savolainen (2011b) and rather in accordance with data from Chuang & Yang (2014). However, this is not an either/or relation. In fact, a demand for personal estimations is most often accompanied by a demand for objective knowledge too. One example is: (translated from German to English) *"… just wanted to ask if anyone of you has experience in this direction and maybe knows therapists who have told their own ways of dealing with this daily mental stress…"*. Such questions requesting both knowledge types made up a share of one third of all thread initiations. In sum, we get a clear order of knowledge types asked for. The majority of questions is asking exclusively for *objective knowledge* (52%), 38% of thread initiations aim for both knowledge types, whereas only a minor fraction of 7% of all initiations reach out solely for *personal estimations*.[1] If we relate this result to the discussion of different knowledge types and their values, we can conclude that these forums indeed are used for needs that could not be fully addressed by typical information systems (e.g. search engines) that rely on existing explicit knowledge. However, even in forums explicit knowledge is the knowledge type users nearly always aim for.

Intents of questions: What about the pragmatic goal of the users? They usually aim for topical information (87% of all thread initiations). 40% of all thread initiations are categorized as *aim*ing *for suggestion*s. *Uncertainty reduction* is also visible to a substantial extend (33%). *Emotional support* is sparsely seen (5%) and seems to play only a subordinated role.

As a whole, nearly every initiation of discussion aims for topical information on a factual answer level. However, these are not the only goals of the users. In roughly two third of the cases we were able to identify additional intents, mainly *aim for suggestion* and *uncertainty reduction.* Here is one example of an information need that reaches beyond a factual answer level. *"I am a teacher student for elementary school in Bavaria and about to begin with practical training in about two weeks. I'm really looking forward to it. I'll teach a 4th grade class. Now, I think about how to introduce myself in the best way, on the first day, but I do not have much experience with regard to this. I would be happy if I could get a few suggestions here!"* (translated from German to English) (intents: *aim for suggestions*, *uncertainty reduction*). If we refer this data to assessments of the role of forums in information seeking as described in section 3.2, we can presume that forums are chosen not only to learn about subject areas but also to directly gain actionable insights or evaluate informational contexts. Therefore, in our investigation we can regard forums as tools used to accomplish higher levels of learning according to Bloom's taxonomy of cognitive learning goals (Krathwohl, 2002

7.1.2 Success of communication (RQ2)

A thread was defined as successful if the questioner explicitly mentioned that at least a part of the problem was solved. *"Thanks for the link! That's exactly what I need!"* (translated from German to English) is an example of such *gratitude*.

[1] Note: Percentages do not sum up to 100% because for one thread initiation the demanded knowledge type could not be categorized.

According to this operationalization, 40% of all threads were successful. With nearly half of the problems being explicitly solved, one can assess these forums as viable tools for an asking-to-learn information seeking approach as stated above. It seems that the success rate is somewhat dependent on the knowledge type aimed for and the intents of questions. With regard to type of knowledge aimed for we get a clear ranking. 55% of the questions asking exclusively for *objective knowledge* are solved. Information needs aiming for both knowledge types have a much lower success rate of 24%. Needs solely aiming for *personal estimations* are on the same level as 25% of them are solved.

Concerning intent, we get the following results. Questions focusing only on the satisfaction of information needs on a factual level of the topics seem to have a higher success rate (57%) than questions also concerned with getting concrete suggestions (32%) or aiming for uncertainty reduction (28%). In sum, questions aiming for implicit knowledge seem to be more difficult to solve. The same is true with regard to pragmatical intents. More complex information needs seem to be harder to satisfy.

7.1.3 Attributes of the course of the discussion (RQ3)

Success seen from a perspective of community oriented knowledge generation is primarily dependent on the cognitive characteristics of the discourse. Table 2 gives a survey of the proportion of the amount of knowledge types and cognitive categories of all postings.

Category	Mean average fraction
Factual answer	49%
Opinion	42%
Suggestion	28%
Back checking	22%
Further inquiry	12%
Affirmation	5%
Opposition	11%
New topical aspect	49%

Table 2 Mean average fraction of knowledge types and cognitive post categories types over all threads (categories are not mutually exclusive)

Every second post contains factual information and/or opinions. That means both types of knowledge are frequently brought into the discussion. In addition, half of the posts provide new topical aspects. Therefore, the discourse in the three forums can be assessed as knowledge generation. One fourth of the contributions provide concrete suggestions for problem solving. Cognitive conflicts occur in half of the threads, but since the table shows a rather low overall fraction of such post types within the threads, one can assess this content type as of subordinated significance. In sum, knowledge generation in these forums basically resembles socio-cultural perspectives of knowledge generation, diminishing knowledge asymmetries by bringing together different knowledge “pieces”. Socio-genetic perspectives of knowledge building were also observed but can be assessed as of lower significance. Therefore we can argue that knowledge creation in the forums follows primarily *integration-oriented consensus building* strategies. With regard to socio-emotional characteristics, discussions can be categorized as sometimes supportive and seldom hostile. Social support is visible in 33% of all threads. Hostility is sparsely seen (5% of all threads).

In table 3 we see occurrences of knowledge types and cognitive post categories of the discussion in dependence on the type of knowledge aimed for on the part of the questioner. Discussions following questions that exclusively aim for

objective knowledge, got a much lower proportion of opinions than discussions that aim for both types of knowledge. Thus, data indicates that the type of knowledge provided roughly corresponds to the type of knowledge asked for.

With regard to cognitive activities there are also differences. The proportion of *suggestions* is higher on needs that include or rely only on *personal estimations*. On *back checking* and *further inquiry*, the picture is not that clear. Furthermore, it seems that questions aiming for personal estimations invoke a much higher grade of socio-genetic activities, at least if we sum up *affirmation* and *opposition.* In contrast to that, differences in the proportion of externalization of *new topical knowledge* are rather low. Hence, although the specificities of discourse are different, the amount of knowledge generation is robust in relation to type of knowledge aimed for.

Type of knowledge / Answer category	Exclusive aim for objective knowledge (n=29)	Aim for objective knowledge and personal estimation (n=21)	Exclusive aim for personal estimation (n=4)
Factual answer	54%	46%	34%
Opinion	31%	56%	66%
Suggestion	21%	36%	43%
Back checking	24%	21%	15%
Further inquiry	12%	14%	10%
Affirmation	6%	2%	13%
Opposition	6%	17%	13%
New topical aspect	48%	53%	51%

Table 3 Mean average fraction of knowledge types and cognitive categories in dependence on type of knowledge aimed for (unless otherwise stated categories are not mutually exclusive)

A segmentation with regard to question intent results in data as presented in table 4. The data is somewhat similar to those in table 3. The proportion of *opinion*s increases with the pragmatic complexity of the information need. The fraction of *factual answer*s decreases. Again, differences in the proportion of cognitive categories are visible. Here, the picture is not so clear. There are deviations with regard to nearly all listed discourse types. The fraction of *suggestion*s rises with the pragmatic complexity of the intent. *Back checking* is stable and *further inquiry* is slightly rising too. Again, we see differences with regard to socio-genetic discourse activities. This time, however, *aim for suggestions* is sticking up, whereas *uncertainty reduction* is somewhere between (exclusively) *aim for topical information* and *aim for suggestions.* Again, differences with regard to the externalization of new topical knowledge are rather low.

Intent / Answer category	Exclusive aim for topical information (n=21)	Aim for suggestions (n=22)	Uncertainty reduction (n=18)
Factual answer	53%	49%	40%
Opinion	31%	46%	61%
Suggestion	21%	36%	33%
Back checking	21%	23%	21%
Further inquiry	10%	15%	16%
Affirmation	7%	3%	4%
Opposition	4%	19%	9%
New topical aspect	47%	54%	53%

Table 4: Mean average fraction of knowledge types and cognitive categories in dependence on intent (unless otherwise stated categories are not mutually exclusive)

Summing up these results, we can conclude that the proportion of types of knowledge brought into discourse is dependent on the type of knowledge aimed for as well as on the intent of the questioner. Nevertheless, although there are differences with regard to the proportion of cognitive categories, the amount of new topical knowledge remains remarkably stable.

7.2 Factors of success and knowledge generation

7.2.1 Predictors of success (RQ4)

To explore possible success factors, statistical correlations and regression analyses were conducted. First, possible relations between thread statistics (number of authors, number of posts, length of initiating questions in words) and success *(gratitude)* were checked. With regard to this simple discourse statistics, there is a negative correlation between the number of authors and success (Spearman -.326*, two sided). This means, the higher the number of participants, the smaller the probability of success. This result is somewhat contradictory to the idea of socio-cultural knowledge building, as one would expect that the larger the number of participants the more comprehensive and versatile the knowledge base. However, this result is somewhat similar to the results of investigations on Wikipedia of Kane (2011) and Liu & Ram (2011) who also state that the sheer amount of contributors is not a criterion for a high quality of knowledge generation. Further segmentation according to type of knowledge and intent (as employed in the chapter above) delivered no significant results.

Second, we tested relations between the types of knowledge *(factual answer, opinion)* and cognitive content types with success. A correlation analysis shows only two significant correlations, a negative correlation between the fraction of *opinion*s within a thread and success (Spearman -.346**, two sided) and a negative correlation between the amount of *opposition* and success (Spearman -.409**, two sided). In addition, binary logistical analysis with regard to type of knowledge and cognitive categories confirms these results.

Again, this result, too, is somewhat opposite to our idea of the value of collaborative knowledge building for information seeking. As a whole, there are no statistically significant positive relations with regard to type of knowledge or cognitive post categories. In contrast, the volume of *opinions* is rather disadvantageous for the success of the questioner as are *oppositions*. Success seems to be more dependent on the provision of the "right" factual answer. One has to keep in

mind that our operationalization of success is an approximation, only measuring explicitly stated success *(gratitude)*. For that reason, data here needs to be interpreted very cautiously.

7.2.2 Predictors of the generation of new knowledge (RQ5)

With regard to knowledge generation, we determine the amount of externalization as the primary success factor, as we were foremost interested in the generation of new knowledge and less in the depth of argumentation seen from a socio-genetic perspective. According to our categorization scheme, the category *new topical aspect* is the variable that measures the increase of knowledge in the course of events of online discussions.

Data shows correlations (Pearson) between *new topical knowledge* and the knowledge types. According to that, 30% (Pearson, .297** two sided) of all posts categorized as *opinions* are also categorized as containing *new topical knowledge*. However, the correlation between *factual information* and *new topical aspect* is by far stronger (Pearson, .617** two sided), indicating that factual knowledge is much more important for knowledge generation than personal views containing tacit knowledge.

A regression analysis supports this argument. The following table shows that new knowledge is primarily generated by *factual information*.

	Non standardized Coefficients		Standardized Coefficient		
	Regression Coefficient B	Std. Error	Beta	T	Sig.
constant	.129	.027		4.780	.000
Factual answer	.581	.036	.580	16.010	.000
Opinion	.168	.03	.168	4.631	.000

Table 5 Regression analysis of knowledge generation

Again, our analysis denotes the inferior value of *opinions* in relation to *factual answers* with regard to the knowledge value of the forums. This result deviates from the argumentation of Savolainen (2011b) who states that knowledge provision relies primarily on personal knowledge. Probably, the specificities of discourse are dependent on the domains observed. Health related information behavior may differ from knowledge building in the educational profession. The first may be more personal, the last more factual oriented.

Finally, we were interested if there are correlations between *new topical knowledge* and other cognitive categories. Here, data reveals correlations with *suggestion*s (Pearson, .427** two sided), *back checking* (Pearson, .145** two sided) and *further inquiry* (Pearson, .195** two sided). There were no correlations with posts categorized as *affirmations* and *oppositions*. It seems, socio-cultural and socio-genetic activities are rather unconnected.

8. Discussion

In the discussion, we first summarize the results of the investigation. Then, we reflect on our research approach, its limits and value for further research.

In this investigation, we analyzed knowledge generation in three forums concerned with professional aspects of teaching and studying. The goal was to get insights of the knowledge value of these forums seen from two perspectives. First, satisfaction of the information need of the initial questioner, second, community oriented knowledge building. The approach was structured with research questions that compile an overview of discourse, factors of success and knowledge generation.

Starting with *types of knowledge asked for and intents of questions* (research question 1), our investigation shows that information needs are often human centric. Questions nearly always aim for objective knowledge, but in half of the

cases they also reach for personal estimations. The questioner nearly always wants to acquire topical knowledge about the specific subject. However, very often discourse initiations also aim for actionable insights or uncertainty reduction. Therefore, information seeking in forums seems to be much more multifaceted than the search for and receipt of topically relevant results. Evaluation and application of knowledge are important dimensions. Therefore, this research adds to results of Savolainen (2011b) and Chung & Yang (2014) denoting that forums are tools or places used for higher level exploratory search in an asking-to-learn approach. But are forums viable tools for such high level information seeking? This leads us to research question 2 *success of communication.* Seen from the perspective of the thread initiator, communication in forums is successful in nearly half of the cases. If we segment the results into the needs solely focusing on factual answers on a topical level and the needs also aiming for suggestions and uncertainty reduction, then we see that the majority of the pragmatically more "simple" needs get solved, whereas the more complex ones got a much lower success rate. Our results here are limited because the category *gratitude* is a dichotomous trait and needs to be explicitly stated. In addition, we did not capture the topical difficulty of questions in our coding scheme. Nevertheless, with nearly one third of needs aiming for suggestions (32%) and/or uncertainty reduction (28%) being solved, we assess these forums as very worthwhile information seeking places even for needs focused on the evaluation of information or applicable knowledge. Furthermore, taking a knowledge generation perspective, the categorization of *attributes of the course of the discussions* (research question 3) discloses that these forums can also be assessed as knowledge communities since half of the posts provide *new topical aspect*s. Discourse corresponds primarily to socio-cultural perspectives of knowledge generation. In contrast, socio-emotional characteristics of communication play a rather subordinated role. In addition, we see that the amount of implicit and explicit knowledge in the discussions relates to the type of knowledge aimed for as well as to the intent of the questioner's need. Discussions to needs with a more complex pragmatic evoke a higher proportion of tacit knowledge. Furthermore, we discovered partly different patterns of cognitive activities. In contrast to that, the amount of knowledge generation (measured with the category *new topical aspect*) remains remarkably stable above all analyzed type of knowledge and intent related segments. Therefore, with regard to consensus building we get a clear picture. Knowledge building follows primarily *integration-oriented consensus building* strategies.

Asking for *factors of success* (research question 4), data shows evidence that is somewhat contradictory to the basic concepts of socio-cultural knowledge building. A higher potential diversity of viewpoints and knowledge is not helpful to solve information needs. In contrast, the more authors participate in a discussion, the smaller the success rate. In addition, the volume of *opinion*s and *opposition*s are rather negative predictors for the success of the questioner. One plausible explanation is that success is more dependent on the provision of the "right" factual answer(s) and not on the discourse itself. From a community perspective, the generation of new knowledge is of utmost importance for knowledge related benefits. Regarding the *predictors for the generation of new knowledge* (research question 5), data indicates that factual knowledge is the most worthwhile knowledge type as it is the best predictor for knowledge generation. Both results question the value of *opinion*s, which we argued as valuable tacit knowledge in reliance on Nonaka & Krogh (2009). Here, we have to keep in mind the limits of our method and keep in mind that the type of question may also be a decisive factor. *Opinion*s have a higher probability to appear in questions aiming for personal estimations. These questions are not that clearly solved as questions aiming for facts.

In sum, by relying on a conceptual framework that combines socio-cultural, socio-genetic and cognitive elaboration perspectives of knowledge building with concepts of tacit and explicit knowledge and by employing a categorization scheme that analyzes the nature and function of contributions, we uncover a complex picture of knowledge generation in forums. We show that these forums have a special role in information seeking, since they address higher levels of exploratory search. Pragmatically complex needs are often solved, indicating that the forums are of high knowledge value for information seeking. With regard to the community oriented perspective, we observe a high grade of knowledge generation. In addition, there are many different perspectives involved in the discussions and knowledge

creation follows primarily socio-cultural perspectives. Opinions seem to be of much lesser knowledge value than factual answers. Although this last result is somewhat unexpected, in sum, seen from the community perspective too, we can assign a high knowledge value to these forums.

As a whole, our research concept is not the end of the discussion but just one of many possible approaches. Seen from an epistemological perspective, it can be assessed as a first draft. Nevertheless, as argued in chapter 4, to our knowledge, our investigation is the first one that combines perspectives of community oriented knowledge generation and information seeking. We think our analytical framework provides a solid conceptual base for the design of further investigations in the field. Nevertheless, we also see a need to redefine and refine our categorization scheme. This is partly caused by the limits of our approach as mentioned above and partly due to the coding category *new topical knowledge*. On the one hand, this category makes it possible to measure knowledge gains in dependence on the state of the discussion in the first place. Accordingly, taking forum analysis forward, this allows for deeper insights than approaches which employ rather static quality criteria that do not reflect the sequence of knowledge generation. Here, we see the central news value of our method. On the other hand, our understanding of knowledge development in the course of the discussions is just a first step. The category *new topical knowledge* is of nominal value, meaning that it is not able to compare or order different kinds and values of knowledge gains in the course of discussion. Thus at current, our understanding of the quality of knowledge processes remains limited. Future investigations should focus on this point. Maybe, extracting the *new topical aspects* from the single posts and then aggregating them in a "knowledge summarization" could serve as a starting point to get a much richer and comparable picture of knowledge processes without losing the balance between accuracy and cost of analysis. Additionally, for future studies we aim to develop analytical perspectives that focus on structures that develop during the sequence of online discussions.

Finally, there is the question if the results are transferable to other forums or communities. As written, we believe that the specificities of discourses are dependent on the domains observed. Our research sample covered communities concerned with factual oriented occupation or education related aspects of users' professional development. We think that results may be transferable to communities focusing on other professional areas, too. However, this still needs to be empirically proven.

9. REFERENCES

Agichtein, E., Castillo, C., Donato, D., Gionis, A. and Mishne, G. (2008), "Finding high-quality content in social media", in *Proceedings of the International Conference on Web Search and Web Data Mining in Palo Alto, CA, 2008*, WSDM '08, ACM, New York, NY, pp. 183-194.

Arazi, O. and Kopak, R. (2011), "On the measurability of Information Quality", *Journal of the American Society of Information Science and Technology,* Vol. 62 No. 1, pp. 89-99.

Aschoff, F.-R., Schaer, V. and Schwabe, G. (2011), "Where should I send my post?: the concept of discourse quality in online forums and its dependency on membership size", in *Proceedings of the 5th International Conference on Communities and Technologies in Brisbane, Australia, 2011*, ACM, New York, NY, pp. 69-78.

Bickart, B. and Schindler, R. M. (2001), "Internet forums as influential sources of consumer information", *Journal of Interactive Marketing*, Vol. 15, pp. 31-40.

Burnett, G. and Buerkle, H. (2004), "Information Exchange in Virtual Communities: A Comparative Study", *J. Comput.-Mediat. Comm,* Vol. 9 No. 2.

Case, D. O. (Ed.) (2012), *Looking for information: A survey of research on information seeking, needs and behavior.* Emerald Group Publishing, Bingley, UK.

Chuang, K. Y. & Yang, C. C. (2014), "Informational support exchanges using different computer-mediated communication formats in a social media alcoholism community", *Journal of the Association for Information Science and Technology*, Vol. 65, pp. 37-52.

Clark, D. B., Sampson, V., Weinberger, A. and Erkens, G. (2007), „Evaluating the quality of dialogical argumentation in CSCL: Moving beyond an analysis of formal structure", in *Proceedings of the 7th International Conference on Computer Supported Collaborative Learning in New Brunswick, NJ, 2007*, CSCL'07, pp. 13-22.

Cvijikj, I. P. and Michahelles, F. (2011), "Understanding social media marketing: a case study on topics, categories and sentiment on a Facebook brand page", in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments in Tampere, Finland, 2011*, MindTrek '11. ACM, New York, NY, pp. 175-182.

De Lisi, R. and Golbeck, S. (2005), "Implications of Piagetian Theory for Peer Learning", in O'Donnell A. M. and King, A. (Ed.), *Cognitive Perspectives on Peer Learning*, Erlbaum, Mahwah, NJ, pp. 3-37.

deVries, E., Lund, K. and Baker, M. (2002), "Computer-mediated epistemic dialogue: Explanation and argumentation as vehicles for understanding scientific notions", *J. Learn. Sci.,* Vol. 11 No. 1, pp. 63-103.

DiMauro, V. (2012), "Nearly 80% of People Participate In Online Community to Help Others", available at: http://blog.leadernetworks.com/2012/07/nearly-80-of-people-participate-in.html (accessed November 3 2014).

Duschl, R. (2007), "Quality Argumentation and Epistemic Criteria", in Erduran S. and Jiménez-Aleixandre M. (Ed.), *Argumentation in Science Education*, Springer Netherlands, Florida State University, pp. 159-175.

Evans, B. M. and Chi, E. H. (2008), "Towards a Model of Understanding Social Search", in *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, ACM, New York, pp. 485-494.

Gibs, J. (2009), "Social Media: The Next Great Gateway for Content Discovery?", available at: http://www.nielsen.com/us/en/newswire/2009/social-media-the-next-great-gateway-for-content-discovery.html (accessed 15 April 2014).

Hasler, L., Ruthven, I. and Buchanan, S. (2014), "Using internet groups in situations of information poverty: Topics and information needs", *Journal of the Association for Information Science and Technology*, Vol. 65, pp. 25–36.

Johnson, D. W. ; Johnson, R. T. and Holubec, E.J. (1998), *Cooperation in the classroom.* Interaction Book, Edina.

Kane, G. C. (2011), "A multimethod study of information quality in wiki collaboration", *ACM Trans. Manag. Inf. Syst.* Vol. 2 No. 1, Article 4.

Kim, K., Yoo-Lee, E. and Sin, S. (2011), "Social Media as Information Source: Undergraduates' Use and Evaluation Behavior", in *Proceedings of the American Society for Information Science and Technology,* Vol. 48 No. 1, pp. 1-3.

Kleimann, B., Özkilic, M. and Göcks, M. (2008), *Studieren im Web 2.0. Studienbezogene Web- und E-Learning-Dienste*, HISBUS-Kurzinformation 21. HIS Hochschul-Informations-System GmbH, Hannover, Germany.

Krathwohl, D.R. (2002), “A Revision of Bloom's Taxonomy: An Overview”, in *Theory into practice*, Vol. 41 No. 4, pp. 212-218.

Kuhn, D. and Udell, W. (2003), “The development of argument skills”, *Child Dev*. Vol. 74 No. 5, pp. 1245-1260.

Liu, J. and Ram, S. (2011), “Who does what: Collaboration patterns in the Wikipedia and their impact on article quality”, *ACM Trans. Manag. Inf. Syst*., Vol. 2 No. 2, Article 11.

Marchionini, G. (2006), “Exploratory Search: From Finding to Understanding”, in *Commun*., ACM, Vol.49, pp. 41-46.

Mayring, P. (2010), *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. 11. Aufl., Beltz, Weinheim.

Newman, M.W., Lauterbach, D., Munson, S.A., Resnick, P. and Morris, M.E. (2011), “It’s not that I don’t have problems, I’m just not putting them on Facebook: challenges and opportunities in using online social networks for health”, in *CSCW '11 Proceedings of the ACM 2011 conference on Computer supported cooperative work,* ACM, New York, pp. 341-350.

Nonaka, I. and Takeuchi, H. (1995), *The knowledge-creating company: How Japanese companies create the dynamics of innovation*, Oxford university press, New York.

Nonaka, I. and von Krogh, G. (2009), „Perspective—Tacit Knowledge and Knowledge Conversion: Controversy and Advancement in Organizational Knowledge Creation Theory”, *Organ. Sci.,* Vol. 20 No. 3, pp. 635-652.

O’Donnell, A. M. (2006), “The role of peers and group learning”, in Alexander, P. A. and Winne, P. H. (Ed.), *Handbook of educational psychology*, 2nd edition, Lawrence Erlbaum Associates, pp. 781-802.

Piaget, J. (1979), *Judgment and reasoning in the child*. Routledge, London.

Polanyi, M. (1958), *Personal Knowledge. Towards a Post-Critical Philosophy*. Routledge, London.

Robson, A.; Robinson, L. (2013), “Building on models of information behaviour: linking information seeking and communication”, *Journal of Documentation*, Vol. 69 No. 2, pp. 169-193.

Riemer, K. and Scifleet, P. (2012). “Enterprise social networking in knowledge-intensive work practices: A case study in a professional service firm” *CIS 2012: Location, location, location: Proceedings of the 23rd Australasian Conference on Information Systems* 2012 , ACIS, pp.1-12.

Savolainen, R. (2010), “Dietary blogs as sites of informational and emotional support”, *Information Research,* Vol. 15 No. 4, Paper 438, available at: http://InformationR.net/ir/15-4/paper438.html (accessed November 07 2014).

Savolainen, R. (2011a), “Judging the quality and credibility of information in Internet discussion forums”, *J. Am. Soc. Inform. Sci*., Vol*.* 62 No. 7, pp. 1243-1256.

Savolainen, R. (2011b), “Requesting and providing information in blogs and internet discussion forums”, *Journal of Documentation*, Vol. 67 No. 5, 2011, pp. 863-886.

Scardamalia, M. and Bereiter, C. (1994), “Computer support for knowledge-building communities”, *J. Learn. Sci*., Vol. 3 No. 3, pp. 265-283.

Shirky, C. (2008), *Here comes everybody*. Penguin Press, New York, NY.

Shirky, C. (2010), *Cognitive Surplus*. Penguin Press, New York, NY.

Slavin, R. (1995), *Cooperative learning. Theory, Research and Practice*. Allyn and Bacon, Boston.

Vygotsky, L. (1979), *Mind in Society. The development of higher psychological processes*. Harvard University Press, Cambridge, MA.

Vygotsky, L. S. (1986), *Thought and language*. Harvard MIT Press, Cambridge, MA.

Weinberger, A. (2003*), Scripts for Computer-Supported Collaborative Learning: Effects of social and epistemic cooperation scripts on collaborative knowledge construction*. Dissertation, LMU München: Faculty of Psychology and Educational Sciences

Weinberger, A. and Fischer, F. (2006), "A framework to analyze argumentative knowledge construction in computer supported collaborative learning", *Comput. Educ*., Vol. 46 No. 1, pp. 71-95.

Wenger, E. (1998), "Communities of Practice. Learning as a Social System". Systems Thinker, available at: www.co-i-l.com/coil/knowledge-garden/cop/lss.shtml (accessed November 3 2014).

White, R. W. and Roth, R. A. (2009), *Exploratory Search: Beyond the Query-Response Paradigm*. Morgan and Claypool , San Rafael, CA.

Wilson, T.D. (1999), "Models in information behaviour research", *Journal of Documentation*, Vol. 55 No. 3, pp. 249 – 270.