LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

FOSTERING AWARENESS AND COLLABORATION IN LARGE-CLASS LECTURES

— Principles and Evaluation of the Backchannel Backstage —

Dissertation

an der Fakultät für Mathematik, Informatik und Statistik der
Ludwig-Maximilians-Universität München

Alexander Pohl

30. Juli 2015

Eidesstattliche Versicherung

Hiermit erkläre ich an Eidesstatt, dass die Dissertation von mir selbstständig, ohne unerlaubte Beihilfe angefertigt ist.

Pohl, Alexander

Name, Vorname

München, 17.12.2015

Ort, Datum   Unterschrift Doktorand/in

Formular 3.2

ABSTRACT

For decades, higher education has been shaped by large-class lectures, which are characterized by large anonymous audiences. Well known issues of large-class lectures are a rather low degree of interactivity and a notable passivity of students, which are aggravated by the social environment created by large audiences. However, research indicates that an active involvement is indispensable for learning to be successful. Active partaking in lectures is thus often a goal of technology-supported lectures.

An outstanding feature of social media is certainly their capabilities of facilitating interactions in large groups of participants. Social media thus seem to be a suitable basis for technology-enhanced learning in large-class lectures. However, existing general-purpose social media are often accompanied by several shortcomings that are assumed to hinder their proper use in lectures. This thesis therefore deals with the conception of a social medium, called *Backstage*, specially tailored for use in large-class lectures.

Backstage provides both lecturer- as well as student-initiated communication by means of an Audience Response System and a backchannel. Audience Response Systems allow running quizzes in lectures, e.g., to assess knowledge, and can thus be seen as a technological support of question asking by the lecturer. These systems collect and aggregate the students' answers and report the results back to the audience in real-time. Audience Response Systems have shown to be a very effective means for sustaining lecture-relevant interactivity in lectures. Using a backchannel, students can initiate communication with peers or the lecturer. The backchannel is built upon microblogging, which has become a very popular communication medium in recent years. A key characteristic of microblogging is that messages are very concise, comprising only few words. The brief form of communication makes microblogging quite appealing for a backchannel in lectures.

A preliminary evaluation of a first prototype conducted at an early stage of the project, however, indicated that a conventional digital backchannel is prone to information overload. Even a relatively small group can quickly render the backchannel discourse incomprehensible. This incomprehensibility is rooted in a lack of interactional coherence, a rather low communication efficiency, a high information entropy, and a lack of connection between the backchannel and the frontchannel, i.e., the lecture's discourse.

This thesis investigates remedies to these issues. To this aim, lecture slides are integrated in the backchannel to structure and to provide context for the backchannel discourse. The backchannel communication is revised to realize a collaborative annotation of slides by typed backchannel posts. To reduce information entropy backchannel posts have to be assigned to predefined categories. To establish a connection with the frontchannel, backchannel posts have to be stuck on appropriate locations on slides. The lecture slides also improve communication efficiency by routing, which means that the backchannel can filter such that it only shows the posts belonging to the currently displayed slide. Further improvements and modifications, e.g., of the Audience

i

Response System, are described in this thesis.

This thesis also reports on an evaluation of *Backstage* in four courses. The outcomes are promising. Students welcomed the use of *Backstage*. *Backstage* not only succeeded in increasing interactivity but also contributed to social awareness, which is a prerequisite of active participation. Furthermore, the backchannel communication was highly lecture-relevant. As another important result, an additional study conducted in collaboration with educational scientists was able to show that students in *Backstage*-supported lectures used their mobile devices to a greater extent for lecture-relevant activities compared to students in conventional lectures, in which mobile devices were mostly used for lecture-unrelated activities.

To establish social control of the backchannel, this thesis investigates rating and ranking of backchannel posts. Furthermore, this thesis proposes a reputation system that aims at incentivizing desirable behavior in the backchannel. The reputation system is based on an eigenvector centrality similar to *Google*'s *PageRank*. It is highly customizable and also allows considering quiz performance in the computation of reputation. All these approaches, rating, ranking as well as reputation systems have proven to be very effective mechanisms of social control in general-purpose social media.

ZUSAMMENFASSUNG

Seit Jahrzenten wird die universitäre Lehre durch Massenvorlesungen, die sich durch sehr große anonyme Hörerschaften auszeichnen, geprägt. Wohlbekannte Probleme von Massenvorlesungen sind ein sehr niedriger Grad an Interaktivität als auch eine augeprägte Passivität von Studenten, die auch durch die sozialen Rahmenbedingungen in großen Hörerschaften begünstigt werden. Dabei ist bekannt, dass eine aktive Auseinandersetzung mit dem Lernstoff für ein erfolgreiches Lernen unabdingbar ist. Eine aktive Teilnahme in Vorlesungen ist daher oft ein Ziel technologieunterstützter Vorlesungen.

Ein herausragendes Merkmal von sozialen Medien ist sicherlich die Fähigkeit, Interaktionen in großen Gruppen zu ermöglichen. Soziale Medien scheinen deshalb eine geeignete Grundlage für technologie-unterstütztes Lernen zu sein. Jedoch sind allgemeine soziale Medien häufig auch mit Unzulänglichkeiten behaftet, die eine zweckmäßige Nutzung in Vorlesungen erschweren. Diese Arbeit beschäftigt sich deshalb mit der Konzipierung eines sozialen Mediums genannt *Backstage*, das speziell für die Nutzung in Vorlesungen zugeschnitten ist.

Backstage ermöglicht sowohl dozenten- als auch eine studenteninitiierte Kommunikation mit Hilfe eines Audience Response Systems und eines Backchannels. Audience Response Systeme ermöglichen die Durchführung von Quizzen in Vorlesungen, beispielsweise um Wissen abzufragen, und können so als eine technologische Unterstützung des Fragenstellen durch den Dozenten betrachtet werden. Diese Systeme sammeln und aggregieren die Antworten der Studenten und liefern in Echtzeit die Ergebnisse zurück an die Hörerschaft. Es konnte gezeigt werden, dass Audience Response Systeme effektive Mittel zur Aufrechterhaltung vorlesungsbezogener Interaktivität sind. Durch einen Backchannel können auch Studenten Kommunikation mit anderen Studenten oder dem Dozenten initiieren. Der auf *Backstage* verfügbare Backchannel basiert auf Microblogging, was sich über die letzten Jahre zu einem sehr beliebten Kommunikationsmedium entwickelt hat. Eine Schlüsseleigenschaft des Microbloggings ist die Kürze von Nachrichten, die aus nur wenigen Wörtern bestehen. Die knappe Kommunikationsform macht Microblogging sehr attraktiv als Backchannel für Vorlesungen.

Eine vorläufige Evaluation des ersten Prototyps, die zu einem frühen Zeitpunkt im Projekt durchgeführt wurde, zeigte jedoch, dass ein konventionelles Backchannel dazu neigt, die Teilnehmer zu überlasten. Sogar der Backchannel-Diskurs einer relativ kleinen Gruppe kann schnell unüberschaubar werden. Die Unüberschaubarkeit hat ihre Ursachen in einer mangelnden interaktionalen Kohärenz, eine vergleichsweise niedrige Kommunikationseffizienz, eine hohe Informationsentropie und eine fehlende Verknüpfung zwischen Backchannel und dem Vorlesungsvortrag.

Diese Arbeit untersucht mögliche Abhilfen für die genannten Probleme. So werden Vorlesungsfolien integriert, um damit den Austausch auf dem Backchannel zu strukturieren und in einen Kontext zu bringen. Die Backchannel-Kommunikation wird zudem neu konzipiert, so dass es ein kollaboratives Annotieren von Folien mit Hilfe von getypten Backchannel-Nachrichten umsetzt. Die Typisierung von Backchannel-Nachrichten

dient dazu, die Informationsentropie zu reduzieren. Um eine Verknüpfung mit dem Vorlesungsvortrag herzustellen, müssen zudem Backchannel-Nachrichten an die betreffenden Stellen auf Folien positioniert werden. Die Vorlesungsfolien verbessern auch die Kommunikationseffizienz durch das Routing, so dass der Backchannel nur die Nachrichten anzeigt, die zur aktuell angezeigten Folie gehören. Weitere Verbesserungen und Anpassungen des Systems, z.B. des Audience Response Systems, werden in dieser Arbeit beschrieben.

Diese Arbeit berichtet über eine Evaluation von *Backstage* in vier großen Vorlesungen. Die Ergebnisse sind vielversprechend. So begrüßten die Studenten den Einsatz von *Backstage*. *Backstage* erhöhte nicht nur die Interaktivität sondern trug auch zur sozialen Awareness bei, die eine Voraussetzung für eine aktive Teilnahme ist. Die Backchannel-Kommunikation war zu einem hohen Grad vorlesungsbezogen. Zudem konnte in einer weiteren Studie, die zusammen mit Pädagogen durchgeführt wurde, gezeigt werden, dass Studenten ihre mobilen Endgeräte in *Backstage*-unterstützten Vorlesungen mehr für vorlesungsbezogene Aktivitäten genutzt haben als in konventionellen Vorlesungen, in welchen die mobilen Endgeräte hauptsächlich für vorlesungsfremde Aktivitäten genutzt wurden.

Um soziale Kontrolle auf dem Backchannel zu etablieren, untersucht diese Arbeit Rating und Ranking von Backchannel-Nachrichten. Darüber hinaus schlägt diese Arbeit ein Reputationssystem vor, das als Ziel hat, einen Anreiz für erwünschtes Verhalten auf dem Backchannel zu schaffen. Das Reputationssystem basiert auf einer Eigenvektor-Zentralität, die an *Google*s *PageRank* angelehnt ist. Es ist zu einem hohen Grad anpassbar und ermöglicht auch die Berücksichtigung von Quizleistungen in der Berechnung von Reputation. Alle diese Ansätze, Rating, Ranking und Reputationssysteme haben sich in allgemeinen sozialen Medien als sehr effektive Mittel für soziale Kontrolle erwiesen.

LIST OF PUBLICATIONS

Some ideas and figures presented in this thesis previously appeared in one or several of the following publications:

Baumgart, Daniel, Alexander Pohl, Vera Gehlen-Baum, and François Bry (2012). "Providing Guidance on Backstage, a Novel Digital Backchannel for Large Class Teaching." In: *Education in a Technological World: Communicating Current and Emerging Research and Technological Efforts.* Badajoz, Spain: Formatex, pp. 364–371.

Bry, François, Vera Gehlen-Baum, and Alexander Pohl (2011). "Promoting Awareness and Participation in Large Class Lectures: The Digital Backchannel Backstage." In: *Proceedings of the IADIS International Conference e-society.* Spain, Avila, pp. 27–34.

Bry, François and Alexander Pohl (2014). "Backstage: A Social Medium for Large Classes." In: *Campus Transformation - Education, Qualification and Digitalization.* Ed. by Frank Keuper and Heinrich Arnold. Berlin, Germany: Logos, pp. 255–280.

Gehlen-Baum, Vera, Alexander Pohl, Armin Weinberger, and François Bry (2012). "Backstage – Designing a Backchannel for Large Lectures." In: *Proceedings of the European Conference on Technology Enhanced Learning.* Saarbrücken, Germany, pp. 459–464.

Gehlen-Baum, Vera, Armin Weinberger, Alexander Pohl, and François Bry (2014). "Technology Use in Lectures to Enhance Students's Attention." In: *Proceedings of the 9th International Conference on Technology Enhanced Learning.* Graz, Austria, pp. 125–137.

Hadersberger, Julia, Alexander Pohl, and François Bry (2012). "Discerning Actuality in Backstage – Comprehensible Contextual Aging." In: *Proceedings of the European Conference on Technology Enhanced Learning.* Saarbrücken, Germany, pp. 126–139.

Pohl, Alexander, François Bry, Jeannette Schwarz, and Marlene Gottstein (2012). "Sensing the Classroom: Improving Awareness and Self-Awareness of Students with Backstage." In: *Proceedings of the International Conference on Interactive and Collaborative Learning.* Villach, Austria, pp. 1–8.

Pohl, Alexander, Vera Gehlen-Baum, and François Bry (2011). "Introducing Backstage – A Digital Backchannel for Large Class Lectures." In: *Interactive Technology and Smart Education* 8.3. Extended Version of "Promoting Awareness and Participation in Large Class Lectures: The Digital Backchannel Backstage", pp. 186–200.

— (2012). "Enhancing the Digital Backchannel Backstage on the Basis of a Formative User Study." In: *International Journal of Emerging Technologies in Learning (iJet)*

7.1. Extended Version of "Assessing Backstage - A Backchannel for Collaborative Learning in Large Classes", pp. 33–41.

ACKNOWLEDGMENTS

CONTENTS

Part I

Introduction

1

SETTING THE SCENE

Current university teaching suffers from a number of shortcomings for which social media seem to offer remarkably fitting remedies. However, general purpose social media introduce new problems into teaching that make the overall benefit questionable. Therefore this thesis proposes a special purpose social medium aimed at overcoming the shortcomings of present university teaching while avoiding the problems of general purpose social media.

1.1 Shortcomings in Current University Teaching

The educational expansion in Europe, which began in the 1950s and 1960s, has opened higher education to the general public. Before, around World War II, only about two percent of the European 20-24 age group, mostly members from the upper and upper middle classes, had access to higher education (Hartmann, 2007). The educational expansion reduced the so far considerable educational inequalities. Yet, the ensuing rapid increase in the number of students quickly turned European universities into "mass universities": In France, the country in which the educational expansion set in first in Europe, the number of enrolled students nearly tripled from 214,000 to 630,000 between 1960 and 1970 (Hartmann, 2007). Starting off with a few years delay, in Germany the number increased from 240,000 to 930,000 between 1960 and 1982 (Hartmann, 2007). A similar increase in the student numbers was not only recognized in many other continental European countries (Hartmann, 2007) but also in the UK (Gibbs and Jenkins, 1992) and in the United States (Juhn, Kim, and Vella, 2005). As for Germany, the Bologna Process[1] has caused further increase in the student numbers (Brugger and Wolters, 2012) and led to an unprecedented peak of 2.5 million enrolled students in 2012 (DeStatis, 2012; BMBF, 2013). For decades now, large-class lectures comprising several hundred enrolled students, particularly in popular studies and introductory courses, have become the rule rather than the exception in European higher education.

The primary method of instruction in higher education is the lecture (Costin, 1972; Gibbs and Jenkins, 1992; Bligh, 1998; Apel, 1999; Cooper and Robinson, 2000). In a

---

[1]The Bologna Process was initiated in 1999 with the goal to harmonize standards and qualifications of higher education across Europe (see Wikipedia, `http://bit.ly/1e8HVeY`, last visited on June 16th, 2015).

lecture an entity, hereafter the lecturer, presents an oral discourse on a particular subject (Verner and Dickinson, 1967), possibly supplemented by written material, to an audience. In this constellation the lecturer is active and an audience member "a passive participant required to listen" (Verner and Dickinson, 1967, p. 85). For quite a while, the lecture has been criticized for the passivity of participants and its ineffectiveness in sustaining attention and active focusing, as Verner and Dickinson describe as follows:

> "In a lecture given by a brilliant scholar with an outstanding topic and a highly competent audience, ten per cent of the audience displayed signs of inattention within fifteen minutes. After eighteen minutes one-third of the audience and ten per cent of the platform guests were fidgeting. At 35 minutes everyone was inattentive; at 45 minutes, trance was more noticeable than fidgeting; and at 47 minutes some were asleep and at least one was reading. A casual check twenty-four hours later revealed that the audience recalled only insignificant details, and these were generally wrong." (Verner and Dickinson, 1967, p. 90)

Although Verner and Dickinson's description might have been influenced by a zeitgeist prone to criticizing traditional educational forms, their criticism of the lecture can hardly be discarded. Maintaining the students' attention during a lecture is challenging for the lecturer and the students alike. Indeed, research on this matter has shown that concentration of students, if not being actively involved, declines significantly within the first 10-30 minutes of a lecture (Stuart and Rutherford, 1978). For example, in four different lecture formats, Young, Robinson, and Alberts (2009) investigated what in ergonomics is termed *vigilance decrement*: the decay of attention of humans when information is processed passively, as it occurs in human supervision of automated systems; in such tasks, performance of humans is typically poor. By measuring the vigilance decrement of students the authors confirm that in a conventional lecture with no particularly stimulating material to capture attention and no further interaction, after about 30 minutes students expose vigilance problems similar to human monitors, with likewise notable adverse effects on learning quality and performance.

Karp and Yoels (1976) investigated the passivity of students on the basis of the social dynamics in class: The authors' observations of ten classes reveal that a great deal of the classroom interactions between the lecturer and the audience is initiated by the lecturer and always involves nearly the same handful of students. In their study the average number of students who took part in the lecture did not vary with class size. It even seems that the large part of the audience relies on those few students to interact with the lecturer in its stead, which the authors refer to as *consolidation of responsibilities*. The lecturer's behavior towards students also favors passivity: The lecturer is highly unlikely to call on specific students as she believes the classroom is fraught with anxiety. Indeed, students are afraid of appearing incompetent in front of peers and the lecturer and thus remain passive to avoid embarrassing situations (Paulsen, Bru, and Murberg, 2006). This may also explain why students having comprehension problems remain silent in class (Schworm and Fischer, 2006). Related to fearing disapproval by others, Karp and Yoels (1976) discovered that students may believe that the arguments and statements made in class need to be as well formulated, sound and complete as the lecturer's statements, not taking into account the time necessary for preparing the lecture. It even seems that the better prepared the lecturer the more are students likely to remain passive. Students are also seldom tested in a lecture and if tests are carried out, these are usually announced in advance. Hence, Karp and Yoels (1976) conclude that the lecture makes it easy for students to opt for non-involvement and to feel little obligation of coming prepared to class.

Although there is wide agreement on the benefits of transforming the lecture to an intellectually involving environment, remedies become considerably more difficult when classes are large: Possibly the biggest issue of large classes is a lack of personal contact the lecturer and the students have to get along with (Gibbs and Jenkins, 1992; Carbone, 1999; Cooper and Robinson, 2000). When classes grow bigger, the maintenance of personal relationships becomes increasingly difficult, and interactivity common in small classes becomes merely impossible. Instead, the audience appears to both the lecturer and students as an anonymous and homogeneous mass (Gibbs and Jenkins, 1992). Findings of psychological and sociometric research underpin the empirical finding that interactivity problems are inherent in large groups. According to Dunbar, Duncan, and Nettle (1995), conversational groups are inevitably limited in size; specifically, the authors suggest an upper bound of four participants in a free conversation. Furthermore, Sommer (1961) states that the maximum social distance for convenient conversation is about two meters. Also, visual cues are essential for regulation of turn taking in conversations, which is why conversation partners tend to sit opposite to each other. In the lecture hall, however, students usually sit next to each other, and the social distances within the audience easily exceed the suggested maximum. For these reasons, in a lecture strict moderation by the lecturer becomes indispensable. According to Nassaji and Wells (2000), classroom interaction often adheres to the triadic Initiation-Response-Evaluation/Feedback (IRE/F) pattern: The lecturer initiates interaction with students typically by asking a question; a few students respond to the question; the lecturer then evaluates or comments the students' response or provides some other kind of feedback. During such an IRE/F-based exchange, the greater part of the audience remains passive. The difficulty of genuine interaction in large classes, for example, is also recognized by students interviewed by Kowalski (1987): Those who welcomed more discussion during the lecture likewise agreed on the fact that this is hardly possible in conventional large-class lectures.

Apart from IRE/F-based exchanges, many lecturers tend to resort to an undisrupted speech with a short break in the middle, which further eliminates occasions of active participation in the lecture (Gibbs and Jenkins, 1992): On the one hand, a larger audience produces a higher level of noise in the lecture hall that may cause a perceived risk of losing control; an undisrupted speech is a way of keeping control. On the other hand, an undisrupted speech increases throughput of information and thus improves the efficiency of the lecture – so it seems. In any case, the resulting low interactivity adversely affects the frequency and quality of feedback: The lecturer receives only few insights into how students can get along with the subject matter (Gibbs and Jenkins, 1992). The students gain only little knowledge about their learning progress until the examinations at the end of a semester. During a semester, they have little sense of how well they are doing and may receive only few advice on how they can improve their own learning progress (Gibbs and Jenkins, 1992). As a ramification, there is a risk of a vicious circle: A lack of interaction causes a shortage of feedback and classroom awareness that, in turn, inhibits interaction. As Bligh (1998, p. 151) puts it provokingly, the lecture is a "[...] form of one-way verbal communication such that neither the lecturer nor his audience has adequate knowledge of what they are doing".

1.2 General Purpose Social Media and Teaching

Social media have been praised as a means to facilitate new forms of student-centered learning experiences (Redecker, Ala-Mutka, and Punie, 2010), particularly in higher education (see Tess, 2013 for review). The elements of active learning, i.e., partaking,

collaborating, sharing, and maintaining relationships are also the hallmarks of social media (Ebersbach, Glaser, and Heigl, 2011; Kaplan and Haenlein, 2010), making social media seem to be a natural fit in this regard. More specifically, social media can be cornerstones in transforming the large-class lecture to a more participatory and intellectually involving learning environment by overcoming spatial, temporal and capacity limitations of face-to-face interactions in the lecture hall. Social media can provide situated feedback on the basis of the users' interactions, improving awareness and, in turn, promoting participation.

Besides praises of social media in higher education, also critical views can be found in the literature. Clark (1994) argues not to overestimate the significance of media for learning; rather, media should be considered as means to support the instructional methods chosen by the lecturer in the most purposeful manner possible. A critical view of using social media in the classroom is also reasonable because of the differences between social media vendors' interests and educational goals. That is, general purpose social media may likely provide interactions and activities that are rather disadvantageous or even inappropriate in an educational setting. For example, Friesen and Lowe (2012) highlight the commercial interests behind *Facebook* and similar social media that inevitably shape the user experiences. Thus, *Facebook* and similar social media exceedingly highlight conviviality and liking, while disagreement and criticism, both valuable, if not fundamental, learning activities, are disfavored and refrained. *Twitter*, as another example, seems to be a communication platform suitable for the use in lectures. However, it is questionable whether the provision of public world-wide communication in an open community as provided by *Twitter* does not rather contribute to student disengagement and distraction than to learning-related interactivity in the classroom. Apart from these issues, the question arises whether students are willing to adopt certain social media for learning in the first place. To investigate this question, Jones et al. (2010) evaluated 76 survey responses and 14 student interviews and discovered that students actually distinguish between social media for private and for non-private uses. In particular, their findings indicate that students do not expose significant learning activities on social media they consider part of their private lives.

In spite of the great potential of social media in education, a multitude of issues may inhibit their successful and optimal use in educational settings, especially in the lecture.

1.3   Backstage: A Social Medium for Teaching

This thesis focuses on the conception of a special purpose social medium specially tailored to the use in lectures. It aims at overcoming the limitations of face-to-face interaction and at scaling learning-related interaction common to small classes up to larger classes, free of potentially conflicting commercial interests and issues regarding private uses. More specifically, this thesis describes *Backstage*, a so-called digital backchannel[2] specially designed for the use in lectures. A backchannel describes computer-mediated communication used as an additional non-disrupting communication channel initiated and mainly maintained by students (Yardi, 2006; Kellogg et al., 2006). On a backchannel students can exchange questions, comments and thoughts on the subject matter synchronously to the lecturer's presentation.

To account for the requirements imposed by the lecture the backchannel capitalizes on features of social media not only for encouraging participation but also for handing responsibilities over to the audience. For example, while presenting the subject matter

---

[2]Hereafter, digital backchannels are simply referred to as backchannels.

the lecturer is not able to maintain a comprehensive overview of an extensive backchannel. However, collaborative filtering of the backchannel allows the audience to direct the lecturer's attention to what they consider relevant and noteworthy. In a sense, while the lecturer moderates the frontchannel, i.e. the traditional lecture hall communication, the moderation of the backchannel relies with the audience.

The backchannel is deliberately complemented with an Audience Response System enabling the conduct of quizzes during the lecture as a means to frequently provide situated feedback regarding learning progress and student retention. The artifacts generated by the two components are integrated with the lecture slides to persistently mirror what has happened in the lecture, making the backchannel a valuable reference for rework for both lecturer and students, and also incentivizing high-quality contributions on the backchannel.

1.3.1   Modus Operandi of a Backchannel for Lectures

The following exemplary, though admittedly simplified, scenario inspired from teaching experiences of lecturers using *Backstage* in large-class lectures shall provide an idea of how a lecture supported by a backchannel may look like (taken from Bry, Gehlen-Baum, and Pohl, 2011 with few modifications):

Like every Monday morning, Sue Byrd is giving the introductory lecture on Linear Algebra for an audience of 150 students. She feels better because the backchannel is running and she as well as the students begin to get used to it. A reminder would not harm, though: "If at some point you get lost, post it on the backchannel" Sue says, "And refrain from messages on yesterday evening's soccer." When it comes to the Laplace formula for computing the determinant of an n-by-n matrix, there is a rush of students' messages. Sue sees that many are posting messages and therefore waits a few seconds for the backchannel to inform her: "What is a matrix' signature?" was supported by 37 per cent of the audience. She recalls the concept needed for understanding the Laplace formula and goes ahead with her explanation. Meanwhile, Bob is engaged in a private chat on the backchannel with Ann: "I thought the formula was called Leibniz formula. Is Dr Byrd confusing names?" Ann finds the remark interesting and suggests to Bob to make it public. He does so and within five minutes more than 20 per cent of the lecture's audience support the question, leading the backchannel to forward it to Sue who mentions that there are two distinct formulas for determinants of n-by-n matrices. At the same time, Charlie asks on the backchannel "Recall the formula for the det of 2x2 matrix". Ann, who often tends to be rather direct, posts a public answer: "Dr Sue explained it last week but you were once again skiing." This sharp answer gets support and Charlie's question is not forwarded to the lecturer. After completing the presentation of the Laplace formula, Sue runs a quiz to see whether the students have well followed her explanations. And indeed, Sue is very satisfied with the result: 86 per cent of the 120 students who participated in the quiz submitted the correct answer. As one of the few students who were wrong, the quiz helped Charlie recognize that he should urgently catch up with the subject matter he missed last week – and that he better refrains from skipping Sue's lectures. After the lecture, Sue has a look at the backchannel's log: She is satisfied that the communication has been limited to nine of the 90 minutes of the lecture and to three issues – among others yesterday's victory of the FC Bayern. She notices that the number of students more interested in FC Bayern's victory than in determinants fits well with that of students who did not do last week's home work.

7

This scenario aims at highlighting the following points: Firstly, even though it is preferable that students promptly ask questions when they do not recall a concept, in a large class lecture this is meaningful only if the question is relevant to a sufficiently large number of students. And even then, only a limited fraction of meaningful questions can be discussed in large classes. It is, hence, desirable to have the questions persisted so as to be able to revisit urgent questions outside class or in the subsequent lectures. Secondly, side remarks to neighbors might be relevant to the whole audience and worth forwarding to the lecturer. Deciding whether this is the case should be a social decision. The mechanism of socially deciding on the relevance also offloads the task of steadily monitoring and moderating the backchannel to the audience. Thirdly, quizzes provide a valuable source of feedback for both the lecturer and the students. The lecturer can quickly gain insights whether students got lost, and the students receive immediate feedback on their learning progress and may accordingly adjust their learning activities. Finally, the existence of a sophisticated backchannel for learning does not automatically entail that every student is able to use the backchannel optimally for learning. A backchannel should address this issue to the greatest extent possible, e.g., guiding students in their interactions on the backchannel and making non-learning related activities on the backchannel difficult.

1.4   Contributions and Outline of this Thesis

This thesis elaborates on the conception of a so-called classroom communication system (see Chapter 2) with the aim to promote active participation of students in (large-class) lectures. Therefore, the system provides quizzes in lectures as well as carefully designed computer-mediated communication. The concept is prototypically realized as a Web application, making it possible to integrate and harness mobile devices students usually bring to lectures, for a better learning in lectures. The prototype was evaluated in four Computer Science courses held by different lecturers at different universities. The findings are quite promising. The prototype was well received by students. Both the data collected on the system as well as the students' impressions about using the system indicate that the system promotes interactivity (making lectures more appealing), increases social awareness (in turn promoting active participation), and caters for lecture-relevant electronic exchange. The main contribution of this thesis is to give evidence that a social medium carefully designed to promote communication favorable to, and to prevent communication unfavorable to, learning can be designed and that its deployment in large-class lectures considerably improves both the activity and the "social quality" of large classes.

This thesis is divided into four parts. Part I is the introductory part of this thesis. Chapter 1 is this introduction. Chapter 2 gives an introduction to research fields relevant to this thesis. Chapter 3 presents a first prototype which was preliminarily evaluated to gain insights for the further conception and development of a classroom communication system (CCS) for large-class lectures. Part II is devoted to *Backstage*, a backchannel and Audience Response System (ARS) targeted at large-class lectures. Chapter 4 describes the concepts and realization of *Backstage*. Chapter 5 reports on an evaluation of *Backstage* in four courses over two years. Part III considers first steps in further extensions of *Backstage* towards social regulation. Chapter 6 describes rating as a means to assess the relevance of backchannel comments and to direct the lecturer's awareness. Chapter 7 presents an eigenvector-based reputation mechanism aimed at incentivizing desirable behavior on *Backstage*. This thesis concludes with Part IV. Chapter 8 considers perspectives and gives an outlook to further developments of *Backstage*.

2

CLASSROOM COMMUNICATION SYSTEMS (CCS)

> Section 2.1.2 is based on Pohl, Bry, et al. (2012). Section 2.6.2 is based on Bry and Pohl (2014).

The more lectures focus on imparting knowledge, the more they tend to exhibit a one-way flow of communication directed from the lecturer to the audience; the audience mainly stays passive and rarely engages in the lecture. Furthermore, students do not exchange information with one another, making learning a solitary task in the lecture. This situation is depicted in Figure 2.1. As outlined in the previous chapter, making the lecture more welcoming to an active involvement of students is impeded by various factors such as the classroom logistics (e.g., the class size, the extent of moderation required due to class size, floor time which is the time available to students to speak publicly), the students' confidence and degrees of communication apprehension, the students' personality traits (e.g., extroversion), the lecturer's work towards student participation, and the classroom climate (as reviewed by Rocca, 2010). Regardless of these limitations, however, lecturers often want greater student participation and also consider it to be a necessary element of lectures (Dancer and Kamvounias, 2005).
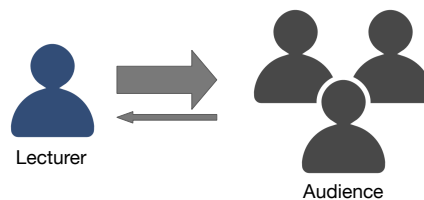
Lecturer

Audience

Figure 2.1: A conventional large-class lecture can be characterized by a mostly uni-directed flow of information from the lecturer to the audience and the isolation of students in the audience.

In spite of its well recognized shortcomings, the lecture remains the standard teaching format of higher education. Arguably, the lecture is better than its reputation among researchers and practitioners. Possibly, it will survive new movements in online educa-

tion, e.g., the MOOCs, as it has survived the paper print which made books considerably cheaper, thus affordable for students, and took the lecture its initial raison d'être: reading to students a book much too precious for being handed out to students. Three aspects may be seen as arguments in favor of the lecture as a convenient form of teaching (Bry and Pohl, 2014):

- The lecture gives an opportunity for students to come together and to be face-to-face with the lecturer.
- The lecture provides a structure and a concise presentation of material students should focus on.
- The lecture gives a learning pace.

An indication of the arguments' validity can be found in students' assessments of the lecture format. For instance, a study conducted by Gysbers et al. (2011) examined why students still attend lectures when, nowadays, ample lecture material in form of videos, audio recordings and texts is available online, e.g., provided by MOOCs and other Web-based lecture technologies. More than 500 undergraduate students who were taught in conventional large-class lectures (Biosciences, mainly presenting *PowerPoint* slides with minimal use of whiteboard or document viewer) participated in the survey. The comments submitted were overwhelmingly positive about lecture attendance. Accordingly, students actually enjoy lectures and consider lectures a part of the "university experience". The students valued physical presence in the lecture hall and being face-to-face with the lecturer, which they find more engaging than just listening to online lectures. The students also appreciated the discipline for learning provided by lectures as compared to online courses. They valued that lectures not only provide external motivation to maintain a structured schedule but also that lectures allows to more easily keep pace with the learning material. That is, despite its limitations and drawbacks, the lecture seems to enjoy a high reputation among students compared to online alternatives.[1]

Today, many students bring their Internet-enabled mobile devices, e.g., laptops, tablets, and smart phones to the lecture. Because of decreasing purchase and running costs more and more students can afford these devices. Also many lecture halls are equipped with wireless access points, certainly encouraging the use of mobile devices in lectures. In conventional lectures a proper use of mobile devices as a learning tool is mainly left to the students and hardly accounted for in the teaching, though. The potential of mobile devices in the lecture as a mediator for student participation and thus facilitating a "coming together" in large-class lectures remains untapped. A well-considered integration of Internet-enabled mobile devices in the lecture as so-called classroom communication systems can remedy shortcomings in classroom logistics and improve the classroom climate by supporting interaction between the lecturer and the audience, as well as among students.

2.1 Using CCSs to Enhance Feedback and Classroom Awareness

Hardware and software used in classroom as a means to increase interactivity are referred to as *classroom communication systems* (Beatty, 2004), or CCS for short. CCSs provide services to support and facilitate various student-to-lecturer and student-to-student interaction with the purpose to help making large classes seem more like small classes

---

[1]Of course, the positive attitude towards the lecture format is the result of general consideration whether presence lectures still have a right to exist, where alternatives are provided by Web-based lecture technologies. Thus, it does not anticipate the students' daily decisions whether and which lectures to attend. Students do not attend lectures mainly due to time clashes with overlapping lectures and jobbing aside.

(Beatty, 2004). CCSs are means to mediate interaction in the lecture hall by providing communication channels, additionally to face-to-face communication, less prone to social inhibitors. CCSs create opportunities for participation and scale up in-group interactions for which the audience might otherwise (i.e., without a CCS) be too large. However, these opportunities are created in an unobtrusive manner, meaning that it does not overly compromise the lecture as the chosen form of teaching (see Figure 2.2). Thus a well-considered use of a CCS to improve interactions may result in benefits for the teaching and learning in large classes by providing valuable feedback to, and enhancing classroom awareness of, both the lecturer and the students.
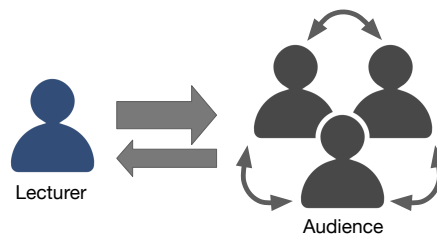
Figure 2.2: The incorporation of Internet-enabled mobile devices and a well-considered designed CCS can be used to scale interactions known from small learning groups to large audiences. These interactions between the lecturer and the audience and among students improve the lecture as a social learning environment.

2.1.1 Improving Feedback

Interactions in the lecture can generate frequent and instant feedback valuable for learning and teaching. The term feedback refers to information provided by some other entity, e.g., the lecturer or peers, about aspects of one's performance or understanding (Hattie and Timperley, 2007).

Feedback is one of the most important influences on learning and achievement (Hattie, 2009). It is essential in eliciting and reducing discrepancies between a student's current understanding or performance and the desired learning goal. For example, feedback can be used to reduce discrepancies by providing the student with a more appropriate learning goal or by assisting the student in reaching the learning goal. Likewise, students may increase effort or use more effective learning strategies to reach the learning goal. Feedback relates to the learning goal, the student performance, and the next steps required to achieve the learning goal. It thus addresses the quality of the student's performance, the processes required to perform a task, and the student's ability to self-regulate learning (Hattie and Timperley, 2007).

Aside from student learning, feedback can also inform the lecturer in her teaching. In fact, question asking as a means to diagnose understanding in formative assessments is the second most common instructional activity of the lecturer besides presenting the lecture content. The purpose of the lecturer's questions is mainly to test whether memorization and recall by students is successful (Hattie, 2009). The student answers, or hesitation to answering the lecturer's questions, may give feedback not only on whether memorization and recall is successful, but also indicate difficulties that should be resolved as soon as possible. The lecturer's questions also guide students in that they indicate which aspects of a topic the lecturer considers important and how well the students can follow the lecture. However, whether students recognize the importance of

11

the lecturer questions as a guide for their learning is at least unclear. Another problem related to lecturers' question asking is that lecturers tend to give students not enough time to elaborate on the question and to formulate an appropriate answer. Lecturers typically wait too short for an answer and ask follow-up questions too early (as reviewed by Rowe, 1986). Helping the lecturer in providing enough time would certainly make question asking more beneficial for both the lecturer and the students.

As for the lecturer, questions asked by students in public can as well be a valuable feedback source, since they enable the lecturer to gain insight in how the students approach the topic, try to solve problems and intend to come to understanding (Chin and Osborne, 2008). This feedback can be accounted for by the lecturer in her responses to student questions. It enables the lecturer to respond to the questions with appropriate feedback to the questioners (Chin and Osborne, 2008) and also inform the lecture's current and future teaching.

A CCS can be used to use the lecturer's questions as opportunities for formative assessments and generation of feedback. Using a CCS can make the lecturer's question explicit as a form of (perceivable) instruction and as learning aid to students. The perception of the lecturer's questions as learning aids can be improved by systematically gathering and evaluating the students' responses and by making the results available in the lecture as well as in the revised lecture material. The results obtained also helps to elicit "hot spots" in the topic, detect fallacies at an early stage during the lecture, and to track and steer the learning in a way that would hardly be possible without the use of a CCS.

2.1.2 Improving Social Awareness

The term social awareness refers to "an understanding of the activities of others, which provides a context for your own activity" (Dourish and Bellotti, 1992, p. 107). In a sense, feedback and awareness can be seen as the two sides of the same coin.

In large-class lectures the lecturer and the students typically have a low degree of awareness. The lack of awareness contributes to barriers that hinder in-class participation. For example, students who have questions may hesitate to ask because they think they are the only ones who have problems in understanding the topic. However, knowledge of the questions and thoughts others have may provide an understanding that "the others are like me" and, hence, conduce to student participation. Similarly, awareness may as well indicate to the lecturer and the students which activities are less desired. Recognizing that others have similar problems or make similar remarks, and noticing peers that help each other, may create a common ground that promotes active involvement. Research has shown that awareness contributes positively to communication and interaction and aids the development of a shared sense of community (Dourish and Bly, 1992), enabling participation in the lecture as well as in the computer-mediated discourse. More participation of students in the lecture engenders an increased disclosure of viable information that in turn is important for the creation of awareness (Rittenbruch and McEwan, 2009).

One can distinguish between passive (Dourish and Bly, 1992) and active, i.e. intentionally and deliberately enriched, awareness (Rittenbruch, Mansfield, and Viller, 2009). The former refers to awareness based upon information that is automatically gathered by the system without one knowing, then correlated and distributed to the others. At first glance, passive awareness does not require an active disclosure of information by others. It focuses on information that emerges unintentionally, e.g., by simply attending a lecture. Active awareness is defined as including information that are actively shared

with others. It is created by intentional communication.[2]

The impact of social awareness in collaborative e-learning has been investigated by Lambropoulos, Faulkner, and Culwin (2012). They report on the integration of awareness tools in the e-learning platform Moodle[3] and on the assessment of the tools' effects on participation. The results reported about in that article indicate that social awareness can indeed contribute to active participation. CCSs, therefore, should focus on gathering information related to social and workspace awareness (Gutwin, Stark, and Greenberg, 1995). According to Gutwin and colleagues, social awareness is the consciousness of the social interrelations in a community. Social awareness refers to the questions "What should I expect from other members of this group?" and "How will I interact with this group?" Thus, social awareness is mainly related to users and a community. Workspace awareness refers to the questions "What are they [the other group members] doing?", "What have they already done?", "Where are they?", and "How can I help other student to complete the project?" (Gutwin, Stark, and Greenberg, 1995).

2.2 Computer-Mediated Communication and Digital Backchannels

Computer-mediated communication (CMC) refers to interpersonal communication using networked computers (Döring, 2003). Obviously, classroom communication systems (CCSs) require some form of CMC, but CMC is not limited to communication in teaching contexts. Although CMC comprises any sort of interpersonal communication including social media and even online multi-player games, the term often refers to an exchange of textual messages between users. Compared to face-to-face communication, CMC has shown to be less rich in terms of additional auditory and visual cues that make face-to-face communication so vibrant and personally immediate. For this reason CMC is also referred to as a lean medium (Herring, 1999). However, a lean medium can be beneficial, especially when the personal immediacy in face-to-face communication becomes a problem (for instance, because of social cues, see below). Thus, it can be sensible to use CMC in lectures to cater for the need of communication channels less prone to social barriers.

2.2.1 Characteristics of Computer-Mediated Communication

CMC can be distinguished by the temporal coherence of exchange it provides. In synchronous CMC, text messages are delivered and received (almost) in real-time, requiring no explicit action of the receiving participant. Synchronous CMC is dialogic and informal. Discourses on synchronous CMC span short durations, usually not more than several hours or days. The messages tend to be brief and less elaborated. Also, to fully understand their meanings, the messages need to be interpreted in the context of the CMC discourse. Primal examples of synchronous CMC are Internet Relay Chats, Instant Messaging, and other systems used as digital backchannels (see below). In contrast, asynchronous CMC is not time-coherent, and resembles (hand-written) correspondence.

---

[2]The concepts of awareness described above are subject to some debate. Schmidt (2002) argues that the term awareness has been used in contradictory ways and that it is in many cases unclear how awareness actually comes to existence. Among others, the distinction between passive and active awareness is questioned. For example, Schmidt (2002) surmises that actors adapt their activities and determine the appropriate degree of obtrusiveness in order to be perceived by others, which does not happen unintentionally. Awareness, he states, "is not the product of passively acquired 'information' but is a characterization of some highly active and highly skilled practices" (p. 292). However, he notes that actors somehow monitor the activities of others to ascertain the state, progress and direction of these activities.

[3]https://moodle.org

Asynchronous CMC messages are not shared in real-time and receiving a message often requires explicit action by the user. The messages are in general more formal, self-contained, elaborated and rather monological. Asynchronous CMC discourses may span longer time periods, i.e., months or years. Primal examples of asynchronous CMC are e-mail and online forums.

2.2.1.1 Reduced Social Cues and Anonymity

Since CMC relies on fewer channels, communication partners are perceived differently from face-to-face communication. In particular, reduced information channels can be advantageous, since also social cues and social background information about the communication partners that tend to have strong influences in face-to-face settings are likewise reduced. For example, an impressive look or confident appearance usually creates communication advantages in face-to-face settings. In CMC, however, these social cues are filtered out. As a consequence, participation in CMC becomes more equal (Döring, 2003).

CMC also enables participants to control and change information available about them much more easily than in face-to-face settings. This may even result in complete anonymity of participants. Theories related to channel reduction and social-cues-filtered-out approaches assume that the equalization effect promotes participation (both pro-social and anti-social) and self-disclosure (Döring, 2003). Anonymity refers to masking one's true identity in CMC, either by using an invented name (a pseudonym) instead of one's real name or by using no name at all. Pseudonyms are mocked identities that allow CMC messages (or other interaction outcomes) to be linked to a sender, whereas a perfect anonymity makes such a linkage impossible (Table 2.1 gives an overview).

	Identifiable	Not Identifiable (Anonymous)
Linkable	Real Name	Invented Name (Pseudonym/Weak Anonymity)
Unlinkable	—	No Name (Perfect or Strong Anonymity)

Table 2.1: Types of Online Identities and Anonymity (built upon Miyazoe and Anderson, 2011). A user is identifiable if her real identity is not masked. Otherwise a user is not identifiable or anonymous. Linkability refers to whether a user can be linked to her online artifacts (e.g., her CMC messages or ratings).

Anonymity can also be useful in online learning. In blended course designs on English writing, Miyazoe and Anderson (2011) conducted an exploratory study of how participation varies between face-to-face (and real names) and pseudonymous on-line exchange. They also investigated how students perceived online writing using pseudonyms. Sixty-three students participated in the study. One of the authors' findings is that twelve out of fifteen students preferred the use of pseudonyms so as to free themselves from evaluations by others and from disclosing their mistakes. Some students also emphasized that anonymity enables the focus on content contributed rather than on the persons who contributed, thus indicating the relevance of anonymity for content-oriented discourse.

2.2.1.2 Interactional Coherence of CMC

From a linguist's viewpoint, CMC does not lend itself to mediation of interpersonal interactions. Linguists come to that conclusion by comparing CMC with face-to-face

communication, which they consider the ideal form of interpersonal communication (Döring, 2003). Accordingly, CMC has several drawbacks, making it inferior to face-to-face communication, two of which are referred to as interactional and topical disjointness (Herring, 1999).

CMC relies on fewer channels to transmit messages than in face-to-face interactions. Particularly when the participants in a CMC environment are not co-located, audio-visual cues are not available that provide information about how others are responding (Herring, 1999). While a lack of audio-visual cues can be advantageous (see Subsection 2.2.1.1), it may raise issues in turn taking of participants and topic development.

The reduced number of available channels affects turn taking, which usually requires simultaneous feedback, i.e. audio-visual cues and other backchannels (Lambertz, 2011) so as to be properly timed and to prevent communication overlap. For instance, a participant expecting a response by some other participant may not be aware about the fact that a message for her is currently being written. The participant may thus impatiently write another message that may render the awaited response redundant or even result in a topic shift. In the latter case communication disruption may be the result when the awaited response is submitted in the meantime.

Turn-adjacency refers to the logical and temporal coherence of messages. If it is assumed that one message corresponds to one turn, two related turns should be grouped by an immediately consecutive display of the corresponding two messages in the CMC environment. That is, the two messages should not be separated by another message of another turn or conversation thread. However, CMC systems often display messages in a linear order as received and processed by the system. Network lags may even produce a temporally inconsistent ordering in which a responding message is displayed before the message to be responded to.

Both characteristics of CMC may yield a disrupted and fragmented communication that becomes incomprehensible for participants. However, Herring (1999) argues that users can adapt to several limitations of CMC. What is more, users consider some limitations as integral part of CMC. Eliminating these limitations, e.g., by facilitating proper turn-taking, would make CMC less enjoyable. Other limitations are remedied by the persistent nature of CMC, allowing users to re-read parts of the communication at any time in order to better comprehend the discourse. Furthermore, special symbol systems are developed and used by users of CMC to provide for social cues where desired (e.g., emoticons) and to manage turn taking (e.g., the use of '%' at the end of a message to indicate that the user is not yet ready to give up the floor).

However, the design of CMC for a use in lectures may require special care. For example, both the lecturer and the students may not have enough time to re-read CMC messages in order to comprehend their meanings. A fragmentation due to a linear ordering of CMC messages makes comprehending a CMC discourse more difficult.

2.2.2   Facilitating and Supporting Student-Initiated Communication

An active engagement of students in large classes may notably benefit from, or even require, computer support of student-initiated communication meaningful for learning in the lecture. As the term suggests, student-initiated communication is initiated by a student and directed to the lecturer or other (possibly more knowledgeable) peers, or the audience in general. As mentioned above, lectures do not provide much room for introducing further interaction besides the lecture discourse. However, using note taking – a quite common student activity in lectures – as a basis for student-initiated communication seems promising.

2.2.2.1 Shared Note Taking as a Collaborative Practice in Lectures

Note taking is perhaps the most frequent learning activity of students in lectures. Students deem note taking a valuable in-class activity (as reviewed by Hartley and Davies, 1978). Thus, the process of note taking, i.e. the creation of notes, is believed to aid recall. But also the product of note taking, i.e. the notes themselves, provide additional information useful for revising and restructuring the lecture material. Although the research findings cited by Hartley and Davies (1978) could not confirm that note taking improves immediate recall (compared to not taking notes), evidence could be found that note taking is positively related to test scores. Furthermore, the usefulness of notes for revision of the lecture material has been shown. Besides, the productivity in note taking is used by students as an indicator for the degree of attention and effort invested in following the lecture. Also, students frequently use note taking as means to stay attentive in class.

Since note taking is so common in lectures, enabling sharing notes among students using mobile devices seems appealing. As Valtonen et al. (2011) point out, interesting opinions and approaches to understanding the lecture's topic are written down by students as personal notes but, despite their potential values for others, are not shared with the audience. Dealing with notes created by peers may be beneficial for a student in that it provides opportunities to contest her own knowledge or may help resolving her own cognitive conflicts. It can be assumed that the possibility to share notes may be well received by students. For example, in a survey of 220 undergraduate students, more than two-third of the respondents expressed interest in what their peers would have to say in class (Aitken and Neer, 1991).

Shared note taking is not only a way to counteract student passivity but it also introduces a form of collaborative practice among students in the lecture. In this way, learning can be more socially involving within the highly structured and moderated environment of a lecture that otherwise puts great emphasis on individual learning. It enables students and the lecturer alike to gain insights into thinking and learning "as it happens" and promotes reflection and exchange within the lecture discourse (Valtonen et al., 2011).

2.2.2.2 Support of Student Questions in Lectures

A more specific form of student-initiated communication is question asking (for an extensive review, see Chin and Osborne, 2008). Usually, questions come up in response to puzzlement when discrepancies or gaps between what is known and the presented information are recognized. Questions indicate that students are dealing with the lecture topic and try to make sense of it. Thus, questions can generate progress and steer individual knowledge construction. Aside from the meaning of question asking for individual learning, student questions may also play an important role at the social level. Student questions can evoke in-class discussions and debate, since questions and respective answers can provide alternative viewpoints. Also, questions may not only stimulate the questioner but also peers, since questions provide insight in how the questioner approaches a problem and intends to come to understanding. It may thus help peers to use similar approaches to advance their own learning.

In several articles (e.g., Ertl and Mandl, 2004; Rosenshine and Meister, 1994), it has been argued that questioning is an important strategy to elaborate new knowledge. Unfortunately, students often find it difficult to ask for help and further explanation (Aleven et al., 2003). This is not surprising, considering that formulating questions already requires some understanding of, and former knowledge about, the topic (Ertl and Mandl,

2004). Questioning is thus subject of various fields of research. For example, reciprocal teaching (Palincsar and Brown, 1984) and script research (King, 1994; King, 2007) deal with questioning or at least highlight its importance (O'Donnell and Dansereau, 1992). Yet, students seldom ask questions in lectures (Pearson and West, 1990; Aitken and Neer, 1991; Fritschner, 2000; Oblinger, 2004). Often students are socially inhibited due to the risk of appearing unintelligent or ignorant of the lecture's topic (Karp and Yoels, 1976; Fassinger, 1995). Many students also perceive the lecture as not receptive to student questions. For example, as Chin and Osborne (2008) point out, students are usually supposed to answer questions asked by the lecturer, not to ask questions themselves. To support asking questions it is thus recommended to make the lecture an environment more receptive and appreciative of student questions.

Contrary to the impressions that a reluctance to asking questions may give, students seem to appreciate opportunities for question asking (Aitken and Neer, 1991). For example, Bauer and Snizek (1989) offered their students to write down emerging questions on paper that were collected after the lecture, evaluated and then answered at the beginning of the next lecture. The authors reported that students highly appreciated the possibility to submit questions in this way. The questions asked were typical for students trying to understand the lectures' topics and thus were considered to conduce to learning. Furthermore, the authors noted that also shy students who otherwise did not participate in the lecture's discourse submitted questions on paper. Nowadays, CMC can be used to facilitate student question asking in a similar fashion. As Chin and Osborne (2008) point out, computers allow to ask questions without the need for face-to-face interaction, making asking questions easier for students. Furthermore, students have time to phrase and revise their questions as well as their answers before making them available to their peers and the lecturer. On CMC, answers do not have to be, but in contrast to Bauer and Snizek's pen-and-paper approach can be, given within the context that provoked the questions, while preserving the anonymity of the questioners.

2.2.3 Digital Backchannels

Originally, the term backchannel stems from the fields of linguistics and politics. In linguistics, a backchannel describes the listener's utterances of so-called acknowledgment tokens, such as "uh hu", to signal interest and engagement in the conversation without claiming the floor (cf. Lambertz, 2011). According to the acknowledgment tokens the speaker can monitor the conversation. In politics, a backchannel describes unofficial interaction in the background to preserve deniability in the public, among other things (Kellogg et al., 2006). Kellogg et al. (2006) also note that the use of digital backchannels can be understood in both linguistic and political terms: as means for the presenter to monitor the conversation, and as means for the audience to signal interest and engagement and to gain social support or agreement prior to interrupting the frontchannel. In the context of CMC, the term digital backchannel is used in a twofold manner: to refer to backchannels (in linguistic terms) based on CMC as well as to refer to specific software used to realize backchannels.

In the latter sense, digital backchannels are (usually text-based) CMC systems used to establish a secondary communication channel complementary to a frontchannel. The frontchannel usually consists of a central presenter, e.g., a lecturer, teacher, or a speaker at a conference, who demands sole attention by the audience. In such settings interactions are typically limited and directed from one (the presenter) to many (the audience). Digital backchannels are used to support participation of the audience by enabling interactions from many to many (Kellogg et al., 2006; Yardi, 2008; Ebner, 2009). CMC software used as digital backchannels thereby shares characteristics of face-to-

face conversations, i.e., the software provides communication that is of a synchronous, ephemeral, and informal nature (cf., Yardi, 2008). Chats and chat-like software fall into these categories, as opposed to e-mail or online forums. Like face-to-face communication, a digital backchannel is governed by social conventions (Cogdill and Kilborn, 2001). Unlike a typical face-to-face conversation, however, a conversation on digital backchannels is usually multi-threaded. Since most CMC systems obviate turn-taking of conversation partners, several possibly unrelated backchannel conversations may take place simultaneously (Herring, 1999; Cogdill and Kilborn, 2001).

Cogdill and colleagues observed that messages on a backchannel[4] usually fall into certain categories. They, therefore, devised a classification of backchannel posts. Accordingly, *process-oriented* backchannel posts analyzes and steers the frontchannel discourse. For example, posts about the frontchannel being too fast belong to this category. Process-oriented posts are often not conversational but rather individual and isolated. *Content-oriented* backchannel posts deal with the content of the frontchannel. It comprises reactions about the frontchannel discourse, such as expressing surprise, agreement, or disagreement. *Participation-enabling* posts are about supporting others in becoming acquainted with the communication software and with the social conventions that govern the backchannel. *Tangential* posts further elaborate on topics introduced but not further pursued in the frontchannel. Finally, *independent* posts are completely unrelated to the frontchannel or participation-enabling discourse.

2.3 Audience Response Systems (ARS)

So-called Audience Response Systems are a much noticed technology by which interactivity of students in lectures can be promoted (as reviewed by Fies and Marshall, 2006; Caldwell, 2007; Kay and LeSage, 2009 and Lantz, 2010), hereafter abbreviated as ARS or simply referred to as quizzes.[5] With an ARS lecturers can deliver short questions, often but not always limited to multiple-choice questions, to the audience. Usually the question together with possible answers to the question are projected on the wall. One or more of the answer alternatives are correct, the other alternatives shown, so-called distractors, are incorrect answers and used to challenge a student's knowledge. A student can choose from the answers using a remote control device. All student responses are collected by the ARS and summarized in real-time, e.g., in form of bar charts. Several extensions are conceivable. Haintz, Pichler, and Ebner (2014) evaluate the functionalities provided by twelve recent ARSs.

Although the concept seems quite simple, ARS have been conceived and used for various educational purposes (Draper and Brown, 2004): ARSs can be used for (1) assessment, e.g., practicing multiple-choice questions used in exams, and instant self-grading, (2) feedback on learning for both the lecturer and students on how well students understood the material, (3) feedback on the teaching (indicating a lecture being too fast, more examples wanted), although this does not strictly comply with multiple-choice questions, (4) peer assessment, where the work of peers are rated by the audience, (5) classroom

---

[4]In the following referred to as backchannel posts.

[5]ARS are also known as Classroom Response Systems, Student Response Systems, Group Response Systems, or Clickers. Sometimes, the term classroom communication system itself is used to refer to ARS. This thesis explicitly distinguishes between CCSs and ARS to emphasize that generic computer-mediated communication may be initiated by students, whereas communication using an ARS is initiated by the lecturer. We also do not follow the classification by Ebner, Haintz, et al. (2014) who consider backchannels to be a kind of ARS as the communication modalities provided by ARS significantly differ from backchannels as described in Section 2.2.3.

awareness, i.e., recognizing if enthusiasm or difficulties are shared with others can improve feeling socially integrated in class, and (6) initiating group discussions.

Several studies referenced in the above-mentioned reviews point to the usefulness of ARSs for learning. Compared to conventional lectures ARSs increase student motivation to partake and to foster in-class discussions (Boyle and Nicol, 2003). The interruptions quiz sessions cause are likely to improve the learning in lectures since student concentration phases are known to span no more than 20 to 30 minutes (Bligh, 1998; Young, Robinson, and Alberts, 2009). Significantly positive effects of ARSs on student learning have been found in teaching first-year physics using peer instruction (Crouch and Mazur, 2001; also see Abrahamson, 1999). However, positive effects of ARSs have also been reported for conventional lectures (Boyle and Nicol, 2003).

One study evaluated the use of an ARS in several lectures of numerous disciplines, including Psychology, Computer Science, Biology, Chemistry, Statistics, and Philosophy (Draper and Brown, 2004). The study did not focus on the assessment of the ARS technology or the didactic approach used in conjunction with ARS alone, but on the combination of the two. Both students and lecturers appreciated ARSs for various reasons. For example, ARSs allowed to break up lectures in smaller parts and made lectures more interesting and interactive. Students like that ARSs enables contributing to the lecture anonymously without worrying about embarrassment. Especially when assessment questions are asked, anonymity is highly appreciated, even in groups in which students know each other well. Students also like seeing what the others think about the lecture, enabling them to find out how well they are doing in relation the the others. They value that ARSs provide feedback about how good they understand the material and allow them to identify problems rather early.

2.4 Deployments of Proprietary Classroom Communication Systems

Several CCSs have been developed and used. Some of these systems focus on shared note taking, others on ARSs, and few systems combine both CMC and ARS into one system.

Dufresne et al. (1996) report on the *ClassTalk* project, a PDA-based classroom communication system set out to support active and collaborative learning in physics classes. ClassTalk is based on a custom developed network of computers and PDAs. Up to four students can share one PDA. The PDAs are connected to adapter boxes that are distributed in the lecture hall. The adapter boxes are connected to the central server, a computer co-located in the classroom and administrated by the lecturer or other teaching staff members. On the server the lecturer can create tasks or questions that can be projected on the public wall and forwarded to the PDAs in order to have students respond to the questions. The lecturer can thereby decide whether each student or each group of students sharing one PDA can submit a response. When only one response per PDA is required, students sharing the PDA have to discuss and to agree upon which alternatives to choose.

Anderson et al. (2003) developed the classroom feedback system *CFS* allowing students to anonymously provide feedback. The system setup consists of students' devices, e.g., laptops, a lecturer's device and the overhead projector. The systems allows for nonverbal feedback or questions that are hand-written on PDAs using stylus pens. Unlike other backchannels, the *CFS* merged the lecture's content with feedback from students. With *CFS* the authors addressed the challenges of in-class interaction, i.e. feedback lag (a fast-paced lecture does not allow for student feedback), student apprehension (social inhibitions to partake), comment verbalization (trouble of communicating confusion in

words) and single-speaker paradigm (frontchannel as a limited resource which can only be used by one person at a time).

Ratto et al. (2003) developed the *ActiveClass* Project, a CCS for PDAs, that allowed students to publicly ask questions and enabled lecturers to electronically poll students using quizzes. The authors examined the use of *ActiveClass* from an ecological perspective taking the interplay between software as well as the concerns and demands of its users into account. For example, it was taken note of the lecturer's concern that using the platform would complicate her lecturing, since apart from lecturing the tool had to be managed. The concern was resolved by having teaching assistants monitor the communication on *ActiveClass*. Although regular student participation was quite low, the discourse in class became broader and student engaged in more question asking. Few especially insightful questions were asked, which the lecturer could not recall to ever being asked in any of his prior lectures. Furthermore, it is noted that also students who do not actively participate in an exchange benefit from such questions. *ActiveClass* supported question asking in the lecture, e.g., by providing anonymity and, thus, facilitating a shift of the focus from the questioners to the questions.

Kam et al. (2005) report on the system *Livenotes* developed for collaborative note taking and discussions in small groups of students alongside a lecture. The system enables students to annotate presentation slides with notes, either handwritten created with stylus pens or using keyboards. In an evaluation of the system in a lecture with 48 participants, distraction was an issue. The authors consider distraction as a result of usability issues that needed to be resolved and the rather low degree of familiarity with tablet PCs. Another issue expressed by participants was the lacking possibility to take private notes. Besides these negative findings, the authors found encouraging results regarding collaborative note taking. The lecture was covered more comprehensively, since students were able to collect notes more efficiently. The approach received great acceptance and participants recognized the benefits of collaborative note taking. However, the participants also recognized issues with synchronization and a considerable amount of cognitive overhead resulting from keeping track of the notes taken by the other group members so as to avoid redundant notes. Students using *Livenotes* took about twice the number of notes than students taking notes individually. Collaborative note-taking allowed students to split up note-taking, which increased the quality of notes as note-takers could take a break and were more focused in their turns. The notes involved personal reflection and internalization of the lecture material. However, tests conducted to assess learning of collaborative note-takers could not show better learning outcomes compared to individual note-takers.

Wilkerson, Griswold, and Simon (2005) present *Ubiquitous Presenter*, a Web-based note taking system that can be used with tablet computers or with browsers. *Ubiquitous Presenter* is an enhancement of the *Classroom Presenter* which has been developed by the authors and successfully used in many lectures. Both systems enable lecturers to present the slides and add handwritten notes to the slide (using stylus pens). The notes are then broadcast to all student. The lecturer may allow students to submit textual comments (typed on keyboard) as slide annotations (although, the authors plan to extend the system so as to allow for handwritten student annotations as well). As an extension of *Ubiquitous Presenter*, Simon et al. (2008) present *Noteblogger*, a Web-based system for shared (public) note taking as a means of student-to-student communication. A number of students is allowed to submit what the authors call noteblogs, hand-written notes on *Ubiquitous Presenter*, similar to those of the lecturer.[6] Students thus see the notes of the

---

[6]Although the term blog is used in the paper, it seems that noteblogs are merely notes. Thus, noteblogs should not be confused with blogs known from social media.

lecturer and, if permitted by the lecturer, the notes of notebloggers on the slide. Also, students can comment on the notes of the selected notebloggers. An evaluation in an introductory Computer Science course has shown that the use of *Noteblogger* positively affects the classroom experience of students.

Wessels et al. (2007) present a generic (Java-based) client-server system providing services for interactive lectures. The system is generic in the sense that service modules can be plugged into the server. For example, the call-in module allows students to enter questions or remarks into their mobile devices and send them to the lecturer. The lecturer can decide whether and when to answer questions. The feedback module enables students to transmit feedback to the lecturer at any time. The lecturer is steadily informed about the feedback which is summarized as bar charts. The feedback categories encompass motivational states of the students, presentation speed or difficulty of the lecture. The quiz module allows the lecturer to run quizzes during the lecture. Besides in-class synchronous communication, the system also provides modules for outside-the-lecture communication, both synchronous (chats) and asynchronous (forum). The acceptance of interactive lectures using the system was experimentally investigated in a lecture. Two groups received an interactive lecture using quizzes, two groups received a conventional lecture without any interactivity measures. A result of the experiment was that students accept interactive lectures. They state that learning does not deteriorate but instead tends to improve (although no significant difference in objective measurements of knowledge gain was obtained).

Harry, Green, and Donath (2009) report on the development and use of the system *backchan.nl*, a Web-based digital backchannel focusing on question asking and social selection of the "hottest" questions by the audience. Although the system has been developed for conferences, it is likewise usable in lectures. The contribution is interesting in that it elaborates more thoroughly on the rating and (time-dependent) ranking of comments (relevant for Section 6.5.1). Audience members can pose questions or submit any other kind of textual comments on the backchannel. Other audience members can then vote for or against these submissions. While all comments are displayed in the audience members' user interfaces, only the four best rated comments as determined by a time-dependent comment ranking are displayed at a public screen for all to see, including the presenter. After dealing with a question the presenter can mark the question. An analysis of the use at IT conferences showed that audience members appreciated the availability of a backchannel. However, in the interviews of users, also concerns were expressed. For example, presenters were concerned that bored audience members use the backchannel or their devices to distract themselves. And in fact, the authors observed audience members who switched between various online applications and rarely used the backchannel. Another concern of presenters is what the authors call "presentation preemption": Audience members may ask questions that are addressed at some later time in the presentation. Although anticipatory questions can be an indicator that the audience is with the presenter and that the presentation ignites interest, panelists may urge the presenter to answer the questions first. Another issue that the authors observed is that audience members manipulated the public ranking for fun.

Ebner (2011) presents *TwitterWall* a system built upon *Twitter* to promote interactivity in the classroom. *TwitterWall* facilitates communication among students and allows the lecturer to tap into the student exchange which took place in a specific time frame. The user interfaces of *TwitterWall* display the tweets[7] of a lecture as marked by predefined hashtags. Apart from indicating affiliation to specific lectures, hashtags are also used as a loose typing of tweets, e.g., into notices or ideas, according to which the

---

[7]For a concise overview of *Twitter* see Section 2.6.

stream of tweets can be filtered. Although students can use their own *Twitter* accounts, Ebner's system also allows students not registered on *Twitter* to partake. The system was used and evaluated in a lecture at TU Graz. Compared to former years, the lecturers were satisfied with the student participation (in average five tweets per lecture). Interviewed participants found the system useful and appreciated the persistent nature of the exchange, allowing to revisit the tweets later on.

Reinhardt et al. (2012) developed the Audience Response System *PINGO* with the goal to facilitate Peer Instruction (Mazur, 1997) in large classes. In Peer Instruction, frequent formative assessment of students in quizzes plays a central role. *PINGO* is realized as a highly scalable Web-based ARS that facilitates a conduct of quizzes in very large groups. The system can be used with any Internet-enabled devices, especially smart phones. The system was evaluated by surveying 438 students participated. The formative evaluation indicated that the system is well received and accepted by students.

Ebner, Haintz, et al. (2014) present an instant feedback system that enables student to communicate feedback along three dimensions: satisfaction, lecturing pace, and comprehension. Unlike conventional text-based backchannels students give numeric feedback by setting a slider for each of the three feedback categories. The provided feedback values are represented as pictograms that reflect the student's current states (e.g., "the lecture pace is ok"). The numeric feedback of individual students are aggregated to an audience-wide feedback presented to both the lecturer and the audience. To better reflect the audience's current states, the aggregated feedback is subject to aging, i.e., recent feedback is weighted more heavily in the aggregation than older feedback.

Haintz, Pichler, and Ebner (2014) argue that existing ARSs are too complicated to use. They surmise that feature richness entails the complexity of commonly known ARSs. Based on this assumption, they justify the development of a Web-based ARS that makes use of mobile devices students bring to lectures. The presented ARS aims at a lean and easy-to-use system. For example, the current systems only support multiple choice quizzes and bar chart summarization.

2.5 Social Media

In the early years of this millennium the use of the World Wide Web has changed significantly. Apart from the publication of content by the respective website owners, websites also started to provide functionalities that enable visitors to collaboratively create and share content themselves, the so-called *user generated content*. The OECD refers to user generated content as "i) content made publicly available over the Internet, ii) which reflects a certain amount of creative effort, and iii) which is created outside of professional routines and practices" (OECD, 2007, p. 11), thus distinguishing it from content published on business websites or as email or instant messages. Additionally, rich interaction among users is foregrounded by new forms of computer-mediated communication and possibilities to establish and maintain relationships in communities of users; the Web of documents has shifted towards a Web of users (Ebersbach, Glaser, and Heigl, 2011). The changed way of using the Web provides new technological and economic opportunities but also raises legal and social issues. Publisher and technologist Tim O'Reilly coined the term *Web 2.0* to capture and to refer to all these facets of the transformed use of the Web (O'Reilly, 2005; Ebersbach, Glaser, and Heigl, 2011).[8]

---

[8]In other words, the term Web 2.0 refers to the Web as a socio-technological system addressing the Web as a technological foundation for new kinds of applications as well as the social implications and challenges associated with the novel use of the Web in economic, legal, ethical, and also educational terms.

Kaplan and Haenlein (2010, p. 61) define social media as "[...] a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content". Kietzmann et al. (2011) understand social media as mobile and Web-based technologies to create highly interactive platforms on which users and communities share, co-create, discuss, and modify user generated content. Ebersbach, Glaser, and Heigl (2011) characterize social media as a part of Web 2.0 and Web-based software that aims at the creation and maintenance of social structures and interaction among users, e.g., sharing information and knowledge among users.

2.5.1 Classifications of Social Media

The descriptions above emphasize the three different dimensions according to which social media can be classified: creation of *information*, establishment and maintenance of *relationships* among users, and *collaborative* creation of medial content. All these dimensions are facilitated by communication. Figure 2.3 gives an informal classification of social media along these three dimensions (adopted from Ebersbach, Glaser, and Heigl, 2011, p. 39).
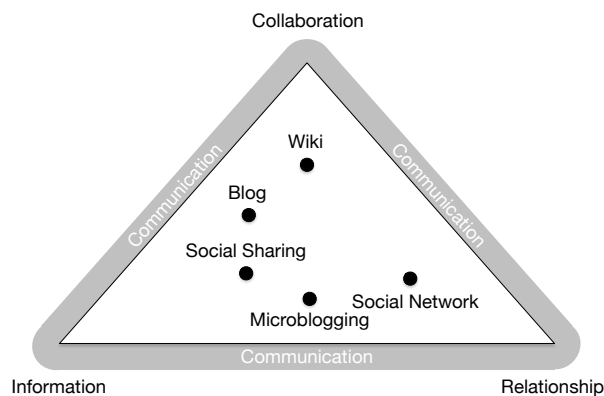
Figure 2.3: The picture shows an informal classification of social software along the three dimensions *Information*, *Relationship* and *Collaboration* (adapted from Ebersbach, Glaser, and Heigl, 2011, p. 39).

It should be noted that the authors provide an informal classification of social media that is neither systematically evaluated nor justified by other means, which is also emphasized by the authors. Moreover, differences between originally intended and actual use of social media that evolves over time makes the development of a clearcut classification of social media hardly possible. For example, according to the classification, microblogging is considered to address sharing information and fostering relationships to an equal extent but, in contrast, does not focus on collaboration. However, *Twitter* is known to be used as a medium for broadcasting disaster forecasting and prevention by ordinary citizens, which is clearly a collaborative task. Also, Twitter is used to create ambient awareness, letting receivers (e.g., family members) perceive closeness to a far-off sender of Tweets (e.g., another family member; see Section 2.6). Clearly, such a use of Twitter focuses on relationship to a greater extent than on mere information

23

sharing. However, the subdivision of social media along the three axes *Information*, *Collaboration* and *Relationship* is still useful for classifying social media.

While the classification of Ebersbach, Glaser, and Heigl (2011) focuses more on the purpose of social media, Kietzmann et al. (2011) investigated functional building blocks of social media. Accordingly, social media provide to a greater or lesser extent functionalities addressing the following aspects: *Identity*, *Conversation*, *Sharing*, *Presence*, *Relationship*, *Reputation*, and *Groups*.

Identity refers to the information users provide about themselves. The information may contain a user's real-name, age, gender, occupation, location as well as tacit information such as one's tastes, likes and moods that can be read out of conversations within social media.

Conversation refers to the extent to which users can communicate with each other. One characteristic of communication on a platform is what the authors term "conversation velocity", the rate of change in and direction of a conversation.

Sharing refers to the extent to which users receive and distribute content by others. Prominent sharing platforms are *Youtube* for sharing videos, and *Flickr* for sharing high-quality photos. The authors note that sharing may but does not necessarily have to lead to conversation or the establishment of relationships among users. Whether this is the case depends on the objective of the social medium.

Presence refers to the extent to which users are provided with information about users who are online and accessible on the platform. For example, *Facebook* and *Google+* highlight "befriended" users who are available for chat conversation. In contrast, users of a blog usually have no knowledge which users are currently accessible on the blog.

Relationship refers to the extent to which users can establish relationships among each other on a platform. On social networking platforms such as *Facebook* users can "befriend" each other, on *Google+* users can be added to circles, and *Twitter* provides followerships. Relationships on a platform determine which and how users can interact with each other. If a platform does not put emphasis on *Identity* it usually does not support *Relationship* to a great extent.

Reputation refers to the extent users can build up a standing within a medium that provides a kind of quality attribute of a user. For example, on an online market place reputation can be a measure of trustworthiness, on a forum reputation can measure frequency and quality of a user's comments.

Groups refer to the extent to which users can create communities and sub-communities by themselves on a platform. For example, *Facebook* allows the creation of special interest groups, e.g., fans of a certain film genre.

2.6 Twitter

Recently, so-called microblogs have gained much popularity as generic social news and communication media. Microblogging is a kind of one-to-many communication that is based on very short posts. The most prominent example of microblogging is *Twitter*[9] and, because of its de facto monopoly on microblogging, it is often considered as microblogging per se (Ebersbach, Glaser, and Heigl, 2011). Also, the use of *Twitter* as a digital backchannel has attracted much attention. This section considers the characteristics of microblogging, and specifically *Twitter*, that makes it appear suitable for the use as digital backchannels.

---

[9] http://twitter.com

A microblog post, on *Twitter* called a tweet, is a short text message that comprises a few words only. Specifically, the length of a tweet is limited to 140 characters.[10] The brevity of microblog posts lowers both the time and thoughts required from the user to create the tweet (Java et al., 2007) and thus promotes frequent, up-to-the-minute posting of posts. The communication microblogging provides is informal and the posts are appellative, expressive, coordinating and referencing (Ebersbach, Glaser, and Heigl, 2011). Microblogging is used to publish breaking news, not only by news agencies but also by lay persons (e.g., observers of natural phenomena and disasters) and can be politically motivated or promotional. In most cases, however, tweets are posted by private persons and often report on utmost trivial events (e.g., what the user is currently doing). Yet, the two properties, up-to-the-minute postings and triviality of posts, are what makes microblogging unique and compelling for so many users[11]. For example, a great number of trivial tweets enables followers to gather *ambient awareness*, i.e., a feeling of closeness and intimacy to a user (Kaplan and Haenlein, 2011).

Tweets are displayed in a so-called time-line in reverse chronological order, i.e., the most recently published tweet is displayed on top. The time-line thus emphasizes the latest tweets and draws attention from old tweets. Hence, microblogging-based communication can be thought of as ephemeral, although, technically speaking, microblog posts are persisted in a database and thus remain retrievable. Tweets can be assigned to topics by tagging, which *Twitter* realizes by so-called hashtags provided in the tweet. Hashtags are freely chosen keywords (tags) prefixed by a hash-sign ("#", e.g., `#elearning`) that allows tags to be distinguished from mere text. Hashtags are also a means to separate tweets into topic threads. Time-lines can be restricted to tweets containing certain hashtags and, thus, hashtags facilitate the creation and maintenance of shared communication channels similar to chat rooms. Often tweets directly address other *Twitter* users by including the addressee's user name preceded by the at-sign ("@", e.g., `@alex`).

The communication provided by *Twitter*'s Web interface is basically asynchronous, however realizing a kind of push-and-pull communication: The client is automatically notified if new tweets are available (i.e. notifications are pushed), but actually loading new tweets and updating the time-line requires the user's active request (the tweets are pulled, see Figure 2.4). However, other client software, e.g., mobile Twitter clients as well as Twitter's embedded time-lines,[12] also provide push-based updates of the time-line. Technically, these pushed updates are achieved by steadily requesting updates of the time-line in short time intervals, e.g., every few seconds, automatically in the background (commonly known as *polling*; see Section 4.4.1). However, this way of frequent updates of a shared time-line can provide a sense of synchrony among communicating *Twitter* users.

Tweets can be subscribed by following users, i.e., by registering for the reception of the tweets of posted by others. Subscribed tweets are displayed in the user's time-line. Following a user is non-reciprocal and does not require agreement of followed users. The follower relationships among users creates a social network that influences the flow of information among users. A user can feed subscribed tweets into her own

---

[10] In 2006 the development of *Twitter* has started with the goal to adopt SMS (short message service in cellular phone networks) as a Web service. In fact, tweets can still be submitted using SMS. Thus, like SMS, a tweet comprises at most 160 characters. However, 20 of the 160 characters are reserved for the sender's user name, leaving 140 characters for text (cf. Ebersbach, Glaser, and Heigl, 2011). A basic compatibility with SMS might as well be the reason why addressing another user, tagging tweets and re-tweeting is done within the tweet's text, although Web interfaces may allow for more convenient solutions.

[11] According to the *Twitter* website, 271 million monthly active users post 500 million tweets a day (`https://about.twitter.com/company`, last visited on 6 August 2014)

[12] `https://dev.twitter.com/docs/embedded-timelines`
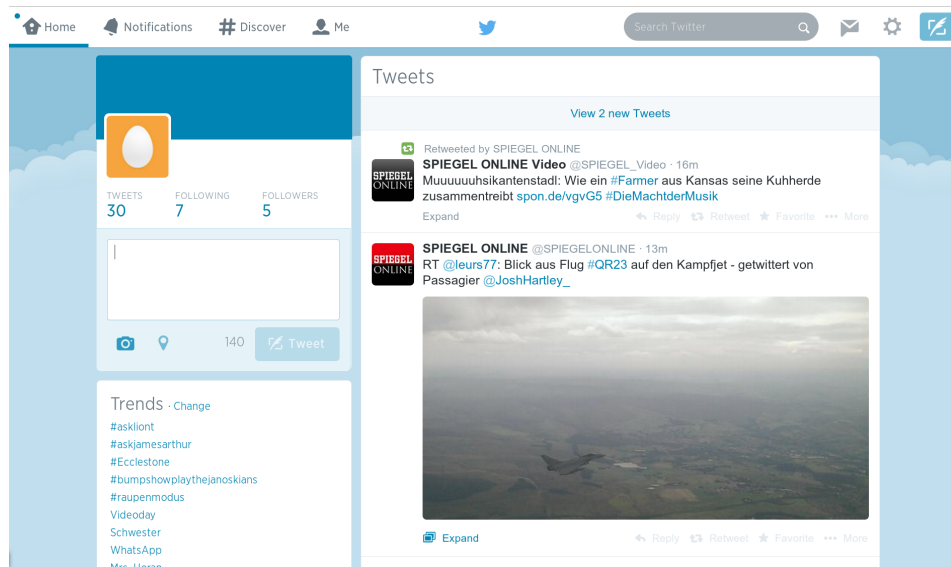
Figure 2.4: The user's dashboard on *Twitter*: The left column shows user-related information (user name, number of tweets, number of followers, number of users following) and the tweet editor. The center column shows the time-line of tweets. At the top of the time-line a notification link indicates that, since signing in, two new tweets have been published. The time-line is updated by clicking on the notification link. (last visited on 4 August 2014)

network of followers by re-tweeting them. In its original form, re-tweeting is a syntactic convention of quoting the original tweet. Accordingly, a re-tweet begins with RT (for "re-tweet") followed by addressing the original tweet's author and the original tweet given verbatim. Additionally, *Twitter* offers a re-tweet button for each tweet. In the time-line shown in Figure 2.4, both kinds of re-tweets are given: The top-most tweet is re-tweeted using *Twitter*'s re-tweet button; the second is a re-tweeted by the syntactic quotation convention.

Kwak et al. (2010) investigated differences between non-reciprocal follower-based networks provided by *Twitter* and reciprocal friendship-based networks provided by social networking sites such as *Facebook*. Although it is possible on *Twitter* that two users mutually follow each other, the reported analyses of follower networks indicate that following is not reciprocated in two-thirds of the cases. What Kwak et al. (2010) rather suggest is that following users is primarily not a means to engage in "social networking" but to connect to relevant information sources. This observation is also reflected in the fast diffusion of information on *Twitter*. According to Kwak et al. (2010), re-tweets typically reach 1000 users on average independently from the number of users following the original tweet's author. Re-tweets are further re-tweeted almost instantly by users up to four hops away from the original source. The fast and broad diffusion of tweets makes *Twitter* an interesting source for event detection. For example, Lampos and Cristianini (2010) investigated the use of *Twitter* as a tracking system for the flu spreadings in the UK by analyzing the tweets of UK citizens. *Twitter* is also investigated in its use by citizens to quickly and easily become aware of natural disasters such as wildfires or earthquakes (Sutton, Palen, and Shklovski, 2008; Earle, Bowden, and Guy, 2011).

2.6.1 The Use of Microblogging and Twitter for Educational Purposes

Several papers addressed potential uses of *Twitter* for educational purposes (e.g., Ebner and Schiefner, 2008; Ebner, 2009; Ebner and Reinhardt, 2009; Ebner, Lienhardt, et al., 2010; Junco, Heiberger, and Loken, 2011, and Kassens-Noor, 2012).

One study investigated the effects of encouraging the use of *Twitter* for educational purposes on student engagement and grades (Junco, Heiberger, and Loken, 2011). In total 132 students participated in the study. The findings are promising: Students who used *Twitter* showed greater engagement. The study also found a positive effect of using *Twitter* for learning on student grades, which was measured by the semester GPAs.[13] The results suggest that *Twitter* can be used to engage students in ways that improve achievement. However, the use of the semester GPA as a measure for student performance was confounded (Tess, 2013),[14] invalidating the causal relationship between *Twitter* use and performance established in the study.

A comparative study investigated the use of *Twitter* as an aide for student learning of a given subject matter (Kassens-Noor, 2012). The study particularly focused on learning contexts in which *Twitter* offers advantages over more conventional teaching methods. In the study students were asked to identify unsustainable practices in cities and suggest remedies. Students could choose among several options: One option was to use *Twitter* as the only medium for group communication, another option was to meet and work in a traditional in-class discussion and keep individual learning diaries. The final third option, which none of the participants chose, was that each student writes an essay on the topic. Students received credit points for the exercise. Tweets had to satisfy several requirements: Each tweet had to describe an unsustainable practice and suggest a remedy. Each re-tweet had to either add an unsustainable practice or refute that the practice described in the original tweet is unsustainable. Students were asked to tweet daily. The study investigated differences in knowledge construction and retention. The analysis indicates that *Twitter* is especially advantageous for collaborative knowledge construction as it allowed students to share ideas and suggestions beyond the classroom at any time. However, students using diaries showed more reflection on the subject. The author emphasizes that in some contexts *Twitter* may aid but in other contexts rather hinder student learning, depending on various factors, such as the course content, the assignment tasks, or the instructor's intent.

2.6.2 Twitter versus Specific Backchannels for Lectures

As Kassens-Noor (2012) suggests in her conclusion, a successful use of *Twitter* for educational purposes demands careful consideration of the context. In general, the choice of technology for learning is not as important as the pedagogy employed (Clark, 1994; Draper and Brown, 2004; Mazur, 2009). No matter what (mix of) media is used, argues Clark, the right instructional methods, not the media, are instrumental in ensuring a better learning. Clark argues that "if learning occurs as result of exposure to any media, the learning is caused by the instructional method embedded in the media presentation" (Clark, 1994, p. 26). Interestingly, Clark's argumentation is the gist of many debates on learning media and, recently, on the MOOCs (cf. Hieronymi, 2012).

---

[13]According to Wikipedia, GPA stands for Grade Point Average and refers to the averaging of all grades a student obtained in the marking period (see `http://bit.ly/1vNXVF1`, last visited on 4 August 2014).

[14]According to Wikipedia, a confounding variable is an extraneous uncontrolled variable that adversely influences the relationship between a dependent and an independent variable (see `http://bit.ly/1CTHoEx`, last visited on 4 August 2014).

As of a general-purpose backchannel like *Twitter*, it is quite unclear whether it might at all support the instructional methods chosen by a lecturer and the learning strategies of the students. Does the social model, especially the kinds of interactions a general-purpose backchannel promotes, fit mass lectures? Arguably, the followership model of Twitter is questionable in lectures where there are no good reasons to incite students to pay special attention to the posts selected by those of their peers they have decided to follow in the past. One could even argue that the follower model of *Twitter* might disrupt the audience cohesion desirable in lectures. Moreover, a backchannel like *Twitter* used for news and personal updates is likely to distract students from the lecture. Recall that the distraction general-purpose social media often cause in lecture halls is a serious concern (cf. Kraushaar and Novak, 2010). The use of a general-purpose backchannel like *Twitter* is, in lectures, surely not inexpensive, since posts irrelevant to the lecture have to be filtered out, nor efficient (cf. Clark, 1994). Even though peer evaluation is a promising approach for an educational backchannel, re-tweeting does not seem appropriate a form of peer evaluation for the lecture hall. Indeed, a repetition of posts is likely to be perceived as distracting noise during a lecture. Keeping repeating, especially on a backchannel complementing the lecturer's discourse, is surely not an appropriate way to stress the importance of an issue during a lecture. While re-tweeting might be excellent at starting viral effects that catch the attention of the many, viral loudening is not desirable in the lecture hall.

Several adaptions of microblogging to better fit the requirements of large-class lectures have been provided by *TwitterWall* (Ebner, 2011), e.g., anonymous tweeting and filtering tweets by hashtags. However, this thesis argues that a general-purpose backchannel like *Twitter*, how useful it might be in lectures if no better medium is available, is far from being a good choice for lectures.

2.7 Research Questions

As described in this chapter, Audience Response Systems (ARS) support the lecturer in question asking in a way that addresses many students in the lecture (as opposed to only a few as is the case for conventional lectures, see Karp and Yoels, 1976) and gathering feedback relevant for teaching and learning. Other systems are dedicated to (shared) note taking or electronic student-to-student communication. Besides proprietary systems, researchers have investigated potential uses of existing general-purpose social media for interactive learning, whose core values lie in providing social services and facilitating interaction among large groups.

This thesis seeks to conceive a classroom communication system (CCS) that integrates an Audience Response System (ARS) with shared note taking and student-to-student communication to support both lecturer-initiated as well as student-initiated activities in lectures. A key issue to be investigated is how to design shared note taking such that it does not distract the lecturer or students and that it focuses on lecture-relevant communication. Therefore, apart from considering structuring and filtering of the backchannel, concepts of social media may come handy. It seems reasonable to investigate means of incentivizing desirable behavior and promoting social regulation of the CCS, which can be assumed to be indispensable for a practical use in large-class lectures. In order to find out whether the concept fits the purpose, a prototype stable enough for experiments needs to be developed and its use extensively evaluated in practical lecture settings.

3

PRELIMINARY EVALUATION OF A FIRST PROTOTYPE

> This chapter is based on Pohl, Gehlen-Baum, and Bry (2012) and Gehlen-Baum, Pohl, and Bry (2011).

A first prototype has been developed to investigate strategies for remedying the drawbacks of general-purpose microblogging for computer-mediated communication in lectures. This chapter presents a first concept, its realization and a preliminary evaluation. The preliminary evaluation aims at investigating the prototype's usability at an early stage of development and the prototype's support of question asking by the audience.

3.1 The First Prototype

The prototype comprises a microblog-based backchannel and an Audience Response System. The two are combined in one view, allowing a simultaneous use of the backchannel and the ARS. In this way, the lecturer is able to skim over the comments while a quiz is running.

The user interfaces of the lecturer and the students differ (for the student's user interface see Figure 3.1, p. 30, for the lecturer's user interface see Figure 3.2, p. 31) in that the lecturer's user interface also provides quiz controls (e.g., selecting, starting and stopping quizzes) and gives real-time overview of student responses to a running quiz. The lecturer and students may register using some freely chosen pseudonym or by real names. A registered student has access to all lectures created on the CCS.

3.1.1 Concept and Realization of a Microblog-Based Backchannel

Microblogs have been shown to be appropriate backchannels, both in educational settings (Costa et al., 2008; Ebner, 2009; Ebner, 2011) as well as in academic conferences
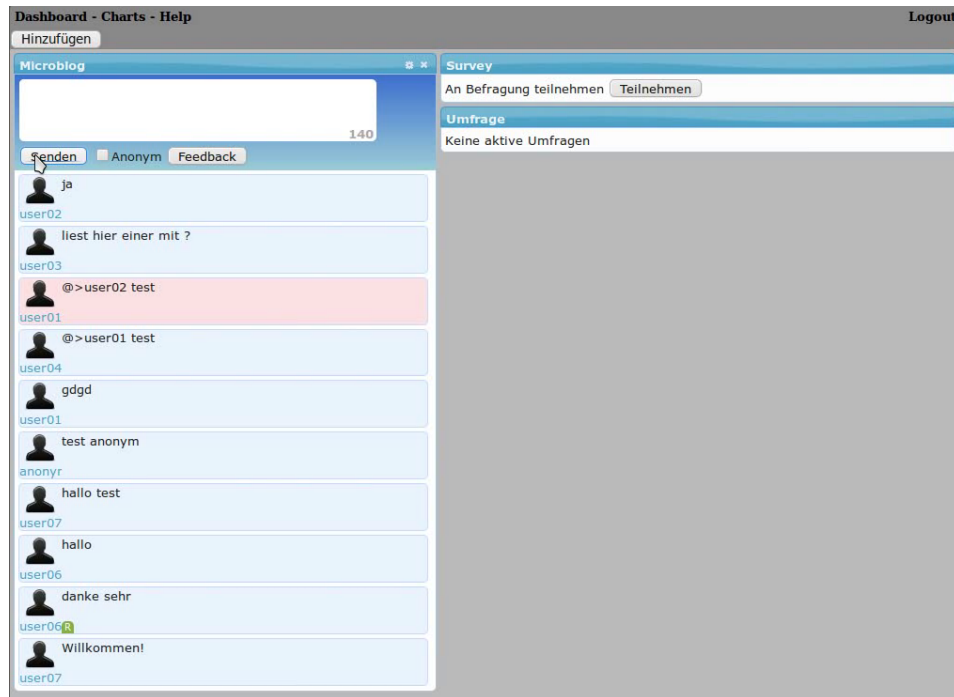
Figure 3.1: The student view of the first prototype in a user study. The left column shows the microblog, the right column the entry point to the survey and the Audience Response System (in the screenshot, no quiz is running).

(Saunders et al., 2009; Ross et al., 2011).[1] As Ebner (2011) points out, a tool for classroom interaction should be easy to use (in the sense that the lecturer and students should not waste their attention on fiddling with the software), permit anonymous partaking of students, provide for communication visible to all, and support the pedagogical approach chosen by the lecturer.

As discussed in Section 2.6.2, not all educational requirements are fulfilled by a general-purpose backchannel such as *Twitter*. However, microblogging provides a viable basis for a backchannel that can be adapted for large-class lectures. Being restricted to a few words only, microblogs can be written and read rather quickly during lectures. The brevity also encourages students to get to the point quickly. Furthermore, it can be assumed that many students are familiar with using microblogs, which may simplify the introduction of a backchannel in lectures.

The prototype realizes a microblog-based backchannel comparable to *Twitter*. Several adaptations have been made to provide for more communication modes, e.g., anonymous communication, that may be useful for an educational backchannel. The main characteristics of the first prototype's backchannel are as follows.

As in *Twitter*, microblog posts are limited to 140 characters. Students may address the lecturer or other students publicly using *Twitter*'s @-syntax (see Section 2.6). Publicly addressed microblog posts are visible to all participants. In addition to public addressing by using the @-syntax, the protoype's backchannel also provides a form of

---

[1] Also see Liz Lawley's blog post "Learning From (and About) the Backchannel" dated April 1st, 2004 (http://bit.ly/19EFgHb, last visited on February 18th, 2015).
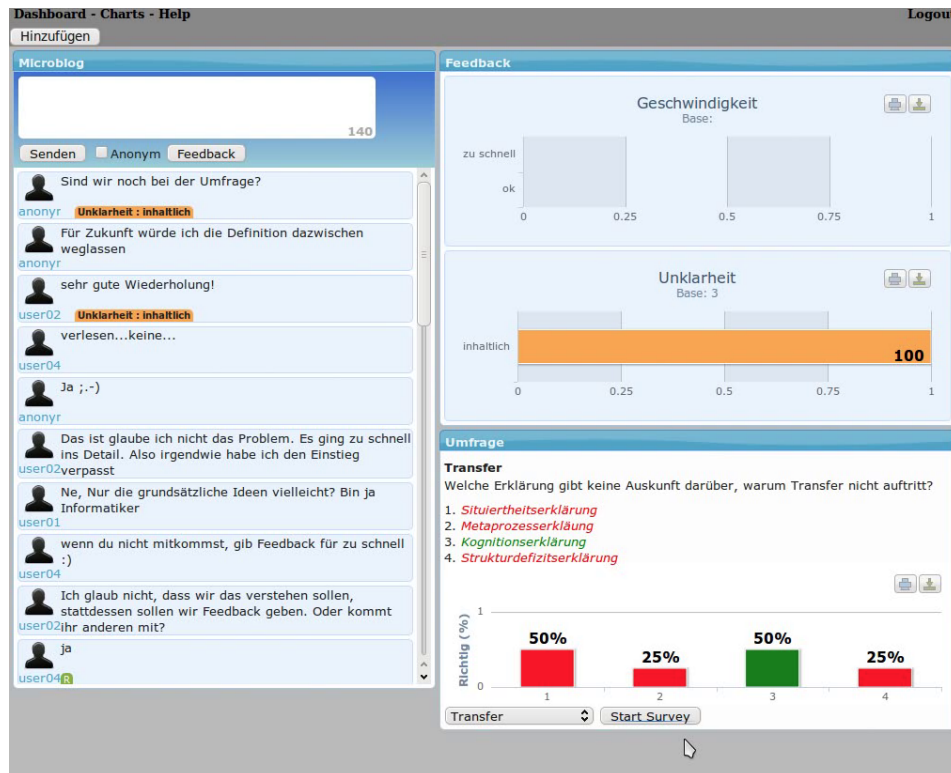
Figure 3.2: The lecturer's view of the prototype in a user study. The left column shows the microblog containing two feedback comments that indicate comprehensions problems (see the comments with orange markings saying "Unklarheit: inhaltlich" which indicates a difficulty in understanding the topic). The right column shows the summary of feedback comments as a chart. Below is the Audience Response System.

private addressing. Private microblogs are only visible to the sender and the addressed participants. A recipient is addressed privately through the use of the syntactic construct `@>` followed by a recipient's user name, e.g., `@>user02`. A microblog post may address more than one user by enlisting all participants in the post. If at least one recipient is addressed privately though, the post is considered to be private and only delivered to the privately addressed recipients. Also, posts can be submitted anonymously, not showing the sender's user name and her avatar along with the anonymous post. In this way, the prototype provides forms of weak and strong anonymity. That is, the backchannel allows a user to be linkable by her pseudonym (and even to be identifiable if she chose her real name as her pseudonym) or to be not linkable (and thus not identifiable) at all; compare Table 2.1. In the prototype, anonymity is not enabled by using a syntactic construct in a text but by selecting anonymity in the editor view. Responding to posts by means of re-posting is discouraged. Instead, a response explicitly references another post. Apart from free-form communication, students may annotate their posts as feedback to the lecturer. The feedback categories are predefined. In the prototype feedback categories are the pace of lecturing and degree of difficulty of the lecture. Feedback-related posts are not only shown at the time-lines but are also aggregated and displayed as a histogram on the lecturer's view (see Figure 3.2). The prototype enables students to give feedback

on the lecture's pace and "wave a flag" if the lecture is perceived as being too difficult.

The backchannel communication is realized as a truly synchronous communication, meaning that a submitted post is promptly visible to all addressed participants. As in *Twitter*, the posts are displayed in a descending chronological order. That is, the newest post is displayed on top, the oldest at the bottom. The post editor is placed above the time-line of posts. Anonymous posts show a default avatar and the user name is replaced by `anonymous`. Private posts are highlighted by a special background color. Responses are highlighted by a green `R`-symbol (`R` for reply) shown next to the author's name in the post. By moving the mouse over the symbol the post responded to is shown in a popup.

3.1.2 The Audience Response System

The second main component is the Audience Response System, which supports multiple-choice quizzes. The lecturer selects from previously prepared quizzes in a drop down list and starts the quiz. In the prototype, the status of a quiz, running or stopped, is provided. When a quiz is started, the students' views display the input form for the quiz. When the quiz is stopped, a final summary is prepared and publicly displayed to all users. The summary is shown to the students until the lecturer closes the quiz. Figures 3.2 and 3.1 show where the Audience Response System is displayed in the respective views. The quiz views during a run are presented in Figure 3.3.

3.2 Preliminary Evaluation of Use

A preliminary evaluation of the first prototype took place at an early stage of development in December 2010. It had two purposes: (1) to assess how far the adjusted backchannel already met the requirements for classroom backchannels; and (2) to indicate the direction of the prototype's further conception and development. Therefore, usability and perceived effectiveness of the prototype were measured. The evaluation was to help answer the following questions:

How do the participants assess the prototype's usability? The lecturer's and the students' attention are to a great extent directed towards the lecture discourse. A use of CMC in lectures is likely to differ from scenarios outside a formal learning setting (outside a lecture) in that the lecturer and the students cannot pay full attention to the backchannel.

Does the backchannel invite participants to ask questions? One of the main purposes of a backchannel is likely to be question asking. It is interesting to investigate whether students' activity on the backchannel influences the quantity of question asking.

Does the questions' quality in a backchannel-supported lecture differ from that of a conventional lecture? Another aspect of question asking worth investigating is if questions that are asked on the backchannel are more or less elaborated than questions asked face-to-face.

3.2.1 Usability

Usability expert Jakob Nielsen defines usability as "[...] a *quality attribute* that assesses how easy user interfaces are to use" (emphasis taken from the original).[2] According to

---

[2]Nielsen, Jakob, "Usability 101: Introduction to Usability", blog post dated January 4th, 2012, `http://bit.ly/1jNaIWu` (last visited on February 19th, 2015).
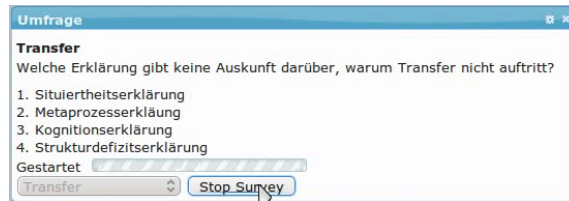
The lecturer's view of a started quiz
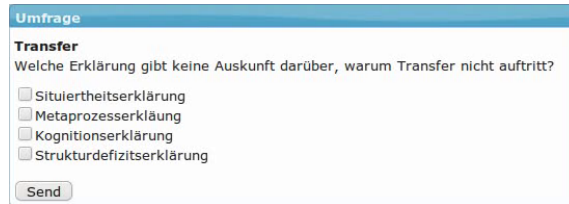
A student's view of a started quiz

The students' final summary of the quiz

Figure 3.3: The Phases of Quiz in the Audience Response System View.

Nielsen, usability consists of five quality components. *Learnability* refers to how easily users can accomplish simple tasks the first time they use a system. *Efficiency* refers to the time needed to perform a task. *Memorability* refers to the ability to reestablish proficiency at using a system after not having used it for a while. *Errors* refers to the severity of errors that users can make and ease of recovering from such errors. Finally, *Satisfaction* captures how pleasant a system is to use. Instead of getting acquainted with a system at first exposure, Tullis and Albert (2008) stresses that it is sometimes more important to consider *Learnability* as the time needed to becoming proficient in using a system. For example, word processors satisfy *Learnability* by allowing more proficient users accelerate tasks by using keyboard shortcut instead of lengthy navigation through menus. Thus, it seems reasonable to adopt user interfaces from well-known applications such as *Facebook* or *Twitter* for learning environments, since they are already familiar to most learners. This allows users to more quickly grasp the way to interact with the application and to accomplish the tasks without much need for settling in.

3.2.2 Effectiveness of the Backchannel for Question Asking

Not only the quantity, but also the quality, of questions is important for learning (King, 1994). In ibid. a distinction is made between three kinds of questions: factual, comprehension, and integration questions. Table 3.1 summarizes the descriptions of questions and their distinctions regarding knowledge.

Type of Question	Description
Factual	Asks for recall of facts or other information explicitly covered in the lesson. *Encompasses knowledge restating*: Simple statements of fact or information gleaned directly from the lesson or prior knowledge.
Comprehension	Asks for a process or term to be described or defined. *Encompasses knowledge assimilation*: Definitions, descriptions, and other material paraphrased in student's own words.
Integration	Goes beyond what was explicitly stated in the lesson, connects two ideas together, or asks for an explanation, inference, justification, etc. *Encompasses knowledge integration*: Makes new connections or goes beyond what was provided in the lesson – explanations, inferences, relationships between ideas, justifications, statements, linking session content to material from outside the lesson (prior knowledge and personal experience).

Table 3.1: Kinds of Questions (by King, 1994, p. 350, with minor adaptations)

3.2.3 Method

Nineteen participants from different fields of studies (mainly Educational Science and Informatics but also Psychology and German as foreign language) and different semesters participated in this study. The mean age of the participants was M = 26.32 years (SD = 3.33).[3] Five (26.32%) participants were female and fourteen (73.68%) were male. Since few is known about the students' prior knowledge in large lectures, a pretest has been omitted for this preliminary study and an experimental design with one control group has been chosen. The participants were picked randomly from the different fields of study. Every participant in the experimental group was assigned to a work place equipped with a PC, resembling traditional computer-based learning environments. To keep the setting of the control group comparable to the experimental group, each participant of the control group was also assigned to a PC work place. To avoid technical problems with the prototype, e.g., due to the amount of data to be processed, the experimental group (N = 14) was subdivided into three subgroups with five and four participants, respectively. The control group was designed to have the same size of five participants in order to ensure that the participants were not inhibited by a bigger group size.

A 30 minutes presentation was held in front of each group. The presentation was about two educational subjects in two talks, which were presented for ten to fifteen minutes each. The first talk was about the notion of transfer: the definition, the reasons why it occasionally fails to take place and the ways to foster it. The second part was about the research on expertise: the definition and a short overview of this field. In the experimental group discussions were supposed to take place only on the backchannel.

---

[3]For the sake of brevity, and by convention, M is used to denote the mean of a sample, SD is used for the standard deviation of a sample.

In the experimental group the participants were logged in to the prototype using pseudonyms prior to each trial. As the study was conducted at an early stage of development, several features were not completely implemented or were unstable and thus not available for the study. The backchannel capabilities as described in Section 3.1.1 were available for use. Although backchannel messages could also be rated, the rating scores of the messages were not aggregated and rendered along with a message and neither were taken into further consideration. The participants could assign messages to predefined feedback categories as described above. No filtering of the backchannel was provided to the presenter. Nevertheless, the presenter tried to keep track of the backchannel and orally answer as many questions as possible during the talks. The ARS was also available for use. In the experimental group the participants were informed about the state of implementation and the features that could be used. The participants of the control group listened to a traditional presentation without a CCS running. The participants of the control group were invited to raise hands in order to participate in the presentation.

3.2.4 Measurements and Instruments

The study consists of analyzing the communication that took place in experimental group (on the backchannel) and in the control group (face-to-face). Furthermore, the participants of the experimental group were asked to fill out a survey at the end of the presentation.

The backchannel communication was exported to log files and the statements of the control group were transcribed accordingly. All recorded and transcribed statements were coded twice by distinct coders into *question, answer, off-topic,* and *feedback* to the lecturer. Since no other kinds of statements ended with questions marks, questions could be easily identified syntactically. A response to a question could be determined by its adjacency to the respective question or by context. Feedback to the lecturer could be extracted in a similar way. In a second step, statements belonging to questions were further subdivided according to their content into the classes given in Table 3.1 on page 34. The inter-rater reliability of the coding yielded $\kappa = 0.74$.[4]

The questionnaire contained fourteen multiple choice questions (6-points Likert scale)[5] and three open items[6] (see Appendix A.7) to measure usability. Since the participants were using the prototype for the first time, assesssing the time needed to get acquainted with the prototype seemed to be more important than other usability aspects such as the number of errors the participants made. *Learnability* was measured by the participants' self-evaluation of the time needed to get acquainted with the prototype. Usability was measured by answers to questions such as "the application's user interface is clear and concise". The questions on usability were tested for reliability ($\alpha = 0.91$). In addition, the three open items were meant to give a deeper insight in what the participants liked or missed. The answers were sighted by two coders and certain categories like questions to the lecturer were counted ($\kappa = 0.92$).[7] The attitude towards question asking was measured by answers to seven questions such as "I felt unsure when raising a question" ($\alpha = 0.77$).

---

[4]For a brief introduction to inter-rater reliability see Section 5.2.1.

[5]The gradations and mappings to numerical values are as follows: strongly agree (6), agree (5), tend to agree (4), tend to disagree (3), disagree (2), strongly disagree (1).

[6]Open items ask participants to answer in free-form text.

[7]See Footnote 4.

3.2.5 Procedure

Each trial took place during a 30 minutes presentation. All participants were assigned to their places. In the experimental group the experimenters logged each participant in to the prototype prior to a trial. The participants in the experimental group received a brief introduction to the prototype and how to use the different functions. Furthermore, the experimenters provided all participants with a short summary about the outline of the study. The places of the control group were equipped with pen and paper. The experimenters instructed participants of the control group to raise hands in order to pose a question during the presentation. The presentation was given in two talks: The first talk was about transfer, the second about expertise. After each talk both groups were invited to ask questions. Immediately after the talks, which were held comparably in contents and style, the participants were asked to fill out the survey.

3.2.6 Results

The participants of the experimental group (using the backchannel) made nine times more contributions to the presentation than the control group (Table 3.2). Only few microblog posts were off-topic (M=2.5, SD=2.82). A considerable amount of the microblog posts was feedback (M=5.42, SD=5.57). Most microblog posts were questions (Table 3.2). This is also reflected by the positive attitude towards asking questions on the prototype (M=4.56, SD=0.87). While few questions were about using the prototype, the majority of questions referred to the content of the presentation. The participants' questions in the control group obviously did not refer to the use of the prototype, but to the presentation's content only. The participants in the control group only asked questions after explicitly being invited to do so by the presenter.

	Comments M (SD)	Questions M (SD)
Experimental Group	14.8 (10.5)	3.8 (3.7)
Control Group	1.6 (2.1)	1.2 (1.3)

Table 3.2: The mean number of comments and questions raised by the experimental group (with CCS) and by the control group (without CCS).

Table 3.3 shows «the distribution of questions regarding the classification of questions according to King (see Table 3.1).

	Factual M (SD)	Comprehension M (SD)	Integration M (SD)
Experimental Group	1.1 (1.4)	0.7 (0.9)	0.4 (1.4)
Control Group	0.4 (0.5)	0.2 (0.7)	0.6 (0.6)

Table 3.3: The mean number of questions as categorized by King (1994) raised by the experimental group (with CCS) and by the control group (without CCS).

As to usability, a mean of M=4.42 (SD=0.78) was measured, which is between "tend to agree" and "agree" and indicates that the participants found it sufficiently easy and intuitive to work with the prototype. Every participant submitted at least one comment during the presentation. Also, the participants' responses to the open item questions indicate a general usability of the prototype. Notably, the time the participants stated to be necessary in order to get acquainted with the prototype highly varied. While the

average was M=8.79 (SD=8.34) minutes, the minimum time stated was one minute and the maximum was 30 minutes. In the open items the participants appreciated the possibility to directly send feedback to the lecturer (mentioned eight times). The process data showed that the experimental groups gave feedback to the lecturer more often than the control group (M=5.43) which gave only short feedback, e.g., "ok" for acknowledgment (M=0.40). Also, the quiz functionality to assess current knowledge was appreciated by the participants of the experimental groups (mentioned five times). However, they missed the easy-to-use communication known from other communication tools. Two participants suggested to integrate the presentations slides into the prototype. Three participants also noted the risk of getting distracted by the use of the prototype.

3.3 Discussion

This chapter reports on a preliminary study that mainly focuses on first impressions of the prototype's use in a presentation. The study was set out to investigate how participants get along with the prototype on first use and how the backchannel supports question asking during presentations. Due to the small number of participants the investigation is confined to a descriptive analysis.

3.3.1 Findings on Question Asking on the Backchannel

According to King (1994), the three question types, factual, comprehension and integration questions, can be ordered by the mental effort required by the learner to create these questions. Factual questions are the easiest type, whereas integration questions the most difficult type of questions. Thus, factual questions produce the weakest learning effect while integration questions produce the strongest effects on learning (King, 1994). However, it appears that students mostly limit themselves to asking factual and comprehension questions. The findings corresponds to King's observation. In the experimental group the majority of questions were factual questions, fewer were comprehension questions, and the fewest integration questions.

The high mean of integration questions in the control group is an unexpected result. However, the result can be explained by the following two circumstances: (1) In a smaller sample a mean is more susceptible to outliers. An integration question asked in a mostly passive group yields a greater mean of integration questions than in a more active group. (2) In the control group, most participants, i.e. three out of five, were students of Psychology and Educational Sciences and about to finish their studies. Since integration questions highly depend on prior knowledge, the described constellation of the group could explain the unexpectedly high mean of integration questions in the control group. It could be the case that the topics chosen for the presentation in combination with the participants' prior knowledge had an influence on the participants' quality of question asking.

3.3.2 Findings on Usability

The average time of nine minutes needed to get acquainted with the prototype during the first talk is rather short, which can be attributed to the similarity to well-known platforms such as *Twitter*. However, the variability around the average could be a ramification of the participants' different notions of "becoming acquainted with the prototype", their backgrounds, or their levels of expertise with new media.

Although the analysis of the questionnaires suggests that the participants liked using the prototype, in the open items, the usability has been remarked to be in need of

improvements. This remark did not come up unexpectedly, though. The responses to the Likert-type questions suggest that most participants did not find the deficiencies too disturbing. The observation that every participant posted at least one message during the presentation indicates that the prototype is, in general, convenient enough to use and has a positive influence on participation.

Although several remarks made about usability issues refer to limitations of the early prototype, it may be the case that a CCS may be less convenient to use than general-purpose communication media, if this is necessary for in-class interactions that conduce to learning. Note that an "inconvenient" use of an application that results from the application's purpose would not be considered as a usability issue. For example, time-on-task, i.e., the time required to accomplish a task, is usually a usability metric to be minimized. However, when using courseware for learning, it may be better to prevent learners from rushing through the learning material but spending more time on specific learning tasks; other examples of "slower may be better"-approaches are games (Tullis and Albert, 2008). Whether learners are satisfied with user interactions that may require more time but are rather more relevant for learning needs to be investigated. Despite a more difficult use of a CCS, the learner may develop satisfaction due to improved learning outcomes and positive feedback. However, one needs to clarify whether satisfaction and impressions of usability may change in the long run. The remarks regarding the quiz and the feedback functionalities of the prototype suggest that the effects of quizzes on active participation also apply to the prototype.

To sum up, the users seemed to get along well with the prototype, even with the first (yet unfinished) prototypical implementation. No fundamental criticism was given, but further functionalities encouraged, e.g., the integration of presentation slides in the prototype. The suggestion was considered and accepted for the next stage of development. Furthermore, the risk of becoming overwhelmed by the backchannel has been noted to require strategies that help reducing information overload.

The comparison of the experimental to the control group indicated positive effects of using the prototype on the activity of the participants. Thus, more questions were raised on the prototype than being asked by the control group. The same applies to feedback addressed to the lecturer. It should be noted that the participants of the control group only asked questions after invitation, maybe to avoid interrupting the lecturer. In how far the increase of questions on the prototype has impact on the learning results needs to be investigated in future studies. Thus, it seems that an increase of activity positively influences learning success and, therefore, motivation (Deci and Ryan, 1993).

As participants stated, however, raising more questions may also contribute to distraction and information overload. This issue is notable, particularly when keeping in mind that the formative evaluation was carried out with small groups of four to six participants. If distraction by too many questions is already considered to be a potential problem in such small groups, further adjustments are necessary.

Part II

1.0

INTRODUCING BACKSTAGE

> This chapter is based on Baumgart et al. (2012), Pohl, Bry, et al. (2012) and Bry and Pohl (2014).

The formative user study described in the previous chapter suggests to investigate the following two issues:

1. How to keep the backchannel communication aligned with the lecture's discourse.
2. How to avoid learners be confronted with too many messages on the backchannel.

The first question addresses the crucial issue of disruption engendered by switching attention between the lecture discourse and the backchannel. The lack of reference points in the prototype's time-lines made re-orientation on the backchannel difficult. The participants of the user study already suggested that an integration of lecture slides into the user interfaces can be part of a solution. The second question is more difficult to answer, since too much communication on the backchannel can have more than one cause. Possible causes are the quality of the lecture, unstructured rendering of the communication, and undesirable communication, e.g., not related to the lecture.

This chapter presents the fundamental modifications made to the first prototype. The resulting revised version is the CCS *Backstage*[1] provides structured backchannel communication on the basis of collaborative annotating of the lecture slides using microblog posts that belong to certain predefined categories. Furthermore, *Backstage* provides an extensible and improved ARS compared to the first prototype.

4.1 The Backchannel Reconsidered

An integration of the lecture slides opens up possibilities to reduce information overload and to increase efficiency of communication on the backchannel. Many improvements can be centered on integrating lectures slides into the CCS and realizing backchannel communication as a means for collaborative annotation of slides.

---

[1] http://backstage.pms.ifi.lmu.de

From an educational viewpoint, Suthers and Xu (2002) have stressed that in face-to-face learning allowing participants to reference and manipulate so-called external artifacts, e.g., documents, pictures and other objects mentioned and used during the lectures, is important. They argue that students' interactions with external artifacts are essential activities in learning. Therefore, referencing artifacts should also be possible in computer-mediated discourses. The primal external artifacts of today's lectures are presentation slides. A CCS for (large-class) lectures should therefore support a collaborative annotation of presentation slides. Contextual annotations can be achieved by requiring from a learner to position her backchannel post on the slide very much like one would write, on a paper slide, her question or comment close to the place where a text or picture prompted that question or comment (Marshall, 1997). Public annotating of learning artifacts has the additional advantage that it may improve the lecture as a learning environment. In empirical studies and self-reports, participants greatly appreciated the collaborative annotation of learning artifacts (Suthers and Xu, 2002; Nokelainen, Kurhila, et al., 2003; Nokelainen, Miettinen, et al., 2005).

4.1.1 Improving Communication Efficiency

Besides educational objectives, annotating slides is also a means to improve communication efficiency and thus to reduce information overload. To improve communication efficiency, Huber (1982) suggests two mechanisms, routing and summarizing of communication.

Routing refers to the automatic separation of communication into relatively few so-called organizational units. The separation of communication by referring to these organizational units reduces information overload in that participants are provided only with the communication addressed to the organizational unit to which they belong. Apart from courses and lectures, lecture slides are such organizational units. Participants are only provided with the backchannel communication that refers to the slide they are currently viewing. As a consequence, a kind of prioritization of communication and a message delay for non-prioritized CMC messages referring to other slides is achieved. Prioritization of communication helps avoid confronting participants with backchannel communication currently not of interest. Among participants viewing the same slide, synchronous communication can be realized, whereas communication between participants viewing different slides is asynchronous, e.g., a kind of push-and-pull communication as provided on *Twitter*'s Web interface (Section 2.6), indicating that new messages are posted on other slides but require explicit action to be retrieved by switching to another slide.

Summarizing is a mechanism to condense communication to its relevant parts without changing its meaning. One way of summarizing is to limit or to structure the form of a participant's input to a communication environment. Limiting the length of CMC messages, as in microblogging, is an instance of summarizing. Another approach is to provide alternative inputs for communication that does not necessarily require full text messages. For example, approval or rejection of communication can be expressed by ratings. Summarizing can also be accomplished by human moderators who keep track of the communication, select relevant messages, e.g., to forward them to the lecturer, and delete irrelevant messages (Hiltz and Turoff, 1985). A similar summarization can be automatically achieved by aggregating communication in a quantitative overview.

As a third mechanism, Hiltz and Turoff (1985) argue for harnessing social control as a means to improve communication efficiency. The authors argue that emergent social norms can play a strong role in preventing undesirable behavior and that the potential of social norms should be taken into account in the design of communication

media. As the authors put it, "to design a CMCS [Computer-Mediated Communication System] is ultimately to alter a social system" (Hiltz and Turoff, 1985, p. 681). The mechanisms of social regulation are employed in social media in a multitude of ways. It is thus reasonable to take elements of social media into account (see Section 2.5) in order to prevent information overload and to make communication more efficient. Examples of such elements are collaborative filtering through rating and ranking of CMC messages (see Chapter 6) and social incentivization of desirable behavior through online reputation of participants (see Chapter 7).

4.1.2 Reducing Information Entropy

Apart from information overload, Hiltz and Turoff (1985) introduce the term information entropy to refer to the variety of CMC messages with respect to message types and content. An insufficient organization of a great variety of CMC messages may as well lead to confused and overwhelmed participants. While routing by lecture slides contributes to a reduction of information entropy, it is reasonable to additionally convey to the participants which content is desirable on the backchannel.

Assigning backchannel posts to categories on the prototype has been an optional activity. Without having to read a post, categories provide valuable content-related information about the post. Introducing generic post categories and prescribing that posts have to be assigned to categories seems to be a reasonable measure in order to reduce information entropy. A middle-way between a strongly structured and free communication is achieved if a limited number of generic, predefined categories, like *Question*, *Answer*, *Remark*, *Too Fast*, and *Too Slow* are offered as post categories. Predefined post categories also implicitly convey to the students what sort of communication is desired. In contrast, unrestricted communication leaves room for off-topic messages. Hence, categorization can be a first step to counteract off-topic messages.

4.2 Overview of Backstage

The user interfaces for students and for lecturers on *Backstage* strive to draw no more attention from participants than necessary. Care has been taken to provide for clear and concise user interfaces strongly focusing on the lecture slides as shared artifacts. A few unobtrusive colors are used to not distract the participants and to direct attention on the CCS to the lecture slides. The user interface shows only those elements that are currently necessary. Further interface elements, e.g. for submitting a post, are only shown on demand. Thus, the user interface follows the principle of quiet design which refers to a user interface design that neither seeks attention of learners for the sake of itself nor carelessly interrupts the activity of learners (Peters, 2014). Quiet designs are particularly preferable, since "one of the biggest ways interface design can contribute to better learning is by getting out of the way" (Peters, 2014, p. 71).

4.2.1 The Backchannel

Figure 4.1 shows the student user interface on *Backstage*. The backchannel time-line can be found in the left column of the view. The center column is used for the display of and navigation through slides. Annotations, symbolized by speech bubble icons indicating the post category, e.g., *Question* or *Answer*, are placed on the slides. The right column view gives the lecturer a virtual presence and provides space for awareness-related information useful during a lecture.
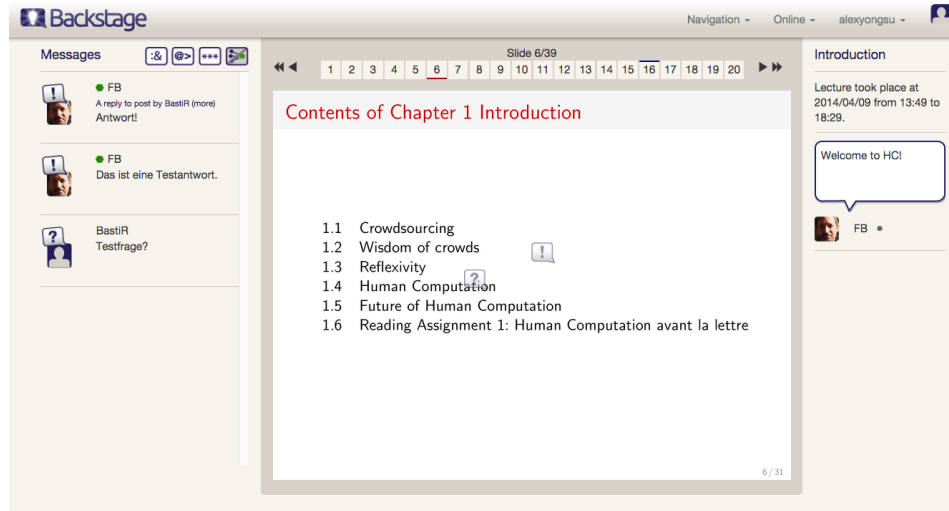
Figure 4.1: The student view on *Backstage*. The left column shows the backchannel time-line. The center area is devoted to the slides. In the picture two annotations are placed on the slide. The right column is used for awareness (not shown) and presence of the lecturer.

The lecturer's user interface provides a view of the slides in a separate window that can be used to project the slides on a public wall. The additional view obviates the use of another presentation tool for this purpose. In the separate window, the backchannel communication is not displayed along with the slides so as to avoid distraction of students who for some reasons decide to not participate in the backchannel.

4.2.1.1 The Annotation Workflow

Participants can annotate slides with backchannel posts (see Section 4.1). For publishing a post, a participant first references the relevant part on the slide by clicking on the slide at a suitable location for the post. As the second step, the participant selects one of the category icons: The question mark stands for the post category *Question*, the exclamation mark for *Answer*, the light bulb stands for a generic *Remark* and the arrow symbols for *Too Fast* and *Too Slow*, respectively. As the final step, the participant writes the post and publishes it. On the lecture view, after clicking on the slide to place the post, the post editor pops up. The editor enlists the possible post categories the participant can choose from (Figure 4.2).

After selecting the appropriate post category, the post editor shows a text field in which text can be entered (Figure 4.3). The process of submitting a post on *Backstage* reflects the communicative act of commencing an utterance with context information like: "I have a question about the fourth statement on this slide".

A student may restrict the recipients of a post and decide whether her user name is shown along with the post or is hidden. These adjustments are referred to as communication modes. As in the first prototype (see Chapter 3), communication modes are expressed using predefined keywords, or modifiers, within the text of the post. Public posts, which are visible to all, are the default mode and therefore require no modifier. The modifiers for public and private addressing of participants are those described in Section 3.1.1, i.e., @ and @>, together with the additional modifiers *** and :&.[2] The

---

[2] Both syntactic modifiers were chosen arbitrarily and can easily be changed.
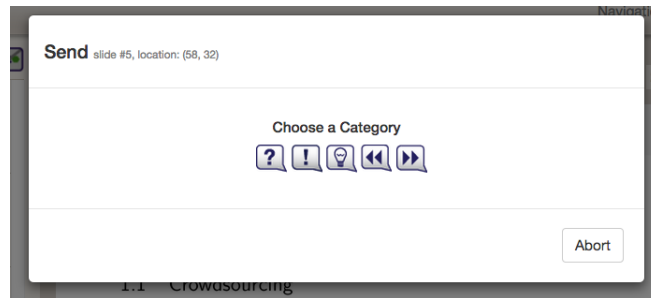
Figure 4.2: After clicking on the slide the microblog editor appears and requests the selection of a category for the post.
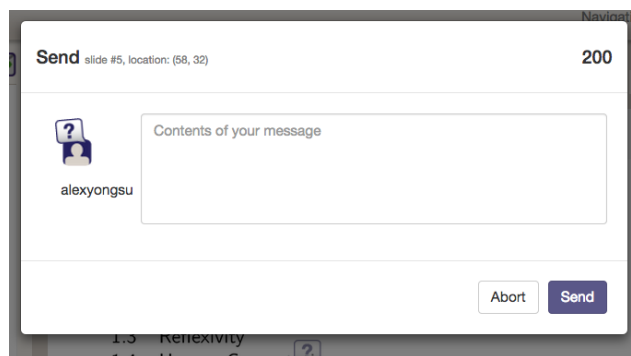
Figure 4.3: After selecting a category the message can be entered and submitted.

modifier ∗∗∗ serves for perfectly anonymous communication in which the sender is neither identifiable nor even linkable (see Section 2.2.1.1). The modifier :& serves for private note taking (private notes are only visible to the notes' author). This functionality is added to offer an alternative to simultaneously working with several media, viz., *Backstage* and paper printouts for note taking.

The backchannel can be configured by the lecturer at any time. She can disable anonymous and private communication. She may furthermore restrict the communication such that the backchannel establishes a student-to-lecturer communication (but not student-to-student). Such a configuration can be made in order to better meet instructional requirements. It may as well be the result of a joint decision of the lecturer with the audience.

The backchannel posts appear in the time-line in a concise fashion, only showing information related to the author, the post category, possibly a reference to another post and the post's content (see Figure 4.1). Further information about and possible actions to a post appear in the expanded view of the post, which can be obtained by selecting the respective entry in the time-line. Figure 4.4 shows all components of a post on *Backstage*. Besides the user name and text, a post includes awareness-promoting information on its author and on the post itself.

4.2.1.2 Maintaining Interactional Coherence by Threads

Although the communication time-line simplifies browsing, it does not reflect the inherent non-linear nature of chats (Herring, 1999), which possibly leads to readers getting
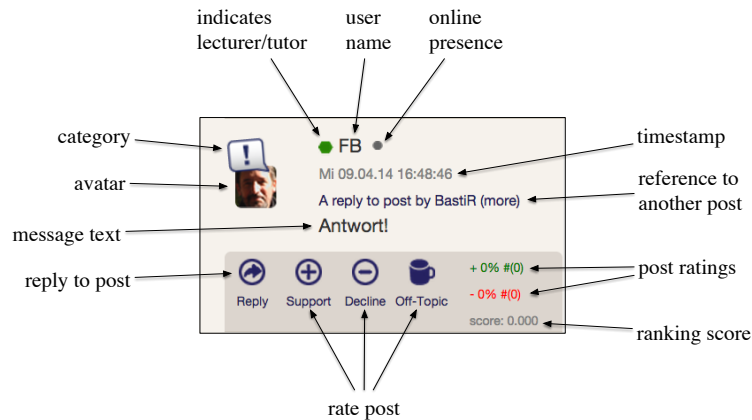
Figure 4.4: All elements of a backchannel post on *Backstage*.

lost. As a remedy, in addition to the chronological ordering of the time-line, one can navigate using the category icons placed on the slides: Clicking an icon on the current slide filters the time-line such that only the selected post together with its replies are shown (Figure 4.5). Thus, *Backstage* complements the time-line navigation with a topic-wise navigation in the time-line of the current slide. Explicit references among posts also maintain interactional coherence. They allow to carry along context with a post. For example, a reply to a post may display the post to which it responds (see Figure 4.6).
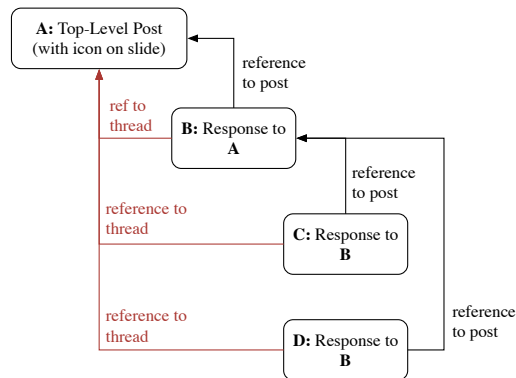
Figure 4.5: A chain of related backchannel posts. All posts **A**, **B**, **C**, and **D** share the same location and are all subsumed by **A** on the slide. Selecting **A** on the slide filters the time-line to only display **A** and all its related posts **B**, **C**, and **D**.

4.2.1.3 The Effort Associated with the Annotation Workflow

In user interface design, the term affordance describes the actions that can be performed on an interface element (cf. Peters, 2014). For example, buttons and links can be clicked. Affordances of an interface element are indicated by so-called signifiers: buttons and links (should) look like they can be clicked (cf. Peters, 2014). The concepts of affordance and signifiers can likewise be extended to functional groups of a user interface. The
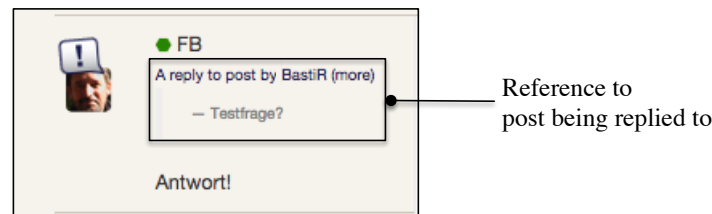
Figure 4.6: A reply to a previous post. A reference to the previous post provides context for the reply.

affordance of *Twitter* is writing short texts. This is signified by an emphasis of the text editor and the button for publishing a post (Figure 2.4). Such a design clearly conveys to a user the ease associated with publishing text.[3]

According to the messaging threshold theory, users consider the utility of a message against the effort required to create the message. If the utility outweighs the associated effort then users invest the effort to create a message (Reid et al., 1996). For example, a student's utility of getting an appointment with the lecturer most likely justifies the effort of logging in to a computer and writing an e-mail to the lecturer. The utility can likewise be of utmost trivial nature, e.g., enjoying small-talk. In this case, the effort of sending messages needs to be comparably low (Döring, 2003). Taking messaging threshold into account, *Twitter*'s user interface invites frequent posting of tweets. Having a similar structure as *Twitter*, the first prototype is likely to invite the same behavior, which however may not be desirable in lectures. A three-step process as described above increases the messaging threshold and may thus help to prevent posting of trivial and off-topic content.

Learning environments often make use of slowing down processes, e.g., by introducing obstacles or providing rigid structures, so as to trigger reflection or other kinds of learning-related mental processes (Peters, 2014). Thus, to evoke reflection and to make thoughtless posting more difficult, a three-step process of publishing a post on the backchannel is deliberately introduced. An increase of effort and a slowing down of processes is also often a product of guidance measures.

4.2.1.4 The Annotation Workflow as a Form of Guidance

According to contemporary research, supporting learners with direct instructions can positively affect learning (King, 2002; Kollar, Fischer, and Slotta, 2005). This is especially the case when learners have insufficient prior knowledge to get along with the topic at hand by themselves or when they are not able to use effective learning strategies (Kollar, Fischer, and Slotta, 2005).

In cognition psychology, the term *script* is used to describe daily life social knowledge. For example, the typical sequence of actions and interactions that occur when dining in a restaurant (e.g., entering the restaurant, being seated at a table, choosing from the menu and so on) can be considered as a script (Schank, Abelson, and Schank, 1977). As another example from day-to-day life, getting to know each other in a chat involves a typical sequence of conversation topics such as asking for the age, gender,

---

[3]Similarly, login pages of social media sites signify what kind of participation is promoted. If media thrive on participation by many, the login pages often emphasize the possibility to register by providing a simple registration form next to the login form (*Facebook* and *Twitter*), or by having a registration page as entry site rather than a login page (e.g., *Tumblr*).

occupation and hobbies (Döring, 2003). In a similar vein, learners may have scripts – in educational settings referred to as internal scripts – that represent plots relevant for learning. If students fail to employ efficient learning strategies it is surmised that no internal scripts are available for the task at hand or that existing internal scripts are insufficiently advanced or even flawed (Kollar, Fischer, and Slotta, 2005). As a remedy, it is suggested to provide learners with so-called external scripts, i.e., plots described on paper or within a computer-mediated learning environment, that instruct students to follow a given sequence of steps (a comprehensive overview is given by Weinberger, 2011). By repeatedly following the instructions provided by an external script, the learner is supposed to internalize the external script, i.e., create new or replace flawed mental structures by what is specified in the external script.

External scripts that focus on coordinating interactions among learners as well as between the instructor and learners can be distinguished from external scripts that are related to the creation of learning artifacts.[4] The former type of external scripts may be referred to as *interaction scripts*, since these external scripts address the interactional settings of instructions. The latter type of external scripts may be referred to as *production scripts*, since these external scripts focus on the activities of individual learners. For example, production scripts may help students to generate self-directed questions (King, 1992), or to construct arguments in a computer-supported collaborative learning (CSCL) environment (Weinberger, Fischer, and Stegmann, 2005). As a more specific example, a production script may help math students to properly construct induction proofs.

Both types of scripts may support learning in two ways: by *engendering* or by *facilitating* mental processes relevant for learning. In this sense, external scripts that are engendering aim at an intensified thinking, whereas external scripts that are facilitating offload tasks from learners so as to allow more cognitive resources to be invested in learning-related activities (Suthers, 2007). For example, a production script that instructs a student in properly constructing an argument may be considered engendering, since it demands the student to carefully consider the elements that constitute an argument, i.e., a claim, a warrant and a qualifier (Weinberger, 2011). As an example of an external script that is facilitating, an interaction script may specify a discussion setting and steer the progress of a discussion.

Although external scripts can be presented in various forms, e.g. on paper, CSCL research especially focuses on computerized forms of external scripts. For example, Weinberger, Fischer, and Stegmann (2005) embed a production script for argumentation in a computer-mediated learning environment. However, whether computerized scripts engender learning activities is likely to depend on how the scripts are realized, and whether learners also recognize these realizations as a support for learning. This may not necessarily be the case. For example, as Herring (1999) points out, rigid control of turn taking, which may be relevant in a collaborative learning setting and thus controlled by an external script, may lead to frustration. Better acceptance of such "shortcomings" due to external scripts may be achieved if the educational purpose of turn taking is made clear to students. That is, it may be reasonable to explicitly state to students which of the user interactions are considered part of a script and is intended to be supported by the script.

The three-step process of annotating slides on *Backstage* is set out to facilitate reflection prior to submitting a backchannel post. In particular its purpose is to prevent, thoughtless posting, especially off-topic posts. By requiring the selection of a suitable post category and a specific location on a slide, the backchannel conveys to students

---

[4]Often, the types of external scripts are distinguished between micro scripts and macro scripts (see Weinberger, 2011). However, for this thesis another distinction seems to be more appropriate.

which communication is desirable. Thus, the process of annotating slides on *Backstage* can be seen as a production script that is facilitating.

4.2.1.5   Rating as an Enabler of Social Control

Participants may also rate backchannel posts so as to express their approval or rejection of some backchannel posts. To this aim, a backchannel posts displays rating elements as well as up-to-date information on the ratings the backchannel post has so far received (Figure 4.4). By rating post, students give a feedback to the authors. Rating can foster a mutual and social regulation of the backchannel communication. Furthermore, rating serves to avoid redundant posts: Instead of posting the same question once again, the rating buttons of a post makes it easier and more natural for a student to support that post (In this way, rating and placing posts on specific slide locations used to avoid redundant communication can be considered as elements of an interaction script that is facilitating). This avoids a "noise" that would be harmful in lectures because repeated or similar posts would be a cause of distraction for students. Thus, the purpose of re-tweeting on *Twitter*, that is, expressing support of a post by repeating it, is achieved with the above described rating in a manner much more convenient to lectures.

Not only support, but also rejection, can be expressed on *Backstage* by ratings. *Backstage* offers two types of rejections of a post: disapproval, used for example for disagreeing with a post, and "off-topic", used for expressing that a post is considered inappropriate during a lecture. Off-topic ratings serve a social control of the backchannel contributing to both, strengthening the classroom community by giving it control and keeping the backchannel communication "on target". Approval can be understood both as "good" or "relevant to me"; the analogous applies to the two forms of rejection.

4.2.2   Promoting Awareness by Summarizing the Backchannel

The annotation of slides with backchannel posts generates a rich content and promotes awareness. The posts' locations on the slides as well as the post categories make students and lecturer alike aware of concerns in the audience at a glance. Awareness is an incentive to properly locate a post on a slide: improperly located, a post is unlikely to receive the attention wished for. Most of the components described in the following aim at improving a lecturer's awareness of her audience's concerns while avoiding an information overload.

Since posts are assigned categories, their category-based distribution can be determined. Participants can therefore be given a concise overview of the prevailing types of backchannel discourse in real-time. For instance, the lecturer can easily find out whether the students mainly post questions, allowing her to answer without delay. Figure 4.7 shows such an overview.

The overview window refers to the global backchannel discourse, independently from the displayed slide. Too frequent a reset, for example on every slide, might update the awareness-promoting information too quickly for the lecturer to take notice of it. For this reasons the lecturer decides when to update the overview.

Another form of summarizing is ranking of backchannel posts. The rating of posts by students assesses the posts' relevance for the lecture. It informs whether a post has been broadly acknowledged. As a consequence, posts' ratings make it possible to filter the time-line of the current slide for those posts considered most relevant. When the communication on the backchannel is high, a rating-based filtering of posts makes it possible for the lecturer to focus on the most relevant posts (Figure 4.8). In a sense, by rating students vote for those posts that are forwarded to the lecturer's user interface.
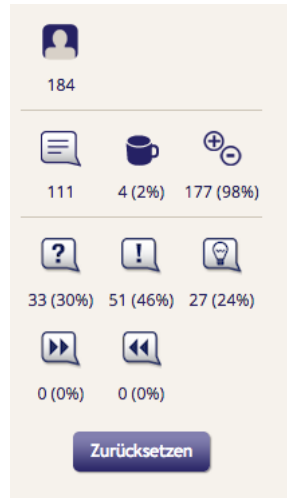
Figure 4.7: Overview window: The window shows the number of students on *Backstage* and the distribution of posts with their categories

That is, rating is a form of social filtering. Since the lecturer can only pay a limited attention to the backchannel, she might rely on this social filtering for selecting those students' backchannel post during a lecture (see Chapter 6). The lecturer can set the threshold of positive ratings a post must have received for being displayed on her user interface.
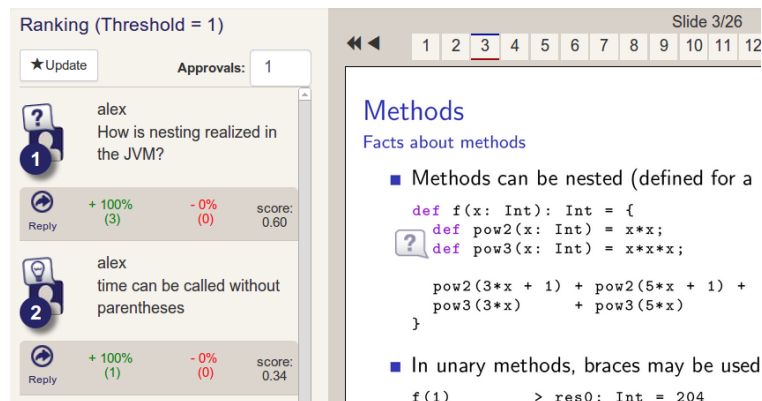
Figure 4.8: Message Ranking in *Backstage*

4.2.3 Online Presence on Backstage

Complementing social awareness, *Backstage* distinguishes between three kinds of online presence: "online", "offline", and "busy". The first two have obvious meanings. The third expresses a limited form of online presence. A student chooses "busy" if she is online but does not wish to participate in the backchannel communication.

On *Backstage*, presence determines the range of functions available at a student's user interface. In state "online", which is indicated by a green dot next to a student's user

name, a student's user interfaces is equipped with all functions. In state "busy", which is indicated by a red dot next to a student's user name, the user interfaces switches into the individual learning mode. That is, the backchannel provides only the backchannel posts of the respective participant and the lecturer. By default, a "busy" participant creates private notes. Partaking in quizzes is not affected by the choice of online presence. Since a post is the central interaction artifact of *Backstage*, the posts also indicates the author's the current online presence (cf. Figure 4.4). In order to reduce distraction and to keep the user interfaces calm, this kind of information is only displayed in the expanded view of the post (see Figure 4.4), which has to be explicitly requested by clicking the respective entry in the time-line.

Backstage supports two kinds of online presence of a lecturer: "online" and "offline". In order to reduce the social distance between students and the lecturer, *Backstage* strengthens a student's awareness of the lecturer's presence (Figure 4.9). To this aim, the lecturer is presented at the students' user interfaces with a photo, her user name, and a current onlie presence. The lecturer may in addition publish a personal statement, an anecdote or an informal definition of learning goals of the current lecture session.

Figure 4.9: A lecturer's online presence at a student's user interface. The lecturer is presented with a photo, his online presence (e.g., "online", as indicated by a green dot), and a personal statement on the lecture topic.

4.3 The Audience Response System (ARS)

The ARS on *Backstage* is built upon a generic architecture that allows the realization of different kinds of quizzes. *Backstage* currently supports multiple-choice quizzes, polls, and open-item quizzes in which students answer with free text. When a quiz is started, the user interfaces switch into an "ARS mode". Figure 4.10 and Figure 4.11 show a lecturer's and a student's views in ARS mode.

The quiz answers received during a running quiz are summarized and displayed at the lecturer's user interface in real-time. Apart from open-item quizzes, such a summary does not only include the temporal development of the total number of answers given by the audience but also the temporal development of each of the answer alternatives students can choose from. A display of these temporal developments can be a valuable source for the lecturer. For example, a rush of correct answers at the beginning of a quiz seems to be a sign that most learners have understood the issue. A rush of wrong responses at the beginning tends to indicate that an essential aspect of the lecture has been misunderstood by most.

Typically, the temporal development of responses takes an S-shaped curve: At the beginning only few responses are submitted. The rate of increase at the beginning mainly

depends on the complexity of the question and the response alternatives provided. Both, the question and the answer options have to be read and understood first. Then, after a short while, the curve shows a relatively steep rise, because many students submit their answers within this time span. At the end, most students have submitted their answers and the curve starts to flatten out. The lecturer can take this as an indicator that the quiz can be stopped, which usually is the case after a few minutes (Draper and Brown, 2004). Besides, the lecturer is provided with information about how many students are currently online and thus invited to participate in a quiz, and how many of them actually gave an answer.

After stopping the quiz, the lecturer can take a look at the results before making them public and decide whether to publish the summary of the quiz responses or not. If she decides to publish the results, a final quiz summary consisting of the quiz slide and a compiled summary slide is prepared and inserted into the set of lecture slides. In this way, the quizzes and their results are persisted and made available as a shared artifact students and lecturers may, like conventional lecture slides, annotate and reflect on. If for some reason the quiz yields useless results, e.g., because a quiz contains errors, the lecturer may decide to not publish the summary.
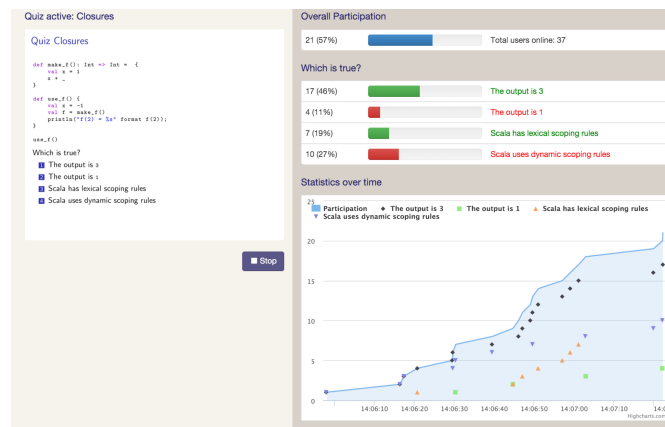
Figure 4.10: The lecturer's user interface while a quiz is running.

Figure 4.11: A student's user interface while a quiz is running.

4.4 System Architecture

This chapter presents the system architecture of *Backstage*. *Backstage* is realized as a Web application built upon conventional Web technologies and protocols. A schematic overview of *Backstage*'s architecture is depicted in Figure 4.12.
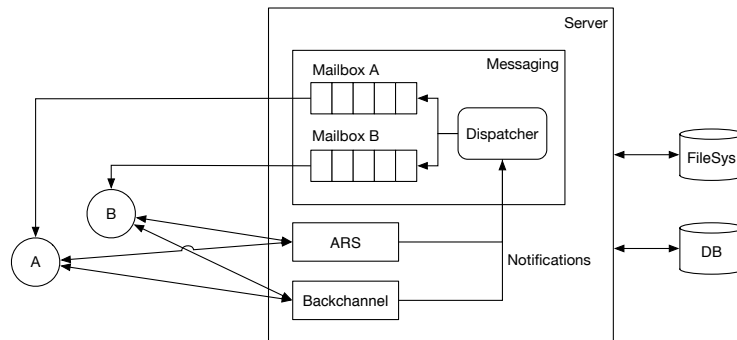
Figure 4.12: Schematic overview of *Backstage*. The circles *A* and *B* represent clients. Clients communicate directly with the Audience Response System and the backchannel (e.g., for retrieving a new post). The modules may notify clients on events using the messaging module (e.g., for notifying clients on a new post). Clients continuously check their mailboxes for notifications

The *Backstage* server consists of three components: An ARS component, a backchannel component and a messaging component. The ARS component realizes all quiz-related functions and the Backchannel component the microblog-based annotation of lecture slides. Clients request and submit data directly to these components. The Messaging component is used to realize synchronous communication among clients using conventional Web technologies.

A desirable feature of Classroom Communication Systems (CCS) is to notify users in real-time of events generated by other users. For example, publishing a public backchannel post in a lecture needs to result in a prompt update of all participants' time-lines. As another example, when the lecturer starts a quiz, the Audience Response Systems at the student user interfaces should promptly be activated. Unfortunately, the conventional Web does not directly support bidirectional real-time communication between clients and a server, which however is a prerequisite for synchronous inter-client communication in a client-server architecture..

4.4.1 Realizing Inter-Client Communication over HTTP

The Web has originally been designed as a distributed, client-server-based document retrieval system among networked computers. In such an architecture, the client is the active component that initiates data exchange with the server. The server is passive in that it only responds to a client's requests, it cannot initiate communication with a client by itself. Furthermore, Web clients can only communicate with each other through a server. The Web's architecture thus makes truly interactive applications and inter-client communication quite difficult to realize.

The key advancements of the Web towards an interactive Web are XHR[5] and Ajax.[6]

---

[5]XHR stands for XMLHttpRequest and is the technological foundation for the transmission of asynchronous HTTP requests. See Wikipedia, http://bit.ly/18KygZe (last visited on March 2nd, 2015).

[6]Ajax stands for Asynchronous Javascript and XML and is used to create asynchronous HTTP requests.

While in the conventional Web a request-response-cycle always ends with a complete reload of a Web page by the requesting client, XHR and Ajax allow submitting requests and receiving responses asynchronously in the background without an ensuing complete page reload. As a consequence, the server does not necessarily have to respond with a complete Web document but may just deliver generic data to the client. The client may then use the data to partially update the locally displayed Web page by manipulating its Document Object Model. This allows Web pages to provide a look-and-feel similar to Desktop applications, which also gave them the name "Webtop applications".

XHR and Ajax allow mimicking bidirectional communication over conventional HTTP (for an overview see McCarthy and Crane, 2008). The fundamental idea is simple: provide the server with frequent occasions to respond with recent events by having the client frequently request updates from the server. If these requests are issued in the background, from an observer's perspective it seems as if the updates are pushed from the server to the client. This is the conceptual basis for full-duplex inter-client communication over the Web. The techniques presented in the following subsections differ in their complexities resulting from balancing network traffic against promptness of updates and thus the quality of synchrony.

4.4.1.1 Piggybacking

In piggybacking, a server's response to a client's request[7] does not only refer to the request but also contains updates of the application's state. That is, updates of the application's state are always added as additional payload to, or piggybacked by, a server's response.
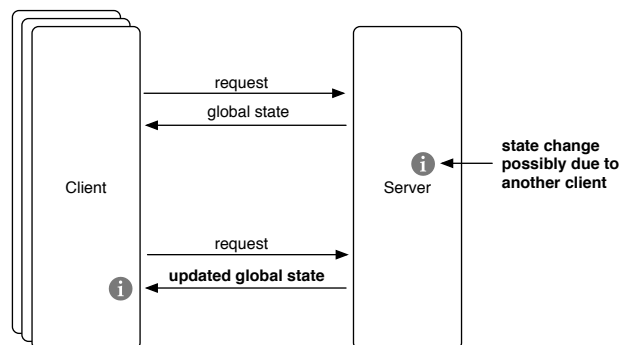
Figure 4.13: Piggybacking. Whenever the client issues a request to the server, additionally to the response to that request, state updates are returned as well.

The communication pattern between a client and the server using piggybacking is described in the Figure 4.13. A "synchrony" between clients largely depends on the clients' activity. The more requests a client issues, the more often is the client provided with updates, and the more do the interactions give an impression of real-time. The time lag between the occurrence of an event and a client's request may be quite large and difficult to predict, however. Thus, piggybacking alone may not be sufficient to

Although the name suggests a use of XML as a data exchange format, nowadays JSON is usually used. See Wikipedia, http://bit.ly/1zVHJ6t (last visited on March 2nd, 2015).

[7] If not stated otherwise, in the following subsections all requests are considered to be asynchronous using XHR and Ajax.

provide real-time bidirectional communication. Nevertheless, piggybacking can be used in combination with other techniques to reduce network traffic.

4.4.1.2 Polling

Possibly the most popular technique for providing a real-time bidirectional communication is polling. At fixed intervals, a client requests updates from the server. The server immediately responds to the request as follows. If a notifiable event has occurred the client is informed about the event. If no event has occurred an empty response is returned to the client. Thus, in polling a request is responded to immediately and the request-response cycle completes. Resembling the custom mode of HTTP, polling does not require any technological modifications on the server side.
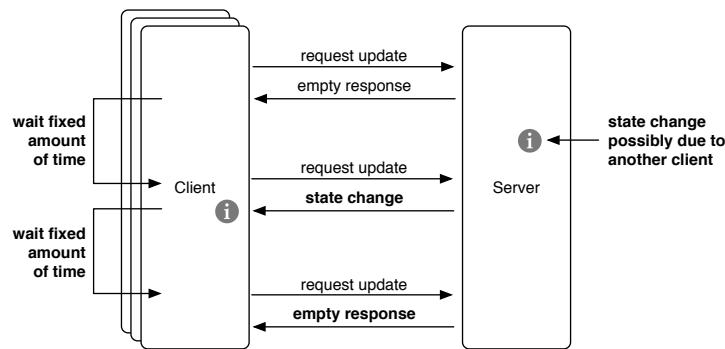
Figure 4.14: Polling. At fixed intervals the client requests the current state. The server immediately responds to the request either by an empty response if no state change occurred since the last request, or transmits the new state.

Polling is quite simple to realize which may be a reason why it is often used. A difficulty associated with polling is the determination of an appropriate interval of update requests. If too long, the time lag between occurrence of an event and notification to a client may be too large for sustaining an impression of real-time communication. If too short, however, too much network traffic may be produced and server resources are wasted in responding to useless requests. Thus, polling appears to be suitable in situations in which events occur rather regularly or in which events occur too frequently to prepare a response for each event. In these cases it is more reasonable to notify the client bulk-wise, with a collection of events that occurred after the last request.

4.4.1.3 Comet

Comet (also known as Long HTTP) has been conceived to avoid too frequent update requests as with polling. In Comet, a client's update request is not responded to immediately but kept open by the server until a notifiable event occurs. This allows the server to instantly respond to a client after an event has occurred. Immediately after receiving the response, the client re-establishes a new connection by issuing a new request for updates. This method ensures minimal delay of updates while avoiding superfluous network traffic.

Comet has been often criticized for requiring the server to manage a large number of open connections. This criticism can be addressed by timing out pending requests

so as to free occupied resources.[8] A similar problem arises on the client. The RFC on HTTP/1.1 stipulates that clients should not keep more than two parallel connections to the same host.[9] Due to the permanent use of one connection for Comet, browsers have to manage the retrieval of data with a single connection, which likely results in slower loading times. For this reason, clients may as well time out a Comet request. However it should be noted that recent browsers do not adhere to the recommendation given in the RFC. Most browsers keep more than two parallel connections to the same host, e.g. six parallel connections per host by recent Firefox and Chrome browsers, up to eight parallel connections by the Internet Explorer, and four parallel connections by Safari browsers (Smith, 2013).
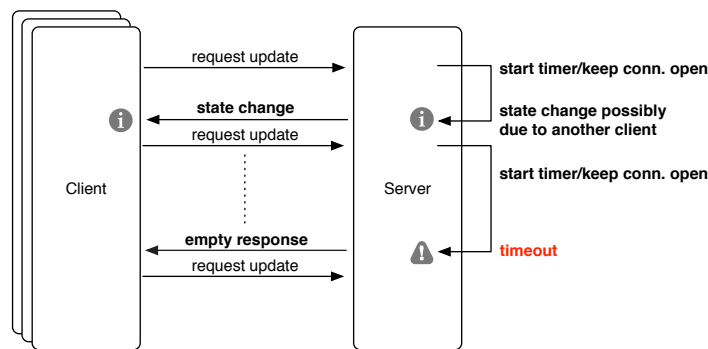
Figure 4.15: Comet/Long HTTP. The client initiates a request that is kept open until the server responds with a new state or times out a request. In both cases, as soon as the request terminates, the client issues a new request.

4.4.1.4 WebSockets

WebSockets are the most recent development of the Web to support genuine bidirectional real-time communication. The term WebSockets refers to both a network protocol extending HTTP[10] as well as to an API for browsers.[11] To establish Websocket communication a handshake is necessary. Therefore, an HTTP request sent to the server contains header fields that ask the server to upgrade the connection to a full-duplex communication channel.[12] The server can accept the request and complete establishing a WebSocket channel by initiating protocol switching. In contrast to the other techniques, however, the application server needs to be WebSocket-compliant.

---

[8]The resource demands of Comet on the server can be quite high. For example, in Java Servlets prior to version 3, requests are processed in separate threads, which can be freed only after the request is responded to. Thus, many or long-lasting requests result in the allocation of many threads. Threads are a limited resource and quite expensive to create. If the maximum number of available threads is reached no requests are accepted and the server stalls. However, as a remedy, the latest application servers decouple the processing of requests from threads by means of scheduling (this is also referred to as asynchronous request processing).

[9]*Hypertext Transfer Protocol – HTTP/1.1*, IETF RFC 2616, dated June 1999, `http://bit.ly/1DOCzhq` (last visited on March 2nd, 2015).

[10]*The WebSocket Protocol*, IETF RFC 6455, dated December 2011, `http://bit.ly/1xA9ZOf` (last visited on March 2nd, 2015)

[11]*The Web Sockets API*, W3C Working Draft, dated September 2011, `http://bit.ly/1DHQLFo` (last visited on March 2nd, 2015)

[12]See Wikipedia, `http://bit.ly/1wYX2Pw` (last visited on March 2nd, 2015).
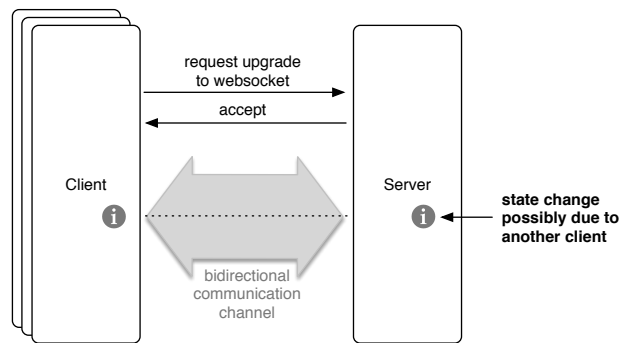
Figure 4.16: WebSockets. The client issues an HTTP request with an additional request to upgrade the connection to a truly bidirectional persistent connection between client and server.

4.4.1.5 Realizing Inter-Client Communication in Backstage

Today, WebSockets may be the preferred choice of technology to realize synchronous inter-client communication. At the time of development of *Backstage*, however, Web-Sockets were in an experimental state and suffered from several security issues. In the rapidly evolving ecosystem of Web application development, technologies quickly come and go, which makes a choice of technology quite difficult. Thus, sometimes it is better to stick to the technologies that have proven reliable in practice and to refrain from early adoption.

Apart from WebSockets, Comet provides the best compromise between synchrony and network traffic. It has thus been decided to use Comet for notifying clients about real-time events. Several elaborate libraries of Comet exist. The Bayeux Protocol, for example, is a comprehensive specification of communication between the client and the server as well as between clients.[13] Direct Web Remoting (DWR) is a JEE-based Comet library that abstracts from concrete communication patterns and realizes remote procedure calls between a client and a server.[14] The basic communication pattern can thereby be configured to either piggybacking, polling or Comet (which DWR refers to as Reverse Ajax). Due to its simplicity, DWR has been used in the first prototype presented in Chapter 3. Due to a change to the non-JEE-compliant Web framework Play![15] the use of DWR had to be renounced. DWR has been replaced with an own implementation of which the messaging component (see Figure 4.12, p. 53) is a part.

Thus, a *Backstage* client is provided with new notifications from the server using Comet. Therefore, new notifications are buffered in mailboxes managed by the server. To inform clients about real-time events, the Backchannel and ARS components create notifications of the kinds "new post with id #9" or "quiz with id #5 started". The messaging component dispatches these notifications to the clients' mailboxes. When a client receives these notifications it communicates directly with the ARS and the backchannel, e.g., to actually retrieve new posts or started quizzes, and to update their views, respectively. From a conceptual perspective, the communication pattern using Comet looks as if clients communicate directly with other clients in a full-duplex mode (e.g., the lecturer's client notifies students' clients about a started quiz).

---

[13]See the Bayeux Website, http://bit.ly/1tgBuWn (last visited on March 2nd, 2015).

[14]See the DWR Website, http://bit.ly/1aKv54w (last visited on March 2nd, 2015).

[15]http://playframework.com

4.4.2 Towards a Scalable and Extensible System Architecture

During the development of *Backstage* the used Web programming framework changed from the JEE-compatible framework Grails[16] to Play!. The NoSQL database MongoDB was used for data storage.[17] All these technologies are relatively easy to learn and to use, which was a decisive factor, since students should be involved in the development.

4.4.2.1 Improving Horizontal Scalability

A strong benefit of the chosen technology can be considered to be the support of a nothing-shared architecture, according to which no state is shared between the client and the server. HTTP sessions, a means to maintain shared state between a client and a server over several request-response cycles are precluded in Play!. Instead, a client transmits all relevant contextual information along with a request, enabling the server to re-create the context required to respond to a request. Avoiding shared state particularly simplifies scaling out, i.e., distributing a Web application over more than one server so as to manage high loads.

However, scaling out requires that a Web application's global state is managed in a database, realizing a nothing-shared architecture among server nodes. Otherwise, the server nodes need to keep the application state in sync, a quite complex task that needs to be accomplished by the Web application itself. Managing shared state also complicates adding and replacing server nodes. Externalizing the global state of a Web application to a database requires fast read and updates by the database system, which most NoSQL database systems provide (by renouncing joins as particularly expensive operations and aggressive caching policies). Also, NoSQL databases are particularly amenable to scaling out by sharding, i.e., fragmenting the data set over multiple database servers, however sacrificing transactional consistency (for a comprehensive overview, see Sadalage and Fowler, 2012).

4.4.2.2 Considering a Microservice-Based Architecture

CCSs such as *Backstage* can be quite elaborate systems difficult to maintain and to extend. Since *Backstage* is a research prototype, extensibility by further components might be of particular interest. Microservices is an architectural pattern that considers software as a collection of collaborating software components deployed independently.[18] Each microservice independently realizes a self-contained task. A request is processed by collaboration of microservices, which communicate over a light-weight communication interface such as REST+JSON.[19] Thus, microservices are an architectural realization of the Single-Responsibility-Principle known from object-oriented design, stating that a responsibility should not be distributed among several components.[20]

For *Backstage* a microservice architecture would be achieved by splitting the server into individually loosely coupled Web applications (Figure 4.17).

---

[16]http://grails.org

[17]http://mongodb.org

[18]James Lewis, Martin Fowler, *Microservices*, Blog post dated 25 March 2014, http://bit.ly/1dI7ZJQ (last visited on March 2nd, 2015).

[19]REST stands for Representational State Transfer. Put simply, REST considers the Web as a distributed (tree-like) data structure in which the nodes are referenced by URIs. The data structure can be operated on using the HTTP request methods. The HTTP methods are given a clear semantics, e.g. GET-requests may be used to retrieve data, POST-requests to create new data. For an overview, see Richardson et al. (2007).

[20]See the Wikipedia article on the Single-Responsibility-Principle, http://bit.ly/RwUy7F (last visited on March 2nd, 2015).
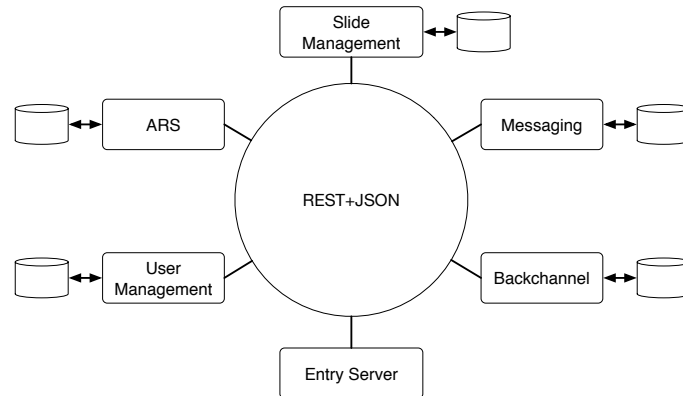
Figure 4.17: The suggested microservice architecture of *Backstage*. The components of *Backstage* are realized as individual Web application loosely coupled by communication based on REST and JSON.

Such a design would be beneficial for various reasons. First, Web frameworks that are most suitable for the task to be accomplished can be used. The same applies to the hardware used. For example, slide management may be deployed on a server optimized for file-based I/O. Databases could be chosen accordingly. For example, user management may be based on graph-based NoSQL databases that are particularly suitable for managing relationships among users ("friendships"), while the slide management server may use a document repository. Second, load is balanced over multiple server nodes and, thus, scaling measures can be restricted to those microservices facing heavy load instead of having to scale out the entire Web application. Third, data access can be controlled. For example, a user management server can provide different views on the user data for different microservices. Finally, new microservices can be developed in isolation and integrated into the architecture without increasing the global complexity of the entire system.

However, this approach also poses several challenges. In particular, complexity is added to each subsystem, since each subsystem needs to be equipped with communication capabilities, including e.g., marshalling and unmarshalling of data. Also, system administration becomes more difficult. Network issues can become a serious problem if the subsystems are deployed on distributed nodes. Also, data privacy concerns may be an issue if microservices are deployed to non-authoritative sites. The architecture may contribute to slower response time, since processing a client's request is likely to entail communication among several microsystems. Finally, care has to be taken to maintain concise REST+JSON communication interfaces that can become quite complex.

EVALUATION OF BACKSTAGE IN PRACTICAL USE

> The surveys used in the analysis reported about in this chapter have been developed in collaboration with Vera Gehlen-Baum.

This chapter reports on the use of *Backstage* in three computer science courses held at Ludwig-Maximilian University of Munich and one computer science course held at Saarland University. The evaluation analyzes the data collected on *Backstage* as well as by surveys administered to the course attendees at the end of each course. The selection of data used for this evaluation is guided by the following questions:

Q1 Is *Backstage* an appropriate means to foster interactivity in lectures?
Q2 Is *Backstage* a tool for reworking lectures?
Q3 What are the students' general impressions about *Backstage*?
Q4 Does *Backstage* help in promoting and directing awareness?
Q5 What kind of communication takes place on the backchannel?
Q6 Are the activities on the backchannel and in the quizzes somehow related?

Question **Q1** is concerned with the actual objective of using *Backstage*, viz. promoting an active involvement of students in (large-class) lectures. Question **Q2** investigates whether *Backstage* supports rework of lectures, e.g. exam preparation. Question **Q3** considers the students' general impressions and perceptions of using *Backstage* in lectures. Question **Q4** considers the role of *Backstage* as a means to promote classroom awareness and to direct the lecturer's awareness to what students find relevant on the backchannel. Question **Q5** considers the quality of activity on the backchannel. A strong focus on lecture-relevant communication on the backchannel would certainly justify usinf a backchannel in lectures. Finally, Question **Q6** investigates potential relationships between a student's backchannel and quiz activities, e.g., do students who often engage in off-topic communication perform worse in the quizzes than other students?

5.1 Overview of the Courses

This section gives an overview of the four courses in which *Backstage* has been used (see Table 5.1). The courses differed in terms of the target audience (first versus advanced

semesters, audience size), and the degree of adaptations of the lectures to a use of *Backstage*.

ID	Univ.	**Course** (titles translated)	**Short Description**
P	LMU	Introduction to Programming and Modeling (summer 2013)	Course on functional programming
W	LMU	Web Information Systems (winter 2013/14)	Advanced course on languages of the (Semantic) Web, search engines and network analysis
L	LMU	Logics and Discrete Structures (summer 2014)	Course on mathematical logic and discrete mathematics
P1	SaarU	Programming 1 (winter 2014/15)	Course on functional programming

Table 5.1: The courses for which the use of *Backstage* is evaluated (LMU = Ludwig-Maximilian University Munich, SaarU = Saarland University).

5.1.1   Courses at Ludwig-Maximilian University Munich (LMU)

In the past two years, *Backstage* has been used in three computer science courses at LMU (see Table 5.1). The courses **P** (held in the summer term 2013) and **L** (held in the summer term 2014) were introductory courses for first-year Bachelor students. Course **W** (held in the winter term 2013/14) addressed students at the end of their Bachelor studies. All three courses were given by the same lecturer.

Apart from the development of quizzes, Course **P** received no other special adaptations to *Backstage*. An overview of the lectures of Course **P** can be found in Appendix A.8. The lectures were on a weekly basis and took three hours with a five to ten minutes break in the middle. The lectures were mainly held by presentation of slides with occasional use of a whiteboard or an overhead projector for demonstration purposes. At the beginning of the term the students received a proper introduction to *Backstage*. During and outside[1] the lectures, an experienced tutor was frequently present in the backchannel to clarify statements made in the lecture and to answer student questions.

In the subsequent winter term 2013/2014, *Backstage* was used in Course **W** (the course outline can be found in Appendix A.9). As a major adaption of Course **W**, lecture slides have been created particularly for the use with *Backstage*:[2] The slides were designed to be clear and concise and provide space for student comments. Furthermore, quizzes were developed for the course. The lectures were on a weekly basis and took three hours with a five to ten minutes break in the middle. The lectures mainly based on presenting the slides with occasional use of whiteboard and an overhead projector for demonstration purposes. An experienced tutor was occasionally present in the backchannel to clarify statements made in the lecture and to answer student questions. Additionally to *Backstage* two other research prototypes, *WIS Market* and *Termina*, were used outside the lectures to promote active involvement with the course topics.

The lectures of Course **L**, which was held in the summer term 2014, underwent complete revision. The material was prepared to be less rigorous and more intuitive than in previous semesters. The revision aimed at making the first encounter with mathematical logics, a branch of mathematics not taught in school and thus mostly unknown to first-year students, easier. During the revision the use of *Backstage* in the

---

[1]In this chapter "outside" a lecture refers to "before or after" a lecture.

[2]In prior semesters, the course was based on lecture notes.

Figure 5.1: Use of *Backstage* in Course **P** ("Introduction to Programming and Modeling" in the summer term 2013 at LMU).

lectures was taken into account. The material was reorganized such that each lecture presented and completed a topic (the course outline can be found in Appendix A.10). The lecture slides contained distinct slides at the beginning that invited students to post initial comments or questions that should be dealt with before presenting the topic of the lecture. Much effort has also been invested in the development of the quizzes. Quizzes were run each 20 to 30 minutes of the lectures. Two lectures were used for mock exams. Students had 60 minutes for solving the exam questions. Sample solutions were presented afterwards. Using the Audience Response System of *Backstage* students were asked to grade their own exam solutions anonymously using a comprehensible grading scheme similar to what was used in the final examinations. The results of the mock exams were reported back to the audience and discussed in the lecture. The lectures were on a weekly basis and lasted three hours with a five to ten minutes break in the middle. The lectures were mainly based on presenting the slides with occasional use of a whiteboard for demonstration purposes. The lecturer was present at fixed times outside the lectures to clarify statements or answer student questions on the backchannel.

At the end of the four courses, the students were invited to take part in an online survey. The survey asked about the students' use of *Backstage* and the perceived usefulness of *Backstage* both in terms of its potential to raise activity in the lecture as well as to support better learning outside the lecture.

Although *Backstage* provided the same basic services across the three courses, different versions of *Backstage* were deployed. For example, in Course **P** and **W**, students were

provided with a much more elaborate navigation of the slides that indicated the number of posts on a slide not yet read by the student. Due to necessary architectural changes and lack of time, this function was not available in the later Course **L**. In contrast, students in the courses **L** and **P1** were provided with the same statistics overview about the activity on *Backstage* as was shown to the lecturer in order to enhance their awareness of the activities on *Backstage*. In Course **L** negative rating was not available.

5.1.2 Course at Saarland University

In the winter term 2014/2015 *Backstage* has been used in the computer science introductory course "Programming 1" at Saarland University (see Appendix A.11). This course is referred to as **P1**. Course **P1** covered the same topics as Course **P** at LMU.[3] **P1** was given twice a week and each lecture took two hours. The students were not specially instructed in how to use *Backstage* but were provided with a help page containing a series of short screencasts explaining how to use *Backstage*. The student questions posted on the backchannel were answered during the lecture both in the frontchannel[4] by the lecturer as well as on the backchannel by a teaching assistant.[5] Besides the creation of quizzes, the course was not specially adapted to a use of *Backstage*. As a significant difference to the courses at LMU, *Backstage* was only used during lectures; a content management system and a forum were used for exchange between students and the teaching team outside the lectures. In particular, the lecturer disabled the commenting functionality after the lectures such that no comments could be posted after the lectures. The lectures were held by presenting lecture slides. Two additional beamers were used for displaying slides of previous lectures for repetition as well as in-class programming demos. Additional examples, proofs and explanations were provided on whiteboard. Most of the time, the ranking of backchannel posts was enabled.[6] At the end of the course, students were asked to participate in an online survey that was a shortened version of the surveys administered in the courses **P** and **L**.

5.2 Preparation of the Data Collected on Backstage

The questions formulated at the beginning of this chapter are investigated on the basis of the following data collected on *Backstage*: (1) The student logins to *Backstage*, (2) the backchannel communication, (3) the ratings of backchannel comments, (4) the student responses to the quizzes, including the correctness of the responses given in percentages, and (5) the student responses to the surveys administered at the end of the courses.

Although the login data is quite noisy,[7] it provides insights into how many participants were present during and outside the lectures. Therefore, a participant is counted as logged in if she has been marked as online during and outside a lecture at least once. In this way an upper bound estimate of the number of online participants during a lecture is obtained.

---

[3]Since more time was available in **P1** than in **P**, some topics could be taught in more detail.

[4]The term frontchannel refers to the face-to-face exchange in the lecture hall.

[5]The teaching assistant registered on *Backstage* as a normal student, not as a member of the teaching team. As a consequence, the teaching assistant's backchannel posts were not specifically marked (see Figure 4.4 on page 46). This may have affected the visibility of the teaching team on the backchannel.

[6]Information on the use of *Backstage* in Course **P1** has been collected through personal correspondence with lecturer Prof. Dr. Finkbeiner via email (April 2nd, 2015).

[7]Due to timeouts, a participant may be recorded as offline although she actually is logged in to *Backstage*. If a participant leaves the lecture but does not properly log out from *Backstage*, she may remain recorded as logged in after the lecture.

The evaluation of the backchannel communication required more sophisticated preparation and thus is described in a separate section (see Section 5.2.1). The same applies to the preparation of the survey responses, which is explained in Section 5.2.2.

In the analysis several notions of activity or engagement are referred to. Thus, the engagement in off-topic communication is measured by the number of comments being coded as *independent* from the lecture's topic (for the coding scheme, see Section 5.2.1). Furthermore, a student's quiz activity is measured as the product of the proportion of quizzes in which the student participated and the correctness rate the student achieved in each quiz. In this way, a student's engagement is not mistakenly valued as "high performant" when the student performed very well but actually participated only in a small number of quizzes.

5.2.1 Preparation and Analysis of the Backchannel Communication

Besides an investigation of how often participants used the provided post categories such as *Question* and *Answer* (see Section 4.1.1), an analysis of the content posted on the backchannel is required in order to better understand how the backchannel was used, e.g., for off-topic or lecture-relevant communication. For this kind of evaluation it is usual to classify the collected data according to certain characteristics, which in this case is the content of the backchannel posts. For the classification, a suitable coding scheme has to be devised according to which the backchannel posts can be classified, or coded. The coding scheme, however, has to satisfy a few properties (Bortz and Döring, 2009):

Exactness: The classes of the coding scheme have to be specified precisely by indicators whose presence or absence are decisive for an observation to be assigned to a given class or not.

Exclusiveness: An observation must not be assigned to more than one class. This is to avoid that two classes are equivalent, i.e. each class has to capture one or more characteristics distinct from the other classes.

Exhaustiveness: An observation must be assignable to at least one class. That is, the coding scheme must be able to "describe" all observations under study, no observations may remain unclassifiable.

5.2.1.1 Coding Scheme

The coding scheme used for the analysis of the backchannel posts is adapted from the classification suggested by Cogdill and Kilborn (2001) (also see Section 2.2.3). The class of *tangential* posts is removed from, and the class of *organizational* posts is added to, the classification for the following reasons.

Cogdill and Kilborn (2001) refer to *tangential* posts as *content-oriented* posts that are continuations of what has been started in the frontchannel, i.e. in the face-to-face exchange between the lecturer and the students. In the post-hoc review of the exchanges without thorough knowledge about what has been said in the lectures' frontchannels, however, it is impossible to discern *tangential* from *content-oriented* posts. That is, considering *tangential* posts in the coding scheme would violate the exactness and exclusiveness properties. Considering the class of *organizational* posts is sensible, since it can be expected that organizational issues are also communicated on an educational. The adapted coding scheme used in the analysis is thus as follows:

Content-Oriented Posts are posts that refer to contents of the lecture, given on the slides, drawn on whiteboard, or exchanged face-to-face in the lecture hall. Examples of content-oriented posts are questions like "Where does that x come

65

from?", confirmation or corrections like "yes, that is right", or posts used to mark content relevant for the exams, which the lecturer has advised students to do. Also, questions regarding the quizzes run in lectures, e.g. "Which is the correct answer?", are considered content-oriented.

Organisational Posts are posts that refer to organizational issues related to a course. They refer to exercises ("When is the next exercise sheet going to be available?"), tutorials ("The room of the tutorials on Mondays is too small"), requirements for passing the exam ("Do we have to be able to know all this for the exam?"), and administrative use of *Backstage* ("Are the exercise sheets going to be available on *Backstage* as well?").

Process-Oriented Posts are posts that refer to the lecturer's presentation. This class of posts includes posts on what the lecturer has missed to say ("What was the answer to the question?"), wishes for more or fewer clarification ("I think this should be clear by now"), difficulties in acoustics ("The microphone is noisy"), and pace of lecturing, also indicated by the comment categories *Too Slow* and *Too Fast* (see Section 4.1.1).

Participation-Enabling Posts are posts that refer to the use of *Backstage* (and possibly other software in lectures). For example, these posts include advice in annotating slides on *Backstage* ("Do not cover important formulas with your post icons!"), help in creating posts on *Backstage* ("How can I type a negation symbol?"), and advice in the proper use of comment modifiers (e.g. "If you want your comment to be a private note, add :&").

Independent Posts are posts not fitting into one of the classes above. This class comprises test posts participants create to get acquainted with *Backstage* ("This is a test"), off-topic posts, and closures of exchanges (e.g. "Thank you").

Reliability refers to the preciseness of a measurement instrument (Bortz and Döring, 2009). A coding scheme is reliable if coding the same observations several times yields sufficiently similar results. This is important because exactness is typically a fuzzy notion. There is often a gray area of observations where a certain indicator may be considered to be absent or present, depending on the coder's considers or even on that coders' mood. Essentially, reliability indicates the degree of this inevitable fuzziness. To measure the reliability of a coding scheme, a set of observations is classified, or coded, either by one coder multiple times (with some time in between) or individually by multiple coders. In case of nominal classes and two codings of observations the κ coefficient can be used to measure reliability (Bortz and Döring, 2009). A good reliability yields a κ coefficient between 0.6 and 0.7. Values above 0.7 indicate very good to excellent agreement (a value of 1 corresponding to two identical codings), and values below 0.6 indicate poor agreement, indicating great dissent between the coders due to flaws in the coding scheme such as overlapping classes (Bortz and Döring, 2009).

In the evaluation of the backchannel communication, two independent coders classified the backchannel communication in the four courses using the coding scheme described above. The κ coefficients indicate very good to excellent agreement between the two coders (see Table 5.2).

Course	κ	Number of Comments (%)
P	0.85	351 (approx. 50%)
W	0.93	282 (100 %)
L	0.89	487 (100 %)
P1	0.96	1744 (100 %)

Table 5.2: The κ Coefficients of the Backchannel Communication in the Courses **P**, **L** and **P1**. The third column reports the number of comments used for determining the κ coefficients.

5.2.2 Preparation of the Survey Responses

The surveys used in the courses **P**, **L**, and **P1** comprised Likert-type items as well as open items. In the open items, the respondents were asked to respond in free-form text. The open items were optional. The Likert-type items asked the respondents to provide a grade to a statement such as "*Backstage* was easy to use". A six-point grading scale was used, where one point corresponded to "strongly disagree" and six points corresponded to "strongly agree". By using a six-point scale for grading the respondents had to express some tendency and could not remain neutral.

Groups of Likert-type items were used to measure four constructs: INTERACTIVITY, RATING, AWARENESS and REWORK (see Table 5.3). Each construct is valued, or scored, by aggregating the grades of the items used to measure the construct. Each of the four constructs was measured by three to five survey items.

Similar to the reliability of a coding scheme, it is common to assess the reliability, or consistency, of a group of Likert-type items. Consistency thereby expresses the degree of interrelatedness of items used to measure a given construct. For example, consider the three Likert-type items "I like bike riding", "I enjoyed bike riding in the past" and "I hate bike riding" when used for measuring one's attitude towards bike riding.[8] A positive attitude is likely reflected by agreement on the first two items and disagreement on the third question. Likewise, a negative attitude is likely reflected by disagreement on the first two items and agreement on the third. It would be questionable whether the three items measure attitude towards bike riding when many responses fundamentally differ from these patterns.

Whether a set of Likert-type items are consistent can be measured by the α coefficient (Bortz and Döring, 2009). As a rule of thumb, a good consistency of items results in an α value between 0.8 and 0.9, whereas excellent reliability is indicated by an α greater or equal to 0.9.[9] Acceptably good α coefficients (see Table 5.3) were obtained in this analysis.

5.2.3 Statistical Analysis of the Data

Most of the observations reported on in this chapter are not normally distributed. Thus, the median (abbrev. Mdn) is used as the central tendency and the median absolute deviation (abbrev. MAD) is used as a measure of spread. Recall that the median of a frequency distribution is the sampled value of an (artificial) instance that bisects the distribution. The median is an interesting representative of the central tendency, since it is quite robust against outliers but likewise sensitive enough to reflect relevant

---

[8]Example taken from Wikipedia, `http://bit.ly/1yoUxqW` (last visited on April 30th, 2015).

[9]Wikipedia, `http://bit.ly/1yoUxqW` (last visited on April 30th, 2015).

Construct	α	Description
INTERACTIVITY	0.84	measures the usefulness of *Backstage* as a means to promote interactivity in lectures
RATING	0.82	measures the students' assessments of rating to mark relevant backchannel comments
AWARENESS	0.84	measures the usefulness of *Backstage* as a means to gather learning-related awareness
REWORK	0.95	measures the usefulness of *Backstage* for reworking lectures

Table 5.3: Overview of the Constructs Measured by the Surveys Administered at the end of the Courses **P**, **L** and **P1**. The survey used in Course **W** did not measure these constructs.

changes in the data (Garcin, Faltings, and Jurca, 2009). Note that in cases in which the data is normally distributed the median is equal to the mean[10] and the median absolute deviation differs from the standard deviation by a constant factor.[11] Non-parametric tests are used in the statistical analysis. For correlation tests Spearman's rho and for significance tests Wilcoxon-Mann-Whitney and Kruskal-Wallis tests are used. The statistical analysis is carried out using the statistics package *R* (version 3.02 for OS X).[12]

5.3 Results

5.3.1 Analysis of the Login Data

In Course **P** a total of 385 students registered on *Backstage*. The number also includes students who merely wanted to have access to the lecture slides but did not intend to regularly attend the lectures and use *Backstage*. on average, about 98 students logged in to *Backstage* during lectures (Mdn = 97.5, MAD = 39.29, Min = 14 in the last lecture, Max = 187 in the second lecture; see Figure 5.2a). Outside lectures, on average, about 18 students logged in to *Backstage* for rework (Mdn = 17.5, MAD = 13.34, Min = 0, Max = 28 for the third lecture). It is notable that actually from the sixth lecture on, no logins of students outside the lectures were recorded. The login data for Course **P** clearly shows a decreasing trend.

In Course **W** a total of 206 students registered on *Backstage*. As in Course **P** the number also includes students not intending to regularly attend the lecture, e.g., repeaters of the course. During the lectures about 34 students were logged in to *Backstage* (Mdn = 33.5, MAD = 9.64, Min = 0 in the fifth lecture, Max = 97 in the first lecture; see Figure 5.2b). Until the seventh lecture the number of student logins decreased and then remained rather constant. It is notable that besides four students who logged in to *Backstage* to revise the first lecture no further student logins outside the lectures were recorded (see Figure 5.2b).

In Course **L** a total of 269 students registered on *Backstage*. The number also includes students who merely wanted to have access to the lecture slides but did not intend to regularly attend the lectures and use *Backstage*. During the lectures, on average 65 students logged in to *Backstage* (Mdn = 65, MAD = 16.31, Min = 34 in the last lecture, Max = 91 in the second lecture; see Figure 5.2c). The decreasing trend of student logins

---

[10]See Wikipedia, http://bit.ly/WRHAnd (last visited on April 30th, 2015).
[11]See Wikipedia, http://bit.ly/1EtPuos (last visited on April 30th, 2015).
[12]www.r-project.org/
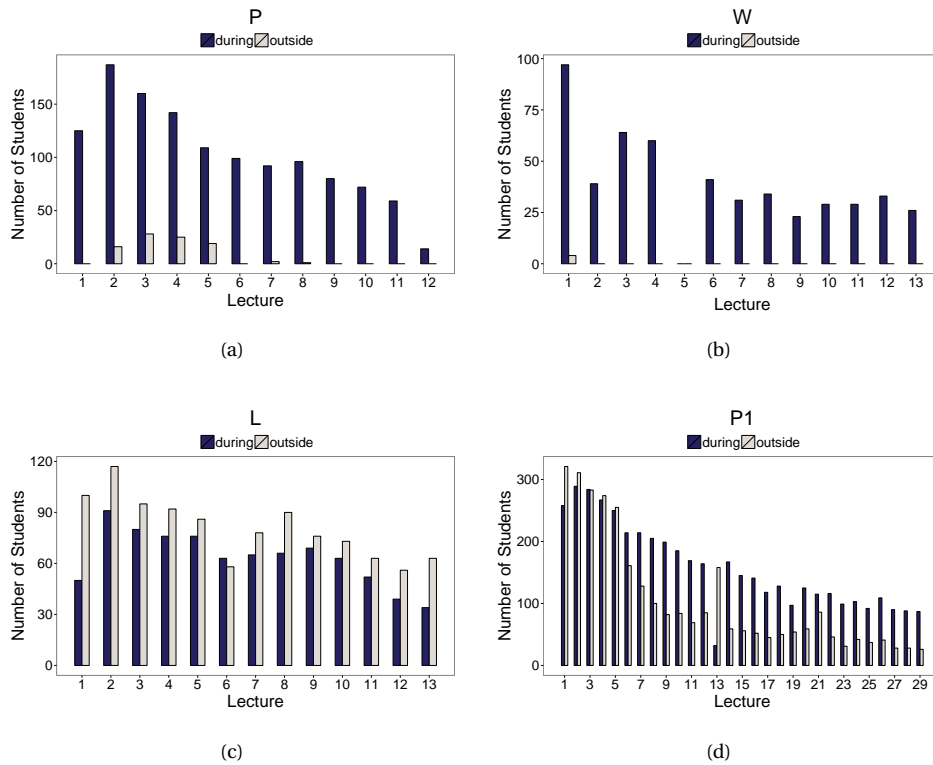
Figure 5.2: Number of Student Logins during (dark-colored bars) and outside Lectures (light-colored bars). The fifth lecture of Course **W** suffered from a breakdown of the WiFi connection in the lecture hall.

during lectures is only weakly pronounced. Compared to the other courses, a great number of students also logged in to *Backstage* outside the lectures, on average 78 students per lecture (Mdn = 78, MAD = 22.24, Min = 56, Max = 117).

In Course **P1** a total of 650 students registered on *Backstage*. As in the courses **P**, **W** and **L** this number is considerably higher than the number of students who regularly logged in to *Backstage* during the lectures. During the 29 lectures, about 141 students logged in to *Backstage* (Mdn = 141, MAD = 65.23, Min = 88 in the last lecture, Max = 290 in the second lecture; see Figure 5.2d). During semester the number of student logins shows a decreasing trend. The strikingly high number of logins outside the 13th lecture is notable. Outside lectures, on average about 59 students logged in to *Backstage* (Mdn = 59, MAD = 39.29, Min = 27, Max = 321; see Figure 5.2d).

5.3.2 Analysis of the Backchannel

Most of the backchannel communication in all four courses took place during lectures. Table 5.4 gives an overview.

5.3.2.1 Use of the Post Categories

The frequencies of the use of the five post categories provided on the backchannel varied across the courses (see Table 5.5). However, in all four courses, the backchannel

		P	W	L	P1
Total Posts		680	282	487	1744
During a Lecture	Mdn (MAD)	14 (20.76)	3 (4.45)	10 (14.83)	47 (40.03)
	(Min, Max)	(0, 116)	(0, 54)	(0, 113)	(0, 147)
Outside a Lecture	Mdn (MAD)	7.5 (11.12)	4 (5.93)	1 (1.48)	0 (0)
	(Min, Max)	(0, 122)	(0, 38)	(0, 17)	(0, 80)
Anonymous Posts	Mdn (MAD)	12 (10.38)	4.5 (2.22)	13 (10.38)	20 (14.83)
	(Min, Max)	(3, 39)	(2, 20)	(1, 28)	(8, 59)
Note Taking	Mdn (MAD)	5.5 (2.97)	1 (0)	1 (0)	1 (0)
	(Min, Max)	(1, 10)	(1, 4)	(1, 9)	(1, 1)

Table 5.4: Overview of the Backchannel Communication per Lecture (during and outside; Mdn = median number of comments per lecture, MAD = median absolute deviation, Min = minimum number of comments, Max = maximal number of comments).

communication is mainly based on the content-related categories *Question*, *Answer*, and *Remark*, whereas the feedback-related categories *Too Fast* and *Too Slow* are virtually not used at all.

		P	W	L	P1
Question	Mdn (MAD)	8 (4.45)	3 (1.48)	6 (7.41)	15 (10.38)
	(Min, Max)	(0, 39)	(1, 13)	(1, 19)	(2, 58)
Answer	Mdn (MAD)	15 (14.83)	13 (15,57)	11 (16.31)	35 (34.10)
	(Min, Max)	(0, 48)	(0, 29)	(0, 78)	(4, 133)
Remark	Mdn (MAD)	22 (11.86)	3 (1.48)	6 (7.41)	1 (1.48)
	(Min, Max)	(7, 49)	(0, 25)	(0, 27)	(0, 11)
Too Fast	Mdn (MAD)	0 (0)	0 (0)	–	0 (0)
	(Min, Max)	(0, 2)	(0, 2)	–	(0, 4)
Too Slow	Mdn (MAD)	0 (0)	0 (0)	0 (0)	0 (0)
	(Min, Max)	(0, 1)	(0, 3)	(0, 1)	(0, 1)

Table 5.5: Distribution of Backchannel Comments along the Post Categories (during and outside; Mdn = median number of comments per lecture, MAD = median absolute deviation, Min = minimum number of comments, Max = maximal number of comments).

In the lectures of Course **P** the use of the post categories differed significantly (Kruskal-Wallis, $N_Q = 145$, $N_A = 232$, $N_R = 299$, $N_F = 3$, $N_S = 1$, df = 4, H = 45.94, $p < 0.001$, see Figure 5.3a). Mostly, content-related post categories were used (post-hoc Kruskal-Wallis pairwise comparisons, $N_Q = 145$, $N_A = 232$, $N_R = 299$, $N_F = 3$, $N_S = 1$, all $p < 0.001$). The category *Remark* was used more often than *Question* (post-hoc Kruskal-Wallis pairwise comparisons, $N_Q = 145$, $N_R = 299$, $p < 0.014$). No other differences were found (post-hoc Kruskal-Wallis pairwise comparisons, $N_Q = 145$, $N_A = 232$, $N_R = 299$, all $p > 0.174$).

Also in the lectures of Course **W** the use of the post categories differed significantly (Kruskal-Wallis, $N_Q = 49$, $N_A = 154$, $N_R = 68$, $N_F = 3$, $N_S = 8$, df = 4, H = 34.47, $p < 0.001$). Among the content-related categories no differences were found (Kruskal-Wallis pairwise comparisons, $N_Q = 49$, $N_A = 154$, $N_R = 68$, all $p > 0.081$; see Figure 5.3b).

In the lectures of Course **L**, too, significant differences were found in the use of the post categories (Kruskal-Wallis, $N_Q = 117$, $N_A = 260$, $N_R = 108$, $N_S = 2$, df = 3, H = 23.36, $p < 0.001$). Among the content-related categories no significant difference in the frequencies of use was found (Kruskal-Wallis pairwise comparisons, $N_Q = 117$, $N_A = 260$, $N_R = 108$, all $p > 0.217$; see Figure 5.3c). The feedback-related category *Too Slow* was not used at all.

Finally, in the lectures of Course **P1** as well the use of the post categories differed significantly (Kruskal-Wallis, $N_A = 1136$, $N_R = 73$, $N_F = 6$, $N_S = 3$, df = 4, H = 120.22, $p < 0.001$, see Figure 5.3d). The content-related categories were used more often than the feedback-related categories (post-hoc Kruskal-Wallis pairwise comparisons, $N_Q = 526$, $N_A = 1136$, $N_R = 73$, $N_F = 6$, $N_S = 3$, all $p < 0.001$). Among the content-related categories, the category *Answer* was used the most (post-hoc Kruskal-Wallis pairwise comparisons, $N_Q = 526$, $N_A = 1136$, $N_R = 73$, all $p < 0.023$), followed by *Question* and *Remark* (post-hoc Kruskal-Wallis pairwise comparisons, $N_Q = 526$, $N_A = 1136$, $N_R = 73$, all $p < 0.001$).
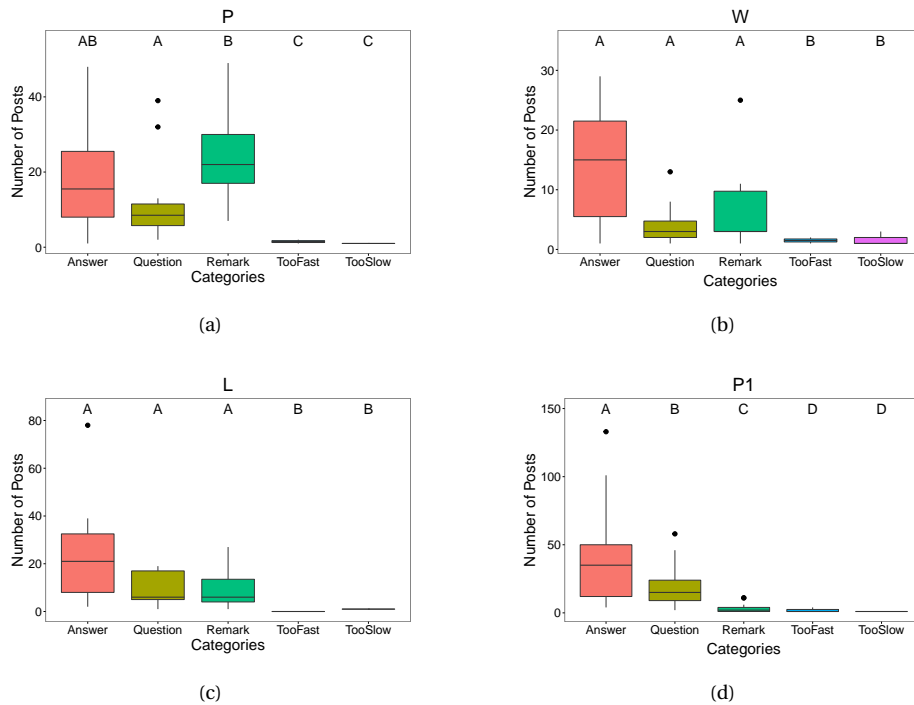
Figure 5.3: Frequencies of Post Categories (box-and-whisker plots with first and third quartiles, median and whiskers (minimum and maximum values); values outside 1.5 times the interquartile range (IQR) are displayed as outliers). Different letters at the top of the boxes indicate significant differences (post-hoc Kruskal-Wallis pairwise comparisons, significance level $p < 0.050$).

5.3.2.2 Analysis of Content

The analysis also investigated the content of the backchannel communication. Table 5.6 summarizes how the backchannel communication distributes along the taxonomy defined by the coding scheme (see Section 5.2.1.1). As can be seen, by far the greatest parts of the backchannel communication consist of *content-oriented* comments, except for Course **P1** in which *content-oriented* and *independent* communication are of the same scale.

		P	W	L	P1
co	Mdn (MAD)	36 (25.20)	14 (13.34)	17 (23.72)	17 (16.31)
	(Min, Max)	(8, 109)	(0, 40)	(0, 63)	(5, 96)
i	Mdn (MAD)	4 (4.45)	0 (0)	4 (4.45)	11 (8.90)
	(Min, Max)	(0, 14)	(0, 18)	(0, 52)	(3, 72)
o	Mdn (MAD)	0 (0)	0 (0)	0 (0)	5 (5.93)
	(Min, Max)	(0, 2)	(0, 3)	(0, 3)	(0, 34)
pe	Mdn (MAD)	0 (0)	0 (0)	0 (0)	0 (0)
	(Min, Max)	(0, 16)	(0, 5)	(0, 6)	(0, 18)
po	Mdn (MAD)	0 (0)	0 (0)	0 (0)	1 (1.48)
	(Min, Max)	(0, 2)	(0, 6)	(0, 2)	(0, 13)

Table 5.6: Overview of the Backchannel Communication by Content (co = *content-oriented*, i = *independent*, o = *organizational*, po = *process-oriented*, pe = *participation-enabling*).

For the lectures of Course **P** there were differences in the kind of communication (Kruskal-Wallis, $N_{co} = 564$, $N_i = 68$, $N_o = 6$, $N_{po} = 9$, $N_{pe} = 33$, df = 4, H = 38.05, $p < 0.001$; see Figure 5.4a). The backchannel communication was mainly content-oriented (post-hoc Kruskal-Wallis pairwise comparisons, $N_{co} = 564$, $N_i = 68$, $N_o = 6$, $N_{po} = 9$, $N_{pe} = 33$, all $p < 0.001$). *Independent* communication was the second largest part of communication and differed from *organizational* and *process-oriented* communication (post-hoc Kruskal-Wallis pairwise comparisons, $N_i = 68$, $N_o = 6$, $N_{po} = 9$, all $p < 0.013$), but did not differ from *participation-enabling* communication (post-hoc Kruskal-Wallis pairwise comparisons, $N_i = 68$, $N_{pe} = 33$, $p > 0.085$).

There were also differences in the kinds of communication on the backchannel in the lectures of Course **W** (Kruskal-Wallis, $N_{co} = 231$, $N_i = 28$, $N_o = 4$, $N_{po} = 12$, $N_{pe} = 7$, df = 4, H = 34.47, $p < 0.001$; see Figure 5.4b). *Content-oriented* communication predominated (Kruskal-Wallis pairwise comparisons, $N_{co} = 231$, $N_i = 28$, $N_o = 4$, $N_{po} = 12$, $N_{pe} = 7$, all $p < 0.001$). No differences were found between *independent*, *process-oriented* and *participation-enabling* communication (post-hoc Kruskal-Wallis pairwise comparisons, $N_i = 28$, $N_{po} = 12$, $N_{pe} = 7$, all $p > 0.085$).

Also for the lectures of Course **L**, differences were found in the kinds of communication (Kruskal-Wallis, $N_{co} = 352$, $N_i = 102$, $N_o = 7$, $N_{po} = 8$, $N_{pe} = 18$, df = 4, H = 12.95, $p < 0.012$). *Content-oriented* comments predominated (Kruskal-Wallis pairwise comparisons, $N_{co} = 352$, $N_i = 102$, $N_o = 7$, $N_{po} = 8$, $N_{pe} = 18$, all $p < 0.019$). Further differences were found between *independent* and the other kinds of communication (post-hoc Kruskal-Wallis pairwise comparison, $N_i = 102$, $N_o = 7$, $N_{po} = 8$, $N_{pe} = 18$, $p < 0.023$). Besides *content-oriented* and *independent* communication, no significant differences were

found among the other kinds of communication (Kruskal-Wallis pairwise comparisons, $N_i = 102$, $N_o = 7$, $N_{po} = 8$, $N_{pe} = 18$, all $p > 0.133$).

Also for Course **P1** differences were found in the kinds of communication (Kruskal-Wallis, $N_{co} = 815$, $N_i = 615$, $N_o = 215$, $N_{po} = 48$, $N_{pe} = 51$, df = 4, H = 82.51, $p < 0.001$). Although *content-oriented* communication still was the largest proportion of the backchannel communication *independent* communication did not significantly differ from it (post-hoc Kruskal-Wallis comparison, $N_{co} = 815$, $N_i = 615$, df = 1, H = 1.99, $p > 0.158$). The high number of *independent* communication was also criticized by the students in the surveys (see Section 5.3.6).
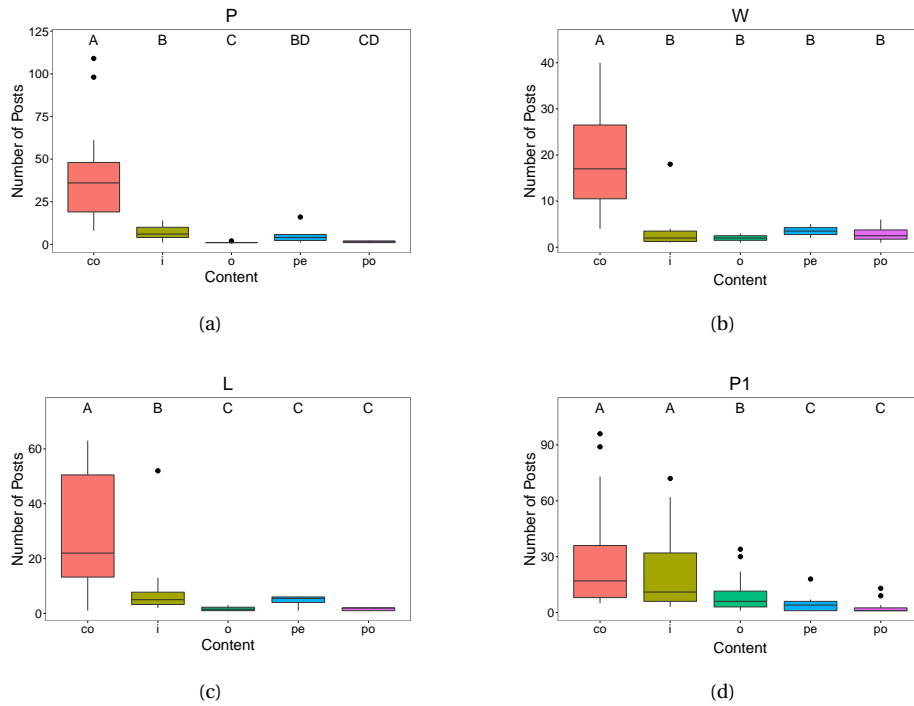
Figure 5.4: The Backchannel Communication by Content (co = *content-oriented*, i = *independent*, o = *organizational*, pe = *participation-enabling*, po = *process-oriented*; box-and-whisker plots with first and third quartiles, median and whiskers (minimum and maximum values); values outside 1.5 times the interquartile range (IQR) are displayed as outliers). Different letters at the top of the boxes indicate significant differences (post-hoc Kruskal-Wallis pairwise comparisons, significance level $p < 0.050$).

5.3.3 Rating

In all investigated courses, students engaged in rating of backchannel comments (see Table 5.7). In most cases, participants rated backchannel comments positively. In the lectures of Course **P1** the differences in ratings were statistically significant (Kruskal-Wallis, $N_{pos} = 2672$, $N_{neg} = 1900$, $N_{off} = 1786$, df = 2, H = 9.869, $p < 0.008$). In particular, positive ratings were most frequent (post-hoc Kruskal-Wallis pairwise comparisons $N_{pos} = 2672$, $N_{neg} = 1900$, $N_{off} = 1786$, all $p < 0.029$). However, no differences were found in the courses **P** (Kruskal-Wallis, $N_{pos} = 447$, $N_{neg} = 104$, $N_{off} = 56$, df = 2, H = 0.444,

$p > 0.801$), **W** (Kruskal-Wallis, $N_{pos} = 266$, $N_{neg} = 54$, $N_{off} = 57$, df = 2, H = 0.786, $p > 0.091$), and **L** (Kruskal-Wallis, $N_{pos} = 170$, $N_{off} = 86$, df = 1, H = 0.452, $p > 0.501$).

Rating activities could be observed in all lectures of the courses **W**, **L** and **P1**, although, over time, the rating activities strongly decreased. Students in the lectures of Course **P1** had the greatest rating activities. In contrast, in the lectures of **P** rating activities came to a complete halt after the fourth lecture.

		P	**W**	**L**	**P1**
Positive Rating	Mdn (MAD)	0 (0)	4.5 (4.45)	8 (8.90)	68 (57.82)
	(Min, Max)	(0, 187)	(0, 142)	(0, 75)	(8, 260)
Negative Rating	Mdn (MAD)	0 (0)	1.0 (1.48)	–	34 (32.62)
	(Min, Max)	(0, 35)	(0, 25)	–	(7, 268)
Off-Topic	Mdn (MAD)	0 (0)	1.0 (1.48)	3 (2.97)	20 (17.79)
	(Min, Max)	(0, 22)	(0, 36)	(0, 75)	(3, 317)

Table 5.7: Overview of Post Ratings by Lecture (Mdn = median, MAD = median absolute deviation, Min = minimum number of ratings, Max = maximun number of ratings).

The large amount of off-topic communication and the high rating activities in Course **P1** provided the opportunity to investigate the extent to which *Off-Topic* ratings correspond to off-topic comments. Therefore, the backchannel communication was divided into two groups: Group *Relevant* comprised all posts not coded as *independent*, and Group *Irrelevant* comprised all posts coded as *independent*. The comparison of the two groups revealed that the posts in Group *Relevant* received significantly fewer *Off-Topic* ratings than the posts in Group *Irrelevant* (Wilcoxon-Mann-Whitney, $N_{Rel} = 1129$, $N_{Irrel} = 615$, $W = 470526.5$, $p < 0.001$).

It might as well be interesting to see whether there is a relationship between engagement in irrelevant communication and the quality of relevant communication, which was measured by the number of positive ratings received for relevant comments (see Figure 5.5). As a result, a weak to moderate positive correlation was found between a student's participation in irrelevant and relevant communication (Spearman's rho, $N_{stud} = 165$, $\rho = 0.44$, $p < 0.001$). In particular no negative correlation could be found as one would have expected.
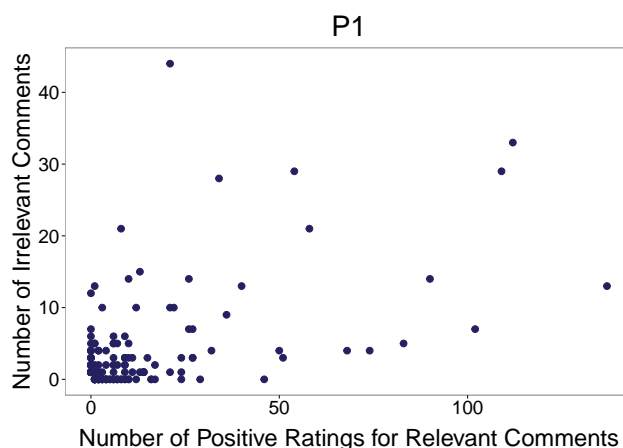
Figure 5.5: Engagement in *Independent* vs. Lecture-Relevant Backchannel Communication in the Lectures of Course **P1** (Spearman's rho, $N_{stud} = 165$, $\rho = 0.44$, $p < 0.001$).

5.3.4 Participation in the Quizzes

In the versions of *Backstage* used in the courses **P** and **W** the numbers of students who were logged in to *Backstage* during the quizzes were not specifically recorded. The analysis of the quiz participation in these two courses thus relies on rough estimates of the login numbers. In the newer versions of *Backstage* used in the courses **L** and **P1** the login numbers during the quizzes were counted more precisely and recorded along with the quiz data. Unlike the login data, the numbers of responses to quizzes were known and did not need to be estimated in all courses.

In all courses the student participation in the quizzes relative to the students logged in to *Backstage* during the quizzes was quite high throughout the semester. This is especially reflected in the more precise measurements of quiz participation in the lectures of the courses **L** and **P1**.

In the lectures of Course **P** a total of 60 quizzes were run. During the quizzes, an estimated median number of 109 students were logged in to *Backstage* (Mdn = 109, MAD = 43, Min = 14, Max = 187) and on average 76 students actually participated in the quizzes (Mdn = 76, MAD = 20.76, Min = 29, Max = 148; see Figure 5.6a).

In the lectures of Course **W** a total of 60 quizzes were run. During these quizzes, an estimated median number of 33 students were logged in to *Backstage* (Mdn = 33, MAD = 8.90, Min = 23, Max = 97). On average 25 students actually participated in the quizzes (Mdn = 24.5, MAD = 5.19, Min = 14, Max = 74; see Figure 5.6b).

In the lectures of Course **L** a total of 60 quizzes were run. A median number of 49 students were logged in to *Backstage* during the quizzes (Mdn = 49, MAD = 8.89, Min = 3, Max = 65). An average of 46 students also participated in the quizzes (Mdn = 46, MAD = 10.38, Min = 0, Max = 64; see Figure 5.6c).

Finally, in Course **P1** a total of 122 quizzes were run on *Backstage*. On average 142 students were logged in to *Backstage* during the quizzes (Mdn = 141.5, MAD = 65.98, Min = 1, Max = 242). About 114 students participated in the quizzes (Mdn = 113.5, MAD = 57.08, Min = 0, Max = 208; see Figure 5.6d).

(a)

(b)

(c)

(d)
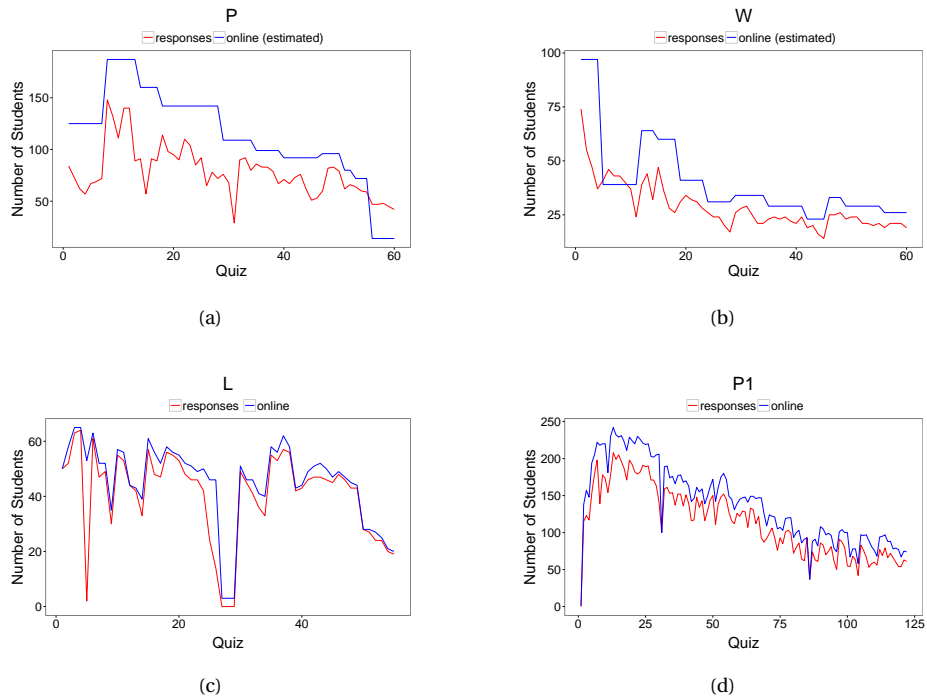
Figure 5.6: Number of Quiz Participants Compared to Students Logged in to *Backstage*. The blue lines in the plots show the numbers of students logged in to *Backstage* during quizzes (roughly estimated for the courses **P** and **W**, more precisely counted in the courses **L** and **P1**), the red lines show the number of students who actually participated in the quizzes. The first sharp downturn in Course **L** was due to technical difficulties with the ARS, the second was due to problems with the Internet connection in the lecture hall.

5.3.5 Comparing Backchannel with Quiz Activities

Another question is whether a student's backchannel activity relates to her quiz activity. As a general result, in all courses weak positive correlations were found between a student's activity on the backchannel and her performance in the quizzes.

First, the relationship between a student's backchannel activity to her performance in the quizzes is investigated. In Course **P** a weak positive correlation was found between the backchannel activity and the quiz performance (Spearman's rho, $N_{stud} = 285$, $\rho = 0.40$, $p < 0.001$). The more a student engaged on the backchannel the better did she peform in the quizzes. Similar results were found for Course **W** (Spearman's rho, $N_{stud} = 131$, $\rho = 0.39$, $p < 0.001$), Course **L** (Spearman's rho, $N_{stud} = 183$, $\rho = 0.44$, $p < 0.001$), and Course **P1** (Spearman's rho, $N_{stud} = 466$, $\rho = 0.43$, $p < 0.001$).

Furthermore, the relationship between a student's quiz performance and the student's engagement in irrelevant communication is analyzed. In Course **P** a weak positive correlation was found (Spearman's rho, $N_{stud} = 285$, $\rho = 0.16$, $p < 0.005$), meaning that the more a student engaged in irrelevant communication the better did she perform in the quizzes. Similar results were obtained for Course **L** (Spearman's rho, $N_{stud} = 183$, $\rho = 0.25$, $p < 0.005$, and for Course **P1** (Spearman's rho, $N_{stud} = 466$, $\rho = 0.39$, $p < 0.001$). In Course **W**, a weak, yet insignificant positive correlation was found (Spearman's rho, $N_{stud} = 131$, $\rho = 0.16$, $p > 0.072$).

5.3.6 Survey Responses

After the courses, students were invited to participate in an online survey. Note that the Likert-type questions are based on a six-point scale where strong agreement corresponds to six points and strong disagreement corresponds to one point, respectively.

Among the courses the administered surveys substantially differed in some of their parts. Except for Course **W**, however, the surveys all contained the questions used to measure the constructs given in Table 5.3. In Course **W** apart from *Backstage* two other social media, *WIS Market* and *Termina*, were used. The survey at the end of the course targeted the use of all these three media. Thus, the items used in the survey were mostly open items. Regarding Course **P1** the survey should be kept rather brief. Thus, this thesis abstains from asking what the participants particularly liked on *Backstage* and what the participants liked about the functions of *Backstage* as in the two courses **P** and **L**. However, criticism of particular functions was also provided by the respondents in the open items. Note that the survey responses to open items reflect individual opinions that were unsystematically collected. Some student responses do not refer to what was being asked in the open items.

5.3.6.1 Course P

In Course **P** a total of 38 students completed the survey (see Appendix A.13). As a general result, the use of *Backstage* was assessed to be good. Thus, the respondents agreed that *Backstage* promotes interactivity in the lectures (INTERACTIVITY; Mdn = 5.25, MAD = 0.62). A tendency towards agreement was found for ratings as a useful means to mark relevant content on the backchannel (RATING; Mdn = 4.33, MAD = 0.99). Agreement was also found on *Backstage* as a useful means to foster awareness (AWARENESS; Mdn = 5.00, MAD = 0.59). The responses also indicate that *Backstage* was considered useful for reworking lectures (REWORK; Mdn = 4.75, MAD = 1.11).

Also in the open items, the respondents in Course **P** largely showed a positive attitude towards *Backstage*. In the following, a selection of answers by respondents when asked

what they particularly liked on *Backstage* is given (translated into English; for the original responses in German, see Appendix A.12.1):

I particularly liked on Backstage *that ...*

> "... one post questions and comments on the current slide."
>
> "... it motivates you to stick to the lecture."
>
> "... quizzes were used. The lecture became more interactive through *Backstage*."
>
> "... the questions of fellow students contributed to one's own understanding, and so did the quizzes."
>
> "... it makes the lecture more interactive and it makes it easier to ask questions that otherwise might not have been asked."
>
> "... you can see the slides of the lecture and others often post comments and questions. The auto navigation always jumps to the current slide."
>
> "... the slides are clear and concise and the quizzes help to consolidate knowledge."

Also most of the respondents (28 of 38 respondents) expressed a positive attitude towards a use of *Backstage* for reworking lectures. For example, respondents highlighted the usefulness of the comments posted on the backchannel and the results in the quizzes. However, some respondents stated that they rather prefer working directly with the PDF slides and hand-written notes.

Most of the criticism referred to usability issues. Thus, some respondents disliked the use of *Twitter*-like textual modifiers, e.g., for creating private notes. The display of lecture slides as raster graphics making selection of text on the slides impossible was criticized as well. One respondent criticized the relatively high efforts required for posting a comment to be a usability issue. Recall, however, that this "high effort" was intended as a means to keep the backchannel focused on the lecture's content. In a similar vein, another respondent disliked the annotation-approach of the backchannel. One respondent criticized that once submitted, a response to a quiz cannot be corrected, which, too, was intended. Another group of usability issues addressed features respondents missed on *Backstage*. For example, students recommended better support for tablets, which was actually lacking in that version of *Backstage*. Other students missed export functionalities allowing them to export the slides together with the backchannel comments and quiz results to PDF for offline use. One student suggested to add audio recordings to *Backstage*. When asked which feature could be abandoned, several students mentioned rating of comments.

Regarding the use of *Backstage* by the teaching team, students criticized that the lecturer could have answered the questions on the backchannel in a more timely manner. Another student encouraged the lecturer to more frequently respond to questions raised on the backchannel after the lectures. When asked what could be improved on the quizzes, many students criticized rather unclear and confusing formulations of the quiz questions and response options. When asked, what could be improved on the lectures, students found that three hours for such demanding lectures is too long and recommend to split the lecture into two 1.5 hours lectures.

As final remarks on the use of *Backstage* two students responded with the following:

> "At the beginning of the course I wouldn't have thought how great *Backstage* is. Big praise and keep it up!"

"An excellent idea that shouldn't be missing in any course."

5.3.6.2 Course W

In Course **W** a total of 11 students completed the *Backstage* part of a survey administered at the end of the course (see Appendix A.13). The respondents of the survey in Course **W** were quite critical of *Backstage*. Only one student thought that *Backstage* is a good idea. Several students mentioned that the low activity on the backchannel prevented them from partaking. Four out of eleven students found a use of *Backstage* in the course merely pointless. A student's response that summarizes the other respondents' views is the following (translated into English; for the original responses in German, see Appendix A.12.2):

> "[...] I do not find it necessary to have the slides digitally in front of me and it distracts me: I can learn very well by listening but when I do not see the lecturer (because of a digital device) I miss a lot more. I think first-year courses are better suited for *Backstage* because the size of the audience makes it more difficult to ask questions [...]. There, I think *Backstage* is more appropriate than in a fifth-term course which 30 to 40 people attend. This is a size of a seminar and not the right environment."

5.3.6.3 Course L

In Course **L** a total of 18 students completed the survey (see Appendix A.13). The responses indicate a good to strong agreement on the use of *Backstage* to foster interactivity in lectures (INTERACTIVITY; Mdn = 5.42, MAD = 0.62). Regarding rating of comments the respondents tended to agree that it is useful (RATING, Mdn = 3.83, MAD = 1.24). In turn, *Backstage* was found to be a good means to foster awareness (AWARENESS; Mdn = 5.30, MAD = 1.04). Similarly, the responses indicate that *Backstage* is useful for reworking lectures (REWORK; Mdn = 5.12, MAD = 0.93).

In the open items, the students of Course **L** largely showed a positive attitude towards a use of *Backstage*. The following responses were given to the question what students particularly liked on *Backstage* (translated into English; for the original responses in German, see Appendix A.12.3):

I particularly liked on Backstage *that ...*

> "... active participation in the lecture was encouraged."

> "–If you cannot keep track once in a while you can go back in the slides.
> –You become awakened by the quizzes in case you did not pay attention.
> –You can follow the lecture from home at least to some extent."

> "... interactivity has been promoted and a step is made away from the standard lecture (lecturer talks)."

> "... you could post questions and answers anonymously."

> "... the students' questions were fixed in written form which supported the preparation to the exam."

> "... you could go back in the slides in case something started off moving too fast; that the mock exams were conducted on it; that you could answer the quizzes with a smart phone."

Most of the respondents (10 of 14 respondents) expressed a positive attitude towards using *Backstage* for reworking lectures. The main reasons were that errors on slides were corrected using backchannel comments and that students wanted to take into account the backchannel exchange as well as the quiz results when reworking lectures.

Criticism of *Backstage* mostly referred to stability and usability issues (although the former was mostly related to issues with the Internet connection in the lecture hall). A respondent suggested to put more focus on rating of comments. The use of *Twitter*-like textual modifiers instead of buttons was also criticized. A respondent criticized the layout of the user interfaces as unclear in parts . Several respondents suggested to improve the summary statistics of the quizzes. Other usability issues referred to missing but desirable functionalities. The lack of export functionalities has been mentioned several times. Also, the possibility to delete or modify submitted comments has been mentioned. Several respondents suggested to give users an overview of the questions that have lately been answered as well as an overview of one's own questions which have not received responses, yet. One respondent suggested to add audio recordings to *Backstage*. Also, one respondent suggested to find alternatives to displaying slides as raster graphics.

Regarding the use of *Backstage*, numerous respondents expressed the wish that more fellow students participate in the backchannel. Also, the respondents would have appreciated if more questions had been answered by the teaching team. Regarding the quizzes, several respondents criticized too long waiting times during the quizzes. Other respondents found the quizzes too easy and suggest more elaborate quizzes.

As final remarks on *Backstage*, the following two responses were given:

> "First I would like to express my gratitude for the software. Despite criticism that is likely to come up, the software changes daily life in lectures clearly into a positive direction. It did not work perfectly smoothly; however, as a current beta it fairly has potential and I would be glad if *Backstage* accompanies me in the one course or the other in the semesters to come."

> "In any case a positive effect on the lectures, I pretty much liked it."

5.3.6.4   Course P1

In Course **P1** a total of 51 students completed the survey (see Appendix A.13). The respondents valued *Backstage* as a useful means to foster interactivity (INTERACTIVITY; Mdn = 5.33, MAD = 0.49). The respondents also agreed on the usefulness of rating on *Backstage* (RATING; Mdn = 4.67, MAD = 0.99). Further agreement could be found on the usefulness of *Backstage* to promote awareness (AWARENESS; Mdn = 4.80, MAD = 0.89). However, the respondents were unsure as to the usefulness of *Backstage* for reworking lectures (REWORK; Mdn = 3.75, MAD = 1.11).

The major criticism expressed by the respondents was the fairly high amount of off-topic comments posted on the backchannel. One respondent suggested a use of visual cues that allow quick recognition of comments that received many positive ratings or that received many off-topic ratings. Other respondents identified the provision of perfect anonymity as a source for the high off-topic communication and suggest to abandon this feature. The responses regarding note taking indicates that the students did not know about this functionality. As for the quizzes, numerous respondents had the opinion that the quizzes could have been more difficult and more related to the exam. Several respondents suggested that the teaching team should be better identifiable on *Backstage*.

As final remarks the respondents expressed mixed opinions on the use of *Backstage* (translated into English; for the original responses in German, see Appendix A.12.4):

"A nice gimmick, not particularly useful for lectures."

"Unfortunately a bit buggy but a nice idea."

"Interesting system, requires decent behavior of students, otherwise it is more distracting than helpful."

"An excellent enhancement."

"A good system for first-year courses when many students are in the lecture hall and thus the communication between students and lecturers is difficult, however still in need of improvements."

"The system is really good! It makes lectures more fun! :)"

"I very much liked it. That it also worked on the tablet was very practical because I did not have to bring a laptop. [...]"

"It worked astonishingly well. I later used it just to participate in the quizzes and to use the automatic navigation in the slides on the laptop."

"Very helpful because of the use by the lecturer, tutors and fellow students. Questions have been answered and thus the lecture material became much more comprehensible. There were enough quizzes yet not disturbingly many of them. Most of all, the explanations to correct answers and why the other answer options were wrong resolved remaining issues. Very useful and desirable in other courses."

5.4 Discussion

The results obtained in the case studies give rise to the following conclusions, which are discussed in more detail:

- The use of *Backstage* in the four courses can be considered successful.
- *Backstage* promotes learning-related awareness and activities.
- The backchannel and the quizzes promoted two distinct kinds of activities in the lectures.
- Rigorous sanctioning of bad behavior could be counterproductive.
- A successful use of *Backstage* also depends on the participants.

The use of *Backstage* in the four courses can be considered successful

The students' responses to the surveys as well as the findings obtained by analyzing the backchannel communication and the quiz participation suggest that *Backstage* is an appropriate means to foster interactivity in lectures (cf. Question **Q1**, p. 61). All four courses exhibited an active exchange on the backchannel. The analysis of post categories suggests that the backchannel communication is not directed towards one particular purpose, e.g., question asking. The use of the two post categories *Answer* and *Remark* varied notably between the courses. For example, in the courses **L** and **P1** *Answer* was used more often than *Remark*. In contrast, students in Course **P** used *Remark* more often than *Answer*. An explanation could be that while students in Course **P** were instructed in how to use the post categories, in the other courses these instructions were omitted. In

this way, the students of the courses **L** and **P1** seemed to develop their own conventions in the use of the post categories, especially for note taking.

In the analysis of the backchannel communication of Course **L** empty *Answer* and *Question* notes (i.e., posts only visible to the respective author) were recognized. It seems that the icons for the two categories were used as mental flags: *Answer* icons seemed to indicate "important" or "notice", whereas *Question* icons seemed to indicate puzzlement requiring rework of the slide after the lecture. It would be interesting to investigate further how students creatively made use of the post categories as such mental flags.

In general, the usefulness of *Backstage* for reworking lectures has been rated relatively high by the students (cf. Question **Q2**, p. 61). This finding is quite surprising, keeping in mind that no functionality has been integrated to particularly support the rework of lectures and that the actual use of *Backstage* outside lectures drastically differed between the courses. The login data suggests that especially in the courses **L** and **P1** students used *Backstage* for rework. This finding was unexpected for Course **P1**. A high number of logins was especially recognized for the 13th lecture which was used for a mock examination. This finding is notable, considering that *Backstage* was set out for a use during lectures as a means to promote interactivity and that other tools were intended for use outside the lectures (see Section 5.1.2).

Finally, the positive feedback of students indicates that the use of *Backstage* was successful (cf. Question **Q3**, p. 61). The criticism by students mainly addressed usability and stability issues which were already recognized. The students of the courses held at LMU were aware that *Backstage* was a research prototype that never claimed to achieve the high convenience and stability levels of a commercial product. Many students recognized the potential of the prototype and gave valuable feedback about possible and necessary improvements (see below).

The observed decrease of student logins throughout the semester may be influenced by the natural loss of lecture attendees in the course of a semester. There is no reason to believe that *Backstage* has any mitigating effects on drop-outs or warrants high rates of lecture attendances. Considering that a use of *Backstage* requires more effort by students during lectures, e.g., by participation in the quizzes, students with little motivation may even become more deeply inclined to stay away from the lectures. Although the logins during lectures of all four courses showed decreasing trends the use of *Backstage* never came to a complete halt.

Backstage promotes learning-related awareness and activities

The surveys in the courses **P**, **L** and **P1** indicate that *Backstage* contributes to learning-related awareness (cf. Question **Q4**, p. 61). That is, both the quiz results and the backchannel communication help the student gather information on their learning in relation to others. The high score achieved in Course **L** may be a result of the revision of the lecture material that better aligned teaching with the use of *Backstage*. In contrast, the large amount of off-topic communication may explain the slightly lower score *Backstage* achieved in Course **P1**.

Moreover, in three of the four courses the backchannel communication was highly focused on content (cf. Question **Q5**, p. 61). Notably small amounts of backchannel comments were off-topic, which frequently occurred at the beginning of a course for testing and playing with the tool. Even in Course **P1** in which the backchannel comprised much off-topic communication, content-related communication was numerically predominant. Nevertheless, it is understandable that students and lecturers can feel disturbed by any numbers of off-topic comments. In Course **P1** the excessive off-topic communication gave rise to students' complaints. It was thus suggested to delete the

Backstage accounts of the disturbing students. As discussed below, however, social control mechanisms and filtering of the backchannel may be more sensible.

In all courses the students used *Backstage* also for question asking. In another evaluation of the student activities during 21 lectures Gehlen-Baum (to appear) reports that students asked the lecturer about 49 questions orally. As a rather coarse estimate, this corresponds to 2.33 oral questions per lecture. In all four courses, students asked notably more than 2.33 questions per lecture on the backchannel, additionally to those questions that students asked orally.

The participation in the quizzes compared to the number of students on *Backstage* was notably high. Besides the participation data, the survey respondents appreciated the quiz functionality as a great means to promote interactivity and participation. The high acceptance of quizzes corresponds to the findings reported on in the literature (e.g., Kay and LeSage, 2009; Lantz, 2010). Thus, also in case of *Backstage*-supported lectures quizzes have shown to be a remarkably useful means to activate students and to help bring the lecture to students' attention again (cf. Young, Robinson, and Alberts, 2009).

There were weak indications that backchannel participation is positively connected to the participation in the quizzes (cf. Question **Q6**, p. 61). More importantly, however, no evidence suggests that the two components, backchannel and quizzes, interfere with each other. That is, if participation in the backchannel exchange distracted students, one could expect a lower quiz performance among the students involved in the backchannel. Also, the performance in the quizzes did not differ between students involved in off-topic exchange and students involved in lecture-relevant communication.

As one might have expected, fewer students participated in the backchannel than in the quizzes. In the surveys, the students welcomed more participation of others on the backchannel, which, however, should not be considered as a negative result. As Peters (2014) points out, many students prefer learning by observing others.

Further indications that *Backstage* supports lecture-related activities are given by Gehlen-Baum, Weinberger, et al. (2014). The authors compare differences in the students' activities in two conventional lectures and one lecture of a *Backstage*-supported course in the summer term 2012. For the investigation the student activities were classified according to four categories: (1) focused (i.e. active) use of media for lecture-related activities, (2) peripheral (i.e. background) use of media for lecture-related activities, (3) focused use of media for lecture-unrelated activities, and (4) peripheral use of media for lecture-unrelated activities. The study observed 41 students in the *Backstage*-supported lecture and 34 students in the two conventional lectures. The authors found that in the conventional lectures the students used their mobile devices mostly for lecture-unrelated activities. However, in the *Backstage*-supported lecture the students used their mobile devices mostly for *Backstage* and, thus, to a lesser extent for lecture-unrelated activities compared to lectures without *Backstage* running.

Rigorous sanctioning of bad behavior could be counterproductive

The relatively high amount of off-topic communication in the lectures of Course **P1** at Saarland University was quite surprising. The teaching teams at LMU (Ludwig-Maximilian University) did not have to deal with similar excessive off-topic communication in any of their courses so far. Even for Course **W** in which the students were most sceptical about the usefulness of *Backstage* the students refrained from partaking in the backchannel at all rather than engaging in excessive off-topic communication.

This evaluation took the opportunity to investigate the overall participation of students in Course **P1** who exposed non-desirable behavior on the backchannel. Surprisingly, only a small number of students appeared to engage in nothing else but excessive

off-topic communication. A comparison of ratings received for relevant comments versus number of off-topic posts showed a weak to moderate correlation. This means that students who contributed many off-topic comments also contributed relevant comments praised by other students. Also regarding the participation in the quizzes, no negative correlation between a student's contribution to off-topic communication and quiz performance was found, contrary to what one would have expected. Thus, students who exposed undesired behavior also exposed desirable behavior. The findings suggest that a rather rigorous sanctioning, e.g., by denying those students further access to *Backstage* is perhaps counterproductive.

As for *Backstage*, social control or sanctioning exerted by the classroom community may thus be more sensible. In social media, desirable behavior is often incentivized by reputation mechanisms (Vassileva, 2012). Users who expose desirable behavior gain reputation, whereas users who behave badly lose reputation. A reputation mechanism on *Backstage* may encourage a collective guidance of users towards desirable behavior. Such a reputation mechanism that might be suitable for *Backstage* is proposed in Chapter 7.

Another approach has been pointed out by Hiltz and Turoff (1985). The authors stress the importance of giving control not only to senders but also to receivers of messages. As for *Backstage* the approach entails that students should be able to restrict their individual backchannel time-lines based on attributes such as the authors' user names, the authors' reputation, certain post categories, keywords, and ratings of comments in a similar manner as the lecturer. Extensive filtering can be used to withdraw attention on an individual basis and thus to de-incentivize undesirable behavior.

It should be noted, however, that social control mechanisms are inherently limited, just like rigorous centralized sanctioning mechanisms.

A successful use of Backstage also depends on the participants

A thriving social medium is not only governed by rules integrated within software but also by social norms that exist outside the social medium. It is impossible to integrate all rules that ensure a proper use within software (Schmidt, 2013). In other words, social media are to certain extents susceptible to misuse. Proper guidance and incentive mechanisms can mitigate misuse but, eventually, the users and their attitudes towards a medium are substantial for the medium's success. Guidance and conventions outside the medium that foster liability are thus indispensable.

A proper use of *Backstage* heavily depends on the lecturer. The evaluated courses were held by two motivated lecturers who took teaching very seriously. Both lecturers received much student praise for their efforts. For example, in the courses at LMU the students praised the conciseness of slides that not only conveyed information in a clear manner but also contributed to a more convenient use of *Backstage*. Additionally, an experienced member of the teaching team regularly participated in the backchannel from remote and answered student questions raised on the backchannel in a timely manner. These efforts certainly contributed to a higher participation of students, to more lecture-relevant exchanges on the backchannel, and to a greater visibility of the teaching team on *Backstage*.

However, proper guidance of the teaching team in using *Backstage* may be necessary. Unlike the courses at LMU, students in Course **P1** used *Backstage* to a notable extent for off-topic communication. Students willing to properly use *Backstage* complained about the high numbers of irrelevant posts. Perhaps, such a misuse could have been mitigated by the teaching team. The lecturer did not make misuse of *Backstage* subject to discussions in the lectures. Also, the backchannel was not configured in a way that would make

misuse more difficult for students. For example, perfectly anonymous communication, which the complaining students considered to promote off-topic communication, could have been disabled on *Backstage*. More drastically, the lecturer could have decided to temporarily disable audience-to-audience communication altogether. Also, the teaching team did not seem to be visible on the backchannel to a great extent. One observation of the analysis was that the teaching assistant, the member of the teaching who participated on the backchannel the most, was logged in to *Backstage* as a student and not as a member of the teaching team.[13]

In the same way, a successful use of *Backstage* depends on the students' attitudes towards a constructive use. In Course **W**, students were completely reluctant to participate. The inclination of a few students in Course **P1** towards off-topic exchange introduced distraction and thus certainly impinged on the success of *Backstage* for learning. To appeal to a proper use, the students in the courses **P**, **W**, and **L** received a special introduction to *Backstage*. The introduction emphasized that a successful use of *Backstage* greatly depends on the students attitude towards proper use and asked them to be nice to each other on the backchannel. The students seemed to accept these suggestions. To encourage proper use by students, the findings suggest that students should receive special and thorough introduction to *Backstage*. Computerized forms of educational scripts (Kollar, Fischer, and Slotta, 2005; Weinberger, 2011) can be an interesting means to guide and help students in appropriately using *Backstage* for learning-related activities.

5.5 Concluding Remarks

The evaluation provided evidence that *Backstage* supports lecture-related interactivity in large-class lectures. Students were satisfied with *Backstage* and found *Backstage*-supported lectures more pleasant. Gehlen-Baum, Weinberger, et al. (2014) could confirm that students engaged in learning-related activities more frequently than in lectures without *Backstage*, since Internet-enabled devices brought into lectures was use to a lesser extent for lecture-irrelevant activities.

Further studies are conceivable and also necessary to fully understand the potential of *Backstage*. Thus, more attention should be paid to the role of the lecturer and the teaching team. However, to have more lecturers use *Backstage*, usability and stability of *Backstage* need to be improved. By now, *Backstage* has been developed foremost for evaluation purposes. Functions indispensable for a platform for public use were implemented only provisionally or omitted altogether. In the courses in which *Backstage* has been used the students as well as the lecturers provided feedback regarding missing and desirable functionalities (for possible improvements and extensions see Chapter 8).

Another question interesting to investigate is whether *Backstage*, in combination with a suitable instructional approach, improves learning. The settings in the four courses made such an investigation impossible. For example, changes to the study regulations over the semesters influenced examination grades. As another example, the lecture contents and material of Course **L** have been considerably revised. Although the revision encouraged a use of *Backstage*, it makes a comparison of examination grades with former semesters impossible. Measuring the effects of *Backstage* on learning outcomes requires controlled conditions that could not be met in the recent semesters. Also, data collection needs to be improved. More precise and thorough measurements are necessary. For example, other variables such as student motivation which is likely to affect both learning outcomes and use of *Backstage* need to be taken into account.

---

[13]Registering as a student was certainly unintended by the teaching assistent. Perhaps the mistake was also caused by usability issues of *Backstage*.

Part III

Beyond Backstage 1.0

RATING AND RANKING OF POSTS

This chapter is a generally revised and extended version of Hadersberger, Pohl, and Bry (2012).

Computer-Mediated Communication (CMC) can quickly become confusing and incomprehensible, even when the number of participants is small. CMC usually varies in both relevance and quality, which makes indicators of relevance and quality desirable, e.g., for automatically filtering out irrelevant communication. To this aim, it is helpful if participants can rate CMC messages.

On a CCS such as *Backstage* student ratings serve as a valuable basis for communication summarization (see Section 4.1.1). Lecturers may rely on summarization to a great extent, since they can only occasionally pay attention to the backchannel. Keeping track of the relevant backchannel communication is hardly possible for them without the help of the audience. Rating makes possible to filter out irrelevant posts as assessed by participants and enables to order, or to rank, backchannel posts by relevance. In this way, rating can be used to collectively direct the lecturer's attention to what the students find particularly relevant for their good reception of the lecture.

6.1 Relationship between Rating and Ranking

Rating and ranking often occur interleaved, since rankings are frequently determined by ratings. Because of their close relationships the concepts of rating and ranking may easily be confused. Rating and ranking can be distinguished as follows: Rating refers to a person assessing an item's quality by assigning some concrete value to that item (e.g., a school grade). Ranking, in contrast, relates two or more posts to each other, thereby specifying a (total) order. Note that rating and ranking are two different and independent concepts. For example, without using rating, a ranking can also be directly determined by pairwise comparisons of the form "post *A* is more relevant than post *B*".

When rating is used as a basis for determining a collective ranking two possible ways exist: (1) aggregate individual ratings to a collective rating of items and derive from this collective rating a collective ranking; (2) determine individual rankings and aggregate these individual rankings to a collective ranking of items (see Figure 6.1). The transition

89

from rating to ranking is achieved by ordering items according to their ratings. The transition from the individual to the collective level is achieved by aggregating individual ratings or rankings. In the Web, a collective ranking is mainly determined by ordering items according to a collective rating. In contrast, economical and political sciences are mainly concerned with the aggregation of individual rankings (or preferences) to obtain collective rankings (e.g., Gaertner, 2009).
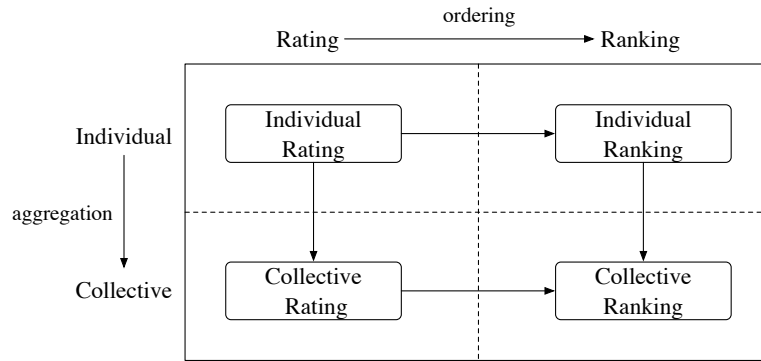
Figure 6.1: Ways to obtain a collective ranking from individual ratings

6.2 Rating

Rating is ubiquitous in the Web and one of the basic forms of empowering users in social media. It is especially known for its use on commercial websites (e.g., rating products or sellers) and in social media as a quick and easy way to get feedback on an item's quality and to allow finding high-quality items (Bian et al., 2008b; Marmolowski, 2008). Rating can be distinguished in two groups: explicit and implicit forms of rating.

6.2.1 Explicit Forms of Rating

Explicit forms of rating comprise all rating systems that require an intentional assessment by a user, meaning that the user is conscious of her assessment (Das Sarma et al., 2010; Lerman, 2009). Possibly the simplest form for explicit rating is to "like" an item (Lerman, 2009). A more elaborate form of liking is to specify a degree to which an item is liked, e.g., as in *Amazon*'s five-star rating.

Normalization is often used to keep the score within a certain range. The downside of normalization is that the reliability of the normalized score is not apparent to the users. For example, an item with an average rating of two out of five stars voted by only one person does not seem as bad as an item with the same average rating voted by, say, twenty people. The opinion of the larger group seems to be more reliable. Another explicit form of rating scheme gives the possibility to not only vote positively for an item, but also negatively or even express neutrality (Agichtein et al., 2008; Bian et al., 2008a; Harry, Gutierrez, et al., 2008; Kamvar, Schlosser, and Garcia-Molina, 2003).

Negative ratings are sometimes desired to give users the possibility to "punish" inappropriate items. However, negative ratings may require more elaborate calculation methods. For example, simply adding +1 for each positive rating and −1 for each negative rating may lead to undesired cancellation effects, which may lead to confusion,

when users are not able to distinguish between items not being rated at all and items assessed discordantly by the raters.

Explicit ratings are often considered to reflect deliberate and intentional assessments of an item's quality, which requires proper incentivization of the associated effort (Nichols, 1998; Vassileva, 2012). Since incentive mechanisms for explicit rating schemes are quite difficult to develop, explict forms of rating often face issues of data sparsity. Services such as collaborative filtering or recommender systems are often built on user ratings and require extensive data to perform well. The quality of such services deteriorates if users are reluctant to rate. As a remedy, explicit forms of rating can either be replaced or extended by implicit forms of rating.

6.2.2 Implicit Forms of Rating

Implicit forms of rating tap into a user's general activities to extract rating-related information. That is, instead of asking users to explicitly assess the quality of an item, the user's actions on the item are interpreted as ratings. For example, purchasing an item or adding an item to one's list of favorites can be considered as a positive rating of that item. In implicit forms of rating, the user is usually unconscious about her quality assessment, since rating happens in the course of merely "using" an online environment (Das Sarma et al., 2010; Lerman, 2009). Apart from purchasing an item or adding an item to a wish list, other kinds of actions can be chosen as a source of implicitly ratings, depending on the context of use. For example, number of clicks on or frequency of revisiting items can be seen as interest and positive feedback.

A difficulty associated with implicit forms of ratings is that these mechanisms are susceptible to misinterpretations. For example, "misclicks" on links might erroneously be interpreted as positive ratings (Agichtein et al., 2008; Bian et al., 2008b). As another example, the frequency of revisiting an item can just as well be the result of a user who got lost on the platform and keeps returning to the item where she started to get confused. Interpretation of actions other than clicking on links might be a more reliable indicator, e.g., measuring the time that was spent on an item (Nichols, 1998). Furthermore, answering questions about an item on a discussion board can also be considered as interest in an item and thus, as a positive rating.

By tapping into a user's actions to extract quality assessment information, implicit forms of rating can be used to overcome data sparsity. In particular, the user does not need incentives to provide for ratings. Hence, incentive mechanisms are dispensable. However, implicit forms of rating can be more complex to handle, e.g., when the extraction of quality assessments requires tracking of a user's actions.

6.2.3 Considering Rating in Backstage

On *Backstage*, a rating scheme with three kinds of ratings for assessing the quality of backchannel posts is implemented: participants can rate backchannel posts positvely, negatively or mark posts as off-topic. A participant may rate a backchannel post at most twice, viz., either positively or negatively and possibly off-topic. The controls for rating are displayed in the expanded view of a backchannel post and shown together with other post-related functions only at explicit request (Figure 4.4, p. 46). Absolute numbers and proportions of ratings are shown along with a backchannel post to provide context for interpretation of a backchannel post's ratings.

This thesis considers rating foremost as a means of communication summarization. Whether rating also serves educational purposes, e.g., as a source of feedback, is questionable. Ratings, or more generally rewards, do not provide sufficient task-related

information to be considered as learning-related feedback (Deci, Ryan, and Koestner, 1999). However, the lack of relatedness to educational purposes provides room for implicit forms of rating, possibly accepting misinterpretations of backchannel activities for the sake of a reduced effort.

Several possibilities of implicit rating in *Backstage* exist. For example, approval and rejection can be redeclared so as to emphasize their purposes for the lecturer. Accordingly, approval can be emphasized as a vote for forwarding a post to the lecturer. Likewise, rejection votes against forwarding of the post to the lecturer. Redeclaring ratings this way may also remedy a feeling of being unjustly punished, which may negatively affect a participant's motivation to continue partaking.

Another form of implicit approval would be to enable participants to maintain favorites lists of backchannel posts. Adding to and removing from these favorites lists is then considered as ratings. Favorites lists create personalized views of the backchannel and in this way enable participants to also exert an individual control of the backchannel. As Hiltz and Turoff (1985) point out, most CMC systems put much control on the senders, but few on the receivers of messages.

In a sense, favorites lists of backchannel posts complement the functionality of forwarding posts to the lecturer in that a ranking of backchannel posts at the lecturer's user interface can be considered as a collectively maintained favorites list. However, note that there is a qualitative difference between explicit rating and maintaining lists of favorite posts. While a rating is given once and for all, posts may be added and removed from favorite lists. Removing posts from a participant's favorites list can either be considered as a negative rating, thus allowing participants to rate posts positively or negatively more than once, or retracting a post's positive rating. That is, an implicit rating mechanism needs to specify the meaning of removing posts from a favorites list.

6.3 Ranking

Ranking can be found in various situations, for example listing the best players of an online game or ordering search results. For *Backstage* a ranking method is needed that aggregates the participants' ratings of backchannel posts into a ranking of posts. Ranking methods fall into two groups, non-parametric and parametric methods.

6.3.1 Non-Parametric and Parametric Ranking Method

Non-parametric ranking methods are methods that do not require externally set parameters. Possibly the simplest non-parametric ranking method is sorting items according to their total or mean ratings. Another form of non-parametric ranking is to have users determine a ranking of items, e.g., by having user pairwise compare items.

The so-called Hasse method (Al-Sharrah, 2010) is another non-parametric ranking method that offers the opportunity to create a ranking of items by directly comparing two or more properties of the items. However, items could be incomparable if they are not better or worse in all relevant properties. Hence, the Hasse method might result only in a partial order.

The so-called Copeland Score (Al-Sharrah, 2010) combines the idea of the Hasse method with the direct comparison of items. Like the Hasse method, items with several properties are compared against each other. The Copeland Score for each item denotes the number of wins minus the number of defeats (a zero score reflect incomparability) while it is compared to all alternatives. Afterwards the items are ranked according to their descending Copeland Score. It has to be noticed that the Copeland Score results in

a total, but not necessarily in a strict order, which means there can be two items with the same score.

Parametric ranking methods, in turn, rely on externally set parameters that significantly influence how a ranking is determined (Al-Sharrah, 2010). Of course, finding appropriate values to those external parameters is an elaborate task that requires expertise and domain knowledge.

Any non-parametric ranking method can be turned into a parametric method by introducing weighting parameters, e.g., weighting the ranking of users by the users' reputation or expertise. PageRank is another ranking method that is part of the *Google* search engine (Brin and Page, 1998; also see Section 7.3.1.2). PageRank computes a ranking of Web pages by the link structure of the Web. The so-called leap factor, denoted by α, is a parameter that can be seen as the degree to which the actual links among Web pages are considered in the determination of the ranking.

6.3.2 Obtaining a Collective Ranking by Aggregating Individual Ratings

On the Web, a ranking of items is typically determined by aggregating the users' individual ratings of items to a collective rating and then sorting the rated items by the aggregated ratings. One problem that needs to be addressed in the aggregation of ratings is finding a way to avoid a potential overestimation of an item's value due to few available ratings and a potential underestimation of an item's value due to many ratings. *Amazon*'s customers trying to rank two items of interest for themselves might be familiar with the problem: Should an item with a full five-star rating by five raters be ranked higher than another item with a five-star rating of twenty raters? Another issue to be dealt with is data sparsity when explicit forms of ratings are used.

In the following two methods of aggregating individual ratings to a collective rating are introduced. It should be noted however, that rating and aggregation of ratings highly depends on the concrete applications and thus is a science of its own. Finding aggregation methods that perfectly suit the needs of *Backstage* requires thorough research and long-term evaluations, which is outside the scope of this thesis.

6.3.2.1 Bayesian Average

The Bayesian Average aggregation method is based on a weighted average that weighs the individual ratings against all ratings an item received. The more ratings an item receives the more does the collective rating deviate from a default average number of positive ratings (in the formula below, denoted as $\overline{n}^+$). The Bayesian Average aggregation $r_b(\overline{n}, \overline{n}^+, x)$ of an item x, e.g., a backchannel post, is calculated as follows (Marmolowski, 2008):

$$r_b(\overline{n}, \overline{n}^+, x) = \frac{\overline{n} \times \overline{n}^+ + n(x) \times n^+(x)}{\overline{n} + n(x)}$$

In the formula above $\overline{n}$ denotes the average number of ratings given by users, $\overline{n}^+$ the average ratings given by users, $n(x)$ the number of ratings of an item x, and $n^+(x)$ the number of positive ratings of x. Positive and negative ratings do not cancel each other out. Instead, the negative ratings weaken the influence of the positive ratings. If the total number of ratings of an item $n(x)$ is much smaller than the average number of ratings $\overline{n}$, the collective rating of item x is dominated by the average rating $\overline{n}^+$. Conversely, if the number of ratings for x is greater than the average positive number of ratings given, the collective rating of x is dominated by the users' positive ratings of x. Thus, the rating scheme is biased in that it favors the average appraisal of the collective over that of the few.

6.3.2.2 Binomial Proportion Confidence Intervals

For a dichotomous rating scheme, i.e., a scheme providing two kinds of ratings, ratings can be aggregated to collective ratings by determining the binomial proportion confidence interval. Evan Miller proposes the use of the so-called Wilson's confidence interval as an aggregation method.[1] According to Miller, the method is also been applied in *Reddit*,[2] *Yelp*,[3] and *Digg*.[4]

The following paragraphs briefly describe the ideas behind and derivation of Wilson's confidence interval (see Equation 6.4, p. 96). The reason why Miller suggests to use Wilson's confidence interval and not the simpler Wald's confidence interval is given below.

If a dichotomous rating scheme of approvals and rejections is assumed, the number of approvals an item i receives can be stochastically modeled as a binomially distributed random variable $X_i \sim \text{Bin}(k; n, p_i)$,[5] where k is the number of positive ratings an item i receives, n is the number of observations and p_i is the probability of i receiving a positive rating in each trial. If known one could use p_i directly as the aggregated rating of item i.

However, the true value of p_i is usually not known. Thus, the common approach is to estimate p_i by determining the observed average number of approvals $\hat{p}_i$ of an item i. For example, if a backchannel post has been approved three out of five times, one would set $\hat{p}_i = 3/5 = 0.6$. Unfortunately, as with all point estimates, few is known about how well $\hat{p}_i$ estimates p_i. Instead, the proposed method constructs a confidence interval $C_i(\alpha, p_i)$ that, in $100(1-\alpha)\%$ of the times the interval is constructed, the interval contains p_i. It should be noted that $1 - \alpha$ is a specification that a confidence interval is supposed to satisfy. If it does, the actual probability of an interval containing p_i, the so-called coverage probability, equals the specified confidence level.

The aggregation method now proposes to take as the collective rating of an item i the lower bound of the constructed confidence interval. The confidence interval is constructed by approximating the binomial distribution that models an item i's probability of being approved with a normal distribution as follows:

1. Approximate the binomial distribution of X_i with a normal distribution.
2. Determine by the thus obtained distribution an approximating normal distribution for $\hat{p}_i$.
3. Transform the values of $\hat{p}_i$ to z-scores that follow a standard normal distribution $\mathcal{N}(0, 1)$.
4. Construct $C_i(\alpha, p_i)$ by determining the critical values $\pm z_{1-\alpha/2}$ (z-scores) as the margins of the confidence interval.

Especially when the sample size n is sufficiently large, a discrete binomial distribution can be approximated by a continuous normal distribution.[6] Recall that the mean of a binomial distribution $\text{Bin}(k; n, p)$ is $\mu_{\text{Bin}} = np$ and the standard deviation is $\sigma_{\text{Bin}} =$

---

[1] Evan Miller, *How not to Sort by Average Rating*, Blog post dated 6 February 2009, http://bit.ly/1dvckNi (last visited on March 2nd, 2015).

[2] http://www.reddit.com/

[3] http://www.yelp.com

[4] http://digg.com/

[5] Recall that a binomial distribution is used to model the number of successes in a sequence of experiments with success or failure as possible outcomes (see Wikipedia, http://bit.ly/1c80KwB, last visited on March 2nd, 2015).

[6] An approximation of a binomial distribution with a normal distribution is reasonable if $n\hat{p}_i > 5$ and $n(1 - \hat{p}_i) > 5$ (Sachs and Hedderich, 2009).

$\sqrt{np(1-p)}$. Thus the binomial distribution may be approximated by the normal distribution $\mathcal{N}(\mu_{\text{Bin}}, \sigma_{\text{Bin}})$. Now, $\hat{p}_i = \frac{X_i}{n}$, since X_i can be approximated by a normal distribution:

$$\frac{X_i}{n} \sim \mathcal{N}\left(\mu_{\frac{X_i}{n}}, \sigma_{\frac{X_i}{n}}\right) \text{ where } \mu_{\frac{X_i}{n}} = \frac{\mu_{X_i}}{n} = p \text{ and } \sigma_{\frac{X_i}{n}} = \frac{\sigma_{X_i}}{n^2} = \sqrt{\frac{p_i(1-p_i)}{n}}$$

In the next step, by using $\mu_{\frac{X_i}{n}}$ and $\sigma_{\frac{X_i}{n}}$ the values of $\frac{X_i}{n}$ are transformed to z-scores, which approximately are distributed according to the standard normal distribution $\mathcal{N}(0,1)$. A transformation to z-scores makes the rating scores of two items comparable with each other. For a random variable $V \sim \mathcal{N}(\mu, \sigma)$ the transformation is defined as:

$$z = \frac{V - \mu}{\sigma}$$

Accordingly, $\frac{X_i}{n}$ can be transformed as follows:

$$z_{\frac{X_i}{n}} = \frac{\frac{X_i}{n} - \mu_{\frac{X_i}{n}}}{\sigma_{\frac{X_i}{n}}} = \frac{\hat{p}_i - p_i}{\sqrt{\frac{p_i(1-p_i)}{n}}}$$

To determine the critical values of the confidence interval, consider the following set-up:

$$\frac{\hat{p}_i - p_i}{\sqrt{\frac{p_i(1-p_i)}{n}}} = \pm z_{1-a/2} \tag{6.1}$$

which can be reformulated as

$$p_i = \hat{p}_i \pm z_{1-a/2}\sqrt{\frac{p_i(1-p_i)}{n}} \tag{6.2}$$

As the simplest confidence interval approximation, Wald's confidence interval suggests to replace all occurrences of p_i on the right hand side of Equation 6.2 by the estimate $\hat{p}_i$:

$$C_i(\alpha, p_i) = \hat{p}_i \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n}} \tag{6.3}$$

However, Wald's approximation has been shown to perform quite badly, which is why the use of Wald's confidence interval is strongly discouraged (Brown, Cai, and DasGupta, 2001): The coverage probability strongly fluctuates independently from the sample size n and the actual coverage probability of Wald's confidence interval has been found to be notably smaller than the specified confidence level, meaning that the error made by assuming that the confidence interval includes p_i is notably greater than specified by the confidence level.

The approach suggested by Wilson (1927) is to solve Equation 6.1 for p_i, resulting in Wilson's confidence interval:

$$C_i(\alpha, p_i) = \frac{1}{1 + \frac{z_{1-\alpha/2}^2}{n}}\left(\hat{p}_i + \frac{z_{1-\alpha/2}^2}{2n} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}_i(1-\hat{p}_i)}{n} + \frac{z_{1-\alpha/2}^2}{4n^2}}\right)$$

Brown, Cai, and DasGupta (2001) verify that Wilson's confidence interval performs much better than Wald's confidence interval in that the coverage probabilities are much closer

to the specified confidence level, even for very small sample sizes. Similar derivations of Wilson's confidence interval can be found in, e.g., Sachs and Hedderich (2009), and Wallis (2013).

In his blog post, Evan Miller[7] now proposes to use as the aggregated rating of an item i the lower bound of the Wilson's confidence interval. That is, the aggregation of individual ratings of an item i, $r_w(i)$ is defined as follows:

$$r_w(i) = \frac{1}{1 + \frac{z_{1-\alpha/2}^2}{n}} \left(\hat{p}_i + \frac{z_{1-\alpha/2}^2}{2n} - z_{1-\alpha/2} \sqrt{\frac{\hat{p}_i(1 - \hat{p}_i)}{n} + \frac{z_{1-\alpha/2}^2}{4n^2}} \right) \tag{6.4}$$

6.3.2.3 Comparison of Bayesian Average and Miller's Method

The following table reflects the rankings of items according to the two presented aggregation methods, Bayesian Average and Wilson's confidence interval.

Item	Approvals	Rejections	r_b	Rank by r_b	r_w	Rank by r_w
1	0	0	0.60	**16**	0.00	**13**
2	0	1	0.60	**15**	−1.00	**14**
3	0	2	0.60	**14**	−2.00	**15**
4	0	3	0.60	**13**	−3.00	**16**
5	1	0	0.73	**12**	0.21	**6**
6	1	1	1.07	**11**	0.09	**10**
7	1	2	1.40	**9**	0.06	**11**
8	1	3	1.73	**8**	0.05	**12**
9	2	0	1.30	**10**	0.34	**2**
10	2	1	1.80	**7**	0.21	**5**
11	2	2	2.30	**5**	0.15	**8**
12	2	3	2.80	**3**	0.12	**9**
13	3	0	2.04	**6**	0.44	**(1)**
14	3	1	2.64	**4**	0.30	**3**
15	3	2	3.24	**2**	0.23	**4**
16	3	3	3.84	**(1)**	0.19	**7**

Table 6.1: Comparison of Bayesian Average (denoted as r_b) and Wilson's confidence interval (denoted as r_w) as methods for aggregating individual ratings. The following parameters were chosen: $\overline{n} = 2$, $\overline{n}^+ = 0.6$, $\alpha = 0.05$, i.e. $z_{1-\alpha/2} = 1.96$.

Since explicit rating needs to perform well also when only few rating data is available, the comparison considers ratings of not more than three positive and negative ratings. Note that the aggregated rating r_w used in the comparison is a slight modification of Equation 6.4. The original equation yields zero when the number of positive ratings is zero. In the comparison the aggregation method is adjusted such that if the number of positive ratings is zero the method yields the negative number of negative ratings.

Table 6.1 suggests that the Bayesian Average favors the total number of ratings over the agreement of the raters. Using Bayesian Average the Item 16 is ranked above

---

[7]See footnote 1 on page 94.

Item 13, although Item 16's ratings reflect discordance among the raters. In contrast, the aggregation method based on Wilson's confidence interval considers to a greater extent agreement of the raters, at least when the total number of ratings for an item is sufficiently large. For example, an item rated positively three times is ranked above an item that received three positive and three negative ratings and thus twice the total number of ratings. An interesting question is whether or not these tentative observations can be theoretically confirmed.

6.4   Ranking Messages in Backstage

This thesis describes two methods for aggregating individual ratings to collective ratings. The collective ratings can be used to sort backchannel posts so as to obtain a ranking of posts. Ranking backchannel posts is considered as a support for the lecturer in keeping track of the relevant backchannel communication. In the current implementation the lecturer can switch between the conventional backchannel view and a ranking of backchannel posts. The ranking is updated only on explicit request by the lecturer in order to keep the lecturer's user interface calm and to leave the decision when to update the ranking to the lecturer.

   The two aggregation methods presented in Section 6.3.2 were implemented in *Backstage* but only used in an incidental manner. A final assessment of the two methods is thus not possible at this stage. In the few practical uses, however the parameters $\overline{n}$ and $\overline{n}^+$ of the Bayesian Average method varied drastically due to a low rating activity of the audience.[8] Thus the Bayesian Average yielded quite counterintuitive and fluctuating rankings. The aggregation method based on Wilson's confidence interval, in contrast, seemed to be more consistent with expectations.

   Which of the methods is preferable from a practical viewpoint likely depends on the context of use (e.g., the preferences of the students and lecturers using the rankings). Thus, besides theoretical considerations, these methods also need to be extensively evaluated in practice. Possibly, the most reasonable approach therefore is to let the backchannel participants decide which aggregation method they want to use. *Backstage* would then resort to collecting statistics about which of the methods are preferred and considered most useful. This unobtrusive evaluation should be accompanied by interviews and surveys to also collect subjective impressions about the methods' appropriateness.

6.5   Aging of Rankings

Relevance of backchannel posts for the lecture is usually of a temporal nature. As the lecture proceeds, topics might change and some backchannel posts might become obsolete with respect to the progress of the lecture (while staying relevant and available for discussions and exchange after the lecture). Thus, ratings and rankings, for that matter, should depend on time. Therefore, some kind of aging is desirable. That is, the importance of backchannel posts should gradually degrade over time. With aging, attention during the lecture is directed to recent and active backchannel posts. Though, determining age on the basis of the physical time does not seem to be reasonable. Lectures usually vary in progress. For example, introductory slides might be presented at much a higher pace than a difficult mathematical proof. Aging should rather depend on a lecture-specific measure of time such as the activity on the backchannel. The

---

[8]When the overall rating activity is low additional ratings have great impact on the parameters $\overline{n}$ and $\overline{n}^+$.

approaches to aging of relevance found in the literature seem to be too involved and difficult to handle in the context of a backchannel for large class lectures.

6.5.1 Considering Time for Aging

Aging is typically a negative process of losing influence as time goes by. Therefore, aging is naturally expressed as some kind of weight decreasing over time and expressing a remaining relevance. The older an item is, the lower its influence on the overall score.

One solution is based on the half-life parameter as known from the modeling of nuclear decay processes. Therefore, a time-dependent monotonic decreasing function $f(t)$ is introduced (Ding and Li, 2005), for example the exponential or logistic function. The authors define the time function as $f(t) = e^{-\lambda t}$, where λ is the decay rate $1/T_0$. This approach depends on the setting of T_0, which specifies how long it takes to reduce the weight by half. The lower T_0, the faster the decay of the weight and the lower the influence.

As an approach to handle "freshness" of items on social tagging sites, Huo and Tsotras (2010) divide the time-line in discrete and equidistant time intervals. Aging is introduced by the function $f_m(s) = a^{m-s}$, where $a \in [0;1]$ is a decay factor. The parameter m counts the number of all time slices until the current time and $s \in [0;m]$ is an index variable that ranges over all time slices. For example, for the current time slice $m = s$. The fresher a tagging the smaller is the exponent, and the greater $f_m(s)$. Thus, fresher items have a bigger influence.

While the aforementioned approaches are comprehensive, the way items are supposed to age on the *backchan.nl*-platform[9] (Harry, Gutierrez, et al., 2008) is unclear. The ranking method of this system comprises two components, the `ageFactor` and the `voteFactor`, which are combined by multiplication. It seems obvious at first sight that the aging here is once again realized by some kind of weight. However, it rather seems that `voteFactor` and `ageFactor` are working against one another. The inconsistency results from the fact that Harry, Gutierrez, et al. (2008) defined the `ageFactor` such that over time the factor increases and, because of how the `ageFactor` and `voteFactor` are combined, the relevance of a post increases as time goes by.

6.5.2 Discerning Actuality in the Ranking of Posts

To better reflect the progress of a lecture the backchannel activity during the lecture can be used to promote aging. The logical time on the backchannel advances after each n-th action of a user on *Backstage*. Both the number n and the specification of what is considered as an action relevant for aging is defined by the lecturer. Actions may comprise sending or rating backchannel posts. Both the number of actions after which the time advances and the specification of activity on *Backstage* are very intuitive parameters that can easily be handled by the lecturer, even during a lecture.

A first idea for measuring the age of backchannel posts would be to calculate the difference between the current time and the time of creation. However, this solution is inappropriate, since backchannel posts that are regularly rated, i.e. active backchannel posts, would age at the same pace as backchannel posts which are disregarded by the audience. On *Backstage*, active backchannel posts ought to age at a lower pace than inactive backchannel posts. Thus, it is reasonable to also consider the age of a backchannel post's ratings. Hence, aging depends on the attention a backchannel post receives: it is promoted when the focus by the audience of a backchannel post

---

[9] http://backchan.nl

recedes. A naive approach to determining age might be to calculate the difference of the current time and the time of the most recent rating a backchannel post received. This is problematic, though. For example, imagine that many students have rated a post a long time ago, i.e. the backchannel post is actually obsolete, but one student revives the backchannel post by rating, and thus the backchannel post would suddenly, and inexplicably, rejuvenate.

The arithmetic mean of time differences over all ratings would solve this issue. However, it is very sensitive to outliers. Many ratings at the same time would be needed to assure that the age of this backchannel post can be considered robust. These difficulties can be mitigated by using the median as the average age of a backchannel post.

Thus, to determine the age of a backchannel post, the median from the ratings' ages and from the creation time of the corresponding backchannel post is determined. Also considering the time of creation is necessary in the case that a backchannel post has not received any ratings at the beginning. Otherwise, the backchannel post would not be considered by aging.

6.5.3 Aging in Backstage

After each n actions on Backstage aging is promoted and the ranking is updated. This thesis therefore propose the procedure given as pseudo-code in Listing 6.1.

Listing 6.1: AgingRank – Ranking with Aging.

```
1    input k: the number of backchannel posts that constitute the ranking
2    input n: the number of actions
3    if clockTick(n) then
4    begin
5      candidatePosts ← getCandidatePosts()
6      (* promote aging *)
7      for p ∈ candidatePosts
8      begin
9        setAge(p, calculateMedianAge(p))
10     end
11     rankingByAge ← sortDescendingByAge(candidatePosts)
12     rankingByRating ← sortDescendingByRating(candidatePosts)
13     (* assume lists to be 1-indexed *)
14     for i ← 1 to maxIndex(rankingByAge)
15     begin
16       post ← getElement(indexAge, rankingByAge)
17       (* get the index of the post in the ranking by rating *)
18       indexRating ← getIndex(post, rankingByRating)
19       setScore(post, indexAge * indexRating)
20     end
21     (* sort the candidates by just updated score values *)
22     relevantPosts ← sortDescendingByScore(candidatePosts)
23     resolved ← resolveConflicts(relevantPosts)
24     result ← firstElements(k, resolved)
25     setRanking(result)
26   end
```

As can be seen in the given procedure the rating score is obtained by multiplying the positions of a post in the two rankings built upon age and ratings. Since it is possible that two posts may be assigned the same rating they may share the same position in the respective ranking. The final top-k ranking is then computed by sorting the list of relevant backchannel posts according to the backchannel posts' scores. However, the given procedure may result in conflicts. For example, two backchannel posts, say, m_1 with indexRating = 2 and indexAge = 3, and m_2 with the positions conversed, that is indexRating = 3 and indexAge = 2 would receive the same score 6. The two backchannel posts m_1 and m_2 would be assigned the same position in the final ranking. Thus, conflict resolution is necessary.

A simple but eligible approach to conflict resolution is proposed: let the lecturer decide which of the conflicting backchannel posts should be given a higher priority. Therefore, the lecturer specifies in her profile, whether she favors a conservative ranking, i.e. older backchannel posts stay in the ranking, or a progressive ranking in which older backchannel posts are replaced by newer ones whenever possible. In case of further remaining conflicts eventually a strict order is established by resorting to the physical age, since the conflicting backchannel posts can then be considered equal in terms of relevance and logical age.

To determine the follow-up ranking it is not necessary to consider the entire backchannel communication. It rather suffices to determine a set of candidates, the number of which depends on the number of actions n by which aging is promoted. Certainly, the backchannel posts listed in the current ranking are also candidates for the follow-up ranking. However, other backchannel posts may be candidates as well. For this purpose, consider the example time-line in Figure 6.2.
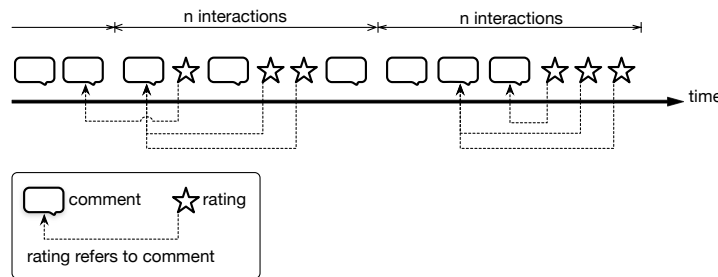

Figure 6.2: Example Time-line of actions. The speech bubbles illustrate the points in time at which backchannel posts are sent, the stars illustrate the points in time at which posts are rated. The dotted arrows connect the ratings with the rated posts.

Between two ticks of the logical clock, n actions are carried out by the users. These actions may comprise the creation of $x \le n$ new posts and $y = n - x$ ratings for existing posts. The ratings may refer to up to y posts created during the recent or some earlier time span. All these posts have recently been in the focus of the audience. Thus, besides the currently ranked posts in the top-k ranking, both the newly created and the newly rated posts are also candidates for the follow-up ranking. Reckoned up, the set of candidate posts comprises not more than $k + n$ posts. For example, in Figure 6.2, the number of actions that make the logical clock advance is $n = 6$. If the ranking of posts comprises $k = 10$ posts, 16 posts belong to the set of candidates and need to be considered in the update of the ranking.

6.6 Discerning Actuality in Twitter-based Tools

Although it was *Backstage* for which temporal ranking by aging was conceived, the aging approach is likely to be of interest for other microblogging platforms as well. To employ the approach it is sufficient to provide for means to determine relevance ratings, to measure activity, and to set the strategy for updating the rankings. This section illustrates possible applications in both e-learning and non-e-learning fields.

Twitter provides a rich Application Programming Interface (API) upon which custom

microblogging applications can be built. *TwitterWall*[10] allows the retrieval and display of multiple post streams filtered by specified hashtags (Ebner, 2011). Furthermore it extends Twitter in that it provides rating of tweets. To extend *TwitterWall* with aging, the rating scheme that is already integrated can be used. Activity can be measured by the number of tweets containing certain hashtags. As for each hashtag *TwitterWall* displays a separate tweet stream, it might also be interesting to provide rankings for each of those streams that underly distinct aging.

Also, discerning actuality in tweet rankings directly on *Twitter* can be accomplished in much the same way as is proposed for *Backstage*. Note that *Twitter* does not provide rating of tweets. However, rating of a tweet can be mimicked by counting the number of users retweeting the tweet. That is, a tweet that is frequently retweeted is heavily focused on by users and may thus be considered relevant. On *Twitter*, activity can be measured by the number of tweets containing certain hashtags and by the number of retweets of those tweets. Obtaining a timely ranking of tweets may provide interesting insights into trends in social news broadcast on *Twitter*.

Another quite interesting field of application might be stock microblogging, e.g., *TweetTrader*[11] (Sprenger, 2011). Among other things, users of *TweetTrader* estimate in tweets the performance of stock quotations. Using special processable syntax, those tweets are evaluated and aggregated to determine the collective estimation of near-future stock developments. Discerning actuality in a ranking of those estimations might be of great interest for stock microblogging. Ratings in this case might be based on the content of the tweets, i.e. the users' assessments of the stock development. The activity may be specified by the number of tweets sent, for example. A progressive update strategy is likely to be preferred for a ranking in order to always be aware of the most recent estimations. Also, timely ranking of stock quotations might yield interesting outcomes in the analysis of trends.

---

[10]http://twitterwall.tugraz.at
[11]http://tweettrader.net

7

REPUTATION

Online communities substantially rely on cooperation among their members.[1] Especially in large and anonymous communities such as online communities, it is not immediately clear which members are cooperative. Reputation is often used to incentivize desired behavior and to collectively regulate the interactions. Reputation thus provides information that allows members to estimate the prospect of cooperation.

In this chapter the algorithmic principle of a reputation system for *Backstage* is described. Students may gain reputation by contributing comments on the backchannel the students consider relevant for the lecture. However, the reputation system is limited to praising, i.e. to positive ratings, for reasons discussed later. Although this reputation system has been designed for *Backstage*, it is likely to be useful for other online platforms as well.

This thesis follows the conceptualization of reputation and its distinction from *status* as made by Vassileva (2012). Members may gain status which, unlike reputation, is not based on the aggregated opinions of other members. For example, a member of a Q&A portal may gain the status of a "power contributor" when the number of submissions exceeds a certain threshold. Whether or not that member's contributions can be expected to be useful, on the other hand, is a question of the member's reputation among other community members. However, status and reputation may coexist as they complement each other. While status can cater for activity of users, which might be of special importance in the development of a community, reputation encourages cooperation and thus contributes to quality of interaction.

7.1 Reputation as a Social Incentive Instrument

In social media information about cooperative behavior is usually provided by online reputation of members (or reputation for short). Reputation is an aggregate value,

---

[1]In this chapter, the term cooperation is used as a reference for interactions that are in a certain sense central for the members of a community. In an online market, for example, cooperation between buyers and sellers refers to agreements of purchases. On a Q&A portal cooperation refers to the submission of questions and answers. In this sense the term cooperation is meant to cover also competitive or even adverse (but not malevolent) interactions, for example in online games.

publicly available to all members of an online community, that summarizes for a member the opinions of others who interacted with the member about how well the member has cooperated (Resnick et al., 2000; Vassileva, 2012). Thus, reputation aggregates and possibly quantifies social recognition. The members' opinions are often expressed by some kind of numeric ratings.

Social recognition can be an extraordinarily effective incentive for cooperation. A rather striking example thereof is constituted by Harvard economist Nava Ashraf. She investigated whether poor people in developmental countries are incited to a greater extent by money or by social reputation.[2] One might suspect that, especially in poor societies, money is a stronger incentive than social recognition. However, Ashraf found that the participants who were promised social recognition significantly outperformed the participants who received additional income.[3]

Reputation enables members of a community to become aware of cooperative behavior. However, this makes reputation also a social incentive. Termed as the *shadow of the future*, the prospective benefit of a potential cooperation in the future is assessed on the basis of a member's past behavior, which is reflected in the member's reputation. To maintain a high chance of drawing benefits in the future, a member is thus incited to cooperate in the present (Resnick et al., 2000; Surowiecki, 2005). In this way, reputation realizes a collective steering of behavior in the sense of a *libertarian paternalism* (Thaler and Sunstein, 2003): The members are basically free in their choices (*libertarianism*); however, their decisions are influenced and steered towards cooperation in order to promote welfare (*paternalism*). However, to maintain an effective reputation mechanism, sufficient information needs to be provided. On the one hand, this can be achieved by having other community members taking explicit action, e.g. providing ratings, which however requires incentivization as well (Jøsang, Ismail, and Boyd, 2007). For example, truly reciprocal interactions can increase the members' willingness to provide ratings (Dellarocas, Fan, and Wood, 2004). On the other hand, reputation can be derived implicitly by observing and evaluating the members' behavior on the platform (Jensen, Davis, and Farnham, 2002).

7.2 Reputation as a Means to Establish Trust

Reputation is also important as a means to reduce uncertainty and to establish trust. In online commercial transactions, the involved parties rarely know each other or have already engaged in previous transactions. As Jøsang, Ismail, and Boyd (2007, p. 618) put it, "The consumer generally has no opportunity to see and try products, i.e. to 'squeeze the oranges' before he buys". For this reason, online reputation and trust are often investigated in the field of e-commerce (e.g., Pavlou and Ba, 2000; Bolton, Katok, and Ockenfels, 2002; Zhou, Dresner, and Windle, 2008). Trust and reputation systems

---

[2]Ashraf's experiment has been reported on in Dr Christoph Kulick's article on poverty economics, "Viel hilft viel. Oder nicht? [A lot helps a lot. Doesn't it?]", published in the German edition of the GEO magazine in May 2012, pp. 98 – 112.

[3]In the experiment hairdressers in South Africa were asked to distribute as many condoms among their female customers as possible, in order to support sexual autonomy of females and to prevent spread of HIV. For the study, all hairdressers received special training in prevention. During the study, several groups of hairdressers received different amounts of money for each distributed condom, another group was promised a public tribute at the end of the study (the reputation-group). About twice the number of condoms were distributed by the hairdressers of the reputation-group compared to the hairdressers of the other groups. Besides the prospect of a public tribute and bragging, the hairdressers were motivated by their service to society and the outlook of being appreciated as prevention counsels. One of the researchers' conclusions that has led to great controversy is that poverty alleviation measures that are centered on social reputation can sometimes be more effective than the mere spending of money.

are so-called *soft security mechanisms*, i.e. social control mechanisms that intend to provide protection against malicious parties. In their summary, Jøsang, Ismail, and Boyd (2007) distinguish between two notions of trust. *Reliability trust* denotes the subjective probability an individual *A* holds that another individual *B* is going to behave as expected by *A* whose welfare depends on *B*. *Decision trust* denotes the extent to which *A* is willing to depend on *B* with a feeling of relative security, even though it is possible that *B* betrays *A*'s trust. Trust systems deal with the measurement of trust among individuals. However, a precise definition of trust systems seems to be difficult. One focus of trust systems is the derivation of transitive *trust paths*.

Reputation plays an essential role when reciprocity is the social norm of collective action in a community. Reciprocity describes a member's positive response to another member's positive action (Falk and Fischbacher, 2006). Since the response may be given at a later time, the connection between one's action and a resulting benefit may not always be clear. That is, a member's action may appear to be altruistic, in which case reciprocity can be characterized as a *reciprocal altruism*. Thus, in a reciprocal community a member may behave cooperatively, expecting the future benefit that others respond to her in kind once they become aware of her initial cooperativeness. A reciprocal action usually involves the following steps: identifying possible interactants, assessing the interactants' likelihoods of being cooperative, deciding to cooperate if cooperative interactants are available (or deciding to refuse cooperation), and punishing members who betray trust (Ostrom, 1998). When reciprocity (or reciprocal altruism) is used by many as the strategy of collective action, there is an incentive to gain reputation.[4]

7.3 Centrality and Prestige

Sociometry, a mathematical branch of sociology, deals with the modeling of social systems as networks consisting of so-called *social actors* and the relationships among actors. Such systems are modeled as graphs in which social actors are represented as nodes. If the relationships under study are non-reciprocal the graph is a directed graph. In case of reciprocal relationships the graph is undirected. One of the basic issues is that of *prominence* of social actors on the basis of the relationships among the actors. Prominence based on reciprocal relationships is also referred to as *Centrality*. For non-reciprocal relationships prominence is also referred to as *Prestige* (Wasserman and Faust, 1994). Conventional prominence measures are based on the immediate ties among actors, e.g. the in-degree and out-degree, as well as closeness and betweenness centrality (Wasserman and Faust, 1994). These measures are simple in that they are solely based on the relationships among actors, leaving the prominence of the "choosing" actors unconsidered. In contrast, so-called *Status Indices* or *Prestige Ranks* take into account the importance of the "choosers" as well (Wasserman and Faust, 1994). These methods are of particular interest for *Backstage*, since it seems reasonable to let students who valuably contribute to the backchannel have greater influence in social control than less active students. However, the mathematical foundations are more sophisticated, since cyclic relationships generally lead to infinite regress.[5] To obtain well-defined importance metrics, these models are required to possess some kind of convergence.

---

[4]Not all online communities use reciprocity or reciprocal altruism as a social norm of collective action. For example, in the Twitter community, following a user, and hence, re-tweeting status updates, is in general not reciprocated (Kwak et al., 2010). Perhaps, the non-reciprocity in the community is one of the reasons why reputation is not provided on Twitter.

[5]For example, consider a social system of three actors *A*, *B*, and *C* and the cyclic relationships $A \rightarrow B \rightarrow C \rightarrow A$. In this system, the determination of *C*'s importance eventually depends on the importance of *C*, leading to an infinite regress.

The proposal of Katz (1953) considers a social system as an undirected graph, whose adjacency matrix is a binary symmetric matrix (i.e., containing only 0's and 1's). The number of times an actor has chosen another actor is not accounted for. Katz proposes that status of actors is "radiating" in the entire network, while the intensity of status attenuates with the number of hops in the network. That is, the status of an actor is computed by the sum of all status that reaches the actor from arbitrary distances: $\mathbf{S} = \sum_{k=1}^{\infty} (\alpha \mathbf{C})^k$, where $0 \leq \alpha < 1$ is the attenuation factor and $\mathbf{C}$ the binary adjacency matrix (the "choice matrix"). If α is smaller than the reciprocal of the largest absolute value of eigenvalues of the choice matrix $\mathbf{C}$, the centrality measure is convergent (Bonacich, 2007). If α is equal to the reciprocal of the largest eigenvalue of $\mathbf{C}$, Katz's centrality is identical to an eigenvalue centrality. The status indices for the actors are obtained by computing the column sums of $\mathbf{S}$, given that the voting is recorded in $\mathbf{C}$ row-wise.

Bonacich (1987) reports on a class of centrality indices $c(\alpha, \beta)$ as a generalization of an eigenvector centrality he had proposed earlier. In the computation of an actor's centrality, the parameter β determines the influence of the other actors' centrality and, thus, the "locality" of centrality. For example, if $\beta = 0$, then the others' centrality is completely disregarded and the obtained centrality measure is proportional to the in-degree centrality index.[6] A positive value of β leads to Katz's proposal. If β approaches the reciprocal of the dominant eigenvalue, $c(\alpha, \beta)$ converges to the corresponding eigenvector of the relationship matrix. Bonacich explicitly allows for negative values for β and thus makes possible to align centrality and power (e.g. bargaining power in exchange networks), which in general are not equal.

Also, Bonacich and Lloyd (2004) presents an approach to eigenvector centrality on binary symmetric matrices that also allows to consider negative relations. Negative relationships are useful when power is to be measured instead of centrality. Power describes an actor's dependence (or independence) on some other actor in a network in which items (e.g. valued items or information) are exchanged (Cook et al., 1983). As it appears, power and centrality are different in that actors with great power do not necessarily have a high centrality score. In Bonacich's proposal, an individual's status reduces accordingly if the individual has a negative or a disassociate tie to another individual with high positive status. Furthermore, the authors use the sign of the obtained status score as a means to distinguish between two cliques in the network.

7.3.1 Eigenvector Centrality in Document Retrieval

Eigenvector centralities have become popular in their uses for the Web, foremost in form of the Google search engine. The eigenvectors are computed for websites on the basis of the Web's link structure. The number of links of a website to another is not accounted for. Thus, the adjacency matrix of the Web is binary. An entry of 1 indicates that a website links to another, an entry 0 indicates absence of such links.

7.3.1.1 HITS

Kleinberg (1999) introduces HITS that operates on a binary hyperlink matrix of web documents. HITS computes two indices for a website, the hub and the authority score. The hub score measures the popularity of a website as a referrer to popular websites measured by the authority score. In HITS, hub and authority are interdependent and computed in alternating fashion. The idea behind HITS is that a website is a good hub if it links to many popular websites and a website is popular if is is linked to by many good

---

[6]In case of $\alpha = 0$, the centrality index equals the in-degree index.

hubs. The two scores are determined on the basis of a (usually sparse) binary adjacency matrix $\mathbf{L}$ of a hyperlink graph. The computation of the hub vector $\mathbf{h}$ and the authority vector $\mathbf{a}$ are as follows:

$$\mathbf{h}^{(k)} = \mathbf{L}\mathbf{a}^{(k)} \quad \text{and} \quad \mathbf{a}^{(k)} = \mathbf{L}^T\mathbf{h}^{(k-1)}$$

which can be simplified to

$$\mathbf{h}^{(k)} = \mathbf{L}\mathbf{L}^T\mathbf{h}^{(k-1)} \quad \text{and} \quad \mathbf{a}^{(k)} = \mathbf{L}^T\mathbf{L}\mathbf{a}^{(k-1)}$$

The solutions correspond to the eigenvectors of the symmetric matrices $\mathbf{L}\mathbf{L}^T$ and $\mathbf{L}^T\mathbf{L}$ which can be shown exist and are both real. HITS is accompanied by several shortcomings. The original proposal requires that a neighborhood graph be constructed that matches a given query. However, a modified query-independent HITS can be found in (Langville and Meyer, 2012). Furthermore, HITS has been found to be vulnerable to spamming and topic drift. Besides these weaknesses, particularly for the use as a reputation measure, obtaining two different indices may be difficult to present and to interpret. HITS was the basic technology for the Temoa search engine[7] (Langville and Meyer, 2012) which has been incorporated by `ask.com`.[8]

7.3.1.2 PageRank

Around the same time of HITS, Sergey Brin and Larry Page have developed the *PageRank* algorithm (Brin and Page, 1998), which is the basis of the Google search engine. *PageRank* is based on the intuitive model of a random web surfer that, at each unit of time, follows the links on a page at the probability $1 - \alpha$ (i.e., clicking on a link embedded in a web page) and with a probability α jumps to an arbitrary page (i.e., entering a new url into the browser's address bar). The intuition behind the model is that of a random surfer that for some time follows links and after a while jumps to an arbitrary website by entering a new URL in the browser's address bar. The stochastic behavior of the random surfer is modeled as a Markov chain, the transition matrix of which is known as the Google Matrix:

$$\mathbf{G} = (1 - \alpha)\mathbf{H} + \alpha\mathbf{e}^T\mathbf{v}$$

The hyperlink matrix $\mathbf{H}$ is a stochastic matrix that is derived from the link structure of web pages, in which "dangling pages", i.e. pages not containing any links, are eliminated by adding uniformly weighted edges to all other web pages. The factor α is the *leap factor*, the probability that the random surfer jumps to another page. The probability of choosing a specific web page as the target of a random jump is given by the *personalization vector* $\mathbf{v}$. In the literature, α is usually assumed to be 0.15 and $\mathbf{v}$ to be the uniform distribution stating that the random surfer does not prefer any web page over another. The *PageRank* vector is characterized as the dominant eigenvector of $\mathbf{G}$. It can be shown that the dominant eigenvector of $\mathbf{G}$ always exists, since $\mathbf{G}$ is irreducible and the graph represented by $\mathbf{G}$ is strongly connected.[9]

Numerous extensions and adaptions of *PageRank* exist. For example, SALSA is a combination of HITS and *PageRank* in that it obtains hub and authority indices on the basis of a Markov Chain (Lempel and Moran, 2001; Langville and Meyer, 2012). Topic-Sensitive *PageRank* is an adaption of *PageRank* that determines a set of *PageRank*

---

[7]`http://www.teoma.com` now redirects to `http://ask.com`

[8]see Wikipedia, `http://bit.ly/1ceVBmG` (last visited on April 30th, 2015).

[9]It can be shown that for the Google Matrix, which is irreducible, meaning that it represents a strongly connected graph, a unique dominant eigenvector exists.

vectors for precompiled topics by specifying topic-specific leap factors (Haveliwala, 2002). PEST is another extension of *PageRank* for graph-structured data such as RDF that propagates term-weight between related data items in form of informed leap factors (Weiand et al., 2012).

7.3.2 Aggregation of Social Choice using Eigenvector Centralities

Variants of *PageRank* have also been developed as an aggregation of collective choice and assessment. Wang et al. (2013) present *ExpertRank* that is used to find expert users in Q&A portals. The approach is based on a combination of what the authors call expertise relevance, which is determined from text messages using information retrieval techniques such as TF-IDF, and a global authority score which is determined on some kind of "interaction graph" that reflects which user has communicated with whom, using *PageRank*. The *ExpertRank* is a convex combination[10] of the two rankings obtained through the expertise relevance and the global authority score. The paper reports on improved results compared to an eigenvector centrality using conventional *PageRank*.

In the field of Social Choice Theory, Altman and Tennenholtz (2005) investigate *PageRank* as an aggregation mechanism to form collective preferences out of the individual preferences of social agents. They consider the Web of documents as a set of agents and the links between the documents define the preferences of an agent: An agent prefers the pages it has links to and does not prefer any of the other pages. The authors present five intuitive aggregation axioms and showed that these are satisfied by *PageRank*. Furthermore, as the central piece of their paper, the authors show that *PageRank* is the only aggregation mechanism that satisfies the five simple axioms. As a consequence, any other aggregation mechanism satisfying the five axioms must coincide with *PageRank*.

Bry (2013) suggests and presents a Human Computation enabled eigenvector centrality measure for a systemic credit risk rating that intends to overcome the difficulties associated with the assessment of credit risks by financial institutions, which have become evident in the 2007 US subprime crisis and the ensued world-wide financial crisis. The fundamental idea of the approach is to let the debtors themselves assess their risk of defaulting. For this purpose, to incite truthful assessment, the deliberate construct of a grace period reward, GPR, is introduced. GPR is a fund maintained by all financial entities which allows debtors to defer repayment in case of payment difficulties. The entities' GPR charges are proportional to the expected risk and the expected duration of payment difficulties. The GPR construct is similar but not identical to an insurance to counteract moral hazard[11] which is problematic for the truthful assessment of one's own risk of defaulting. A graph is constructed where the nodes are financial entities (e.g., individuals, financial institutions) and the edges represent credit transactions among the entities. The edges are weighted by the aggregated individually assessed risks.[12] The graph is made strongly connected by feeding back to all entities the accumulated risk of central banks that are reflected, e.g. in key interest rates or the LIBOR and the EURIBOR.[13] On this graph an eigenvector centrality is computed so as to obtain global

---

[10] A convex combination is a linear combination where all coefficients are non-negative and sum to one. See Wikipedia, http://bit.ly/1AudDNn (last visited on April 30th, 2015).

[11] Moral hazard describes a person's increased readiness to take more risks because someone else has agreed on repaying for potential losses. For example, see Wikipedia, http://bit.ly/1Ba1hq4 (last visited on April 30th, 2015).

[12] In contrast to the aforementioned eigenvector centrality indices, which are defined on unweighted edges, Bry's index is defined on weighted edges.

[13] The LIBOR and the EURIBOR denote the interest rates on loans among banks.

risk assessment scores for all entities. The vision of the proposal is to obtain a more flexible risk assessment instrument that enables a better analysis and prediction of financial bubbles.

7.4 A Reputation System for Backstage

The conceptual model behind the reputation system can be described, and formalized, as a hypothetical "game" played by the *Backstage* users in which a badge is passed after each unit of time from student to student as follows. The current badge holder i passes the badge to a student j of her choice, possibly herself (in which case j is equal to i), who in her opinion most deserves a praise because of contributing relevant backchannel comments.[14] However, not only does receiving the badge express social recognition; it also gives the possibility of praising a student of one's choice. The empowerment for praising others can as well be seen as a social recognition. The more often a student holds the badge, the more does the student exerts influence on the praise (and, hence, on the reputation) of others. Thus, the probability of holding the badge is considered as a student's reputation.

The reader familiar with eigenvectors recognizes in the informal description so far the positive reinforcement typical of eigenvector centralities. Indeed, the reputation system described in this chapter is an eigenvector centrality index which, as one might expect, is well described as a stochastic process.

Definition 7.1 (Community of Students). *Define the community, that is the group of students belonging to a lecture, as the set $\mathcal{S} = \{1,\ldots,n\} \subset \mathbb{N}$ with $n > 1$.*[15]

Definition 7.2 (The Reputation Model as a Stochastic Process). *Let $t \in \mathbb{N}_0$ be the time units or rounds in the badge-passing model. The reputation model can be described as a collection of random variables $(B_t)_{t \in \mathbb{N}_0}$ ranging over $\mathcal{S}$, yielding the current badge holder at time t. Since the time t is discrete and the number of states (i.e., the community of students $\mathcal{S}$) is finite, the model of passing a badge corresponds to a so-called finite-state discrete-time stochastic process (Schickinger and Steger, 2002).*

In the model it is assumed that every student s holds a probability of praising for each student i, denoted as $p_{s \to i}$, expressing the chance of i receiving the badge from s.

Definition 7.3 (Praise Probability Distribution). *Define the praise probability distribution vector of a student $s \in \mathcal{S}$ as $\mathbf{p}_s = (p_{s \to 1}, \ldots, p_{s \to n})$. Let $p_{s \to i}$ denote the probability of a student s to praise student i. That is, the coefficients of the praise probability distribution vector denote the conditional probabilities in the badge-passing model: $p_{s \to i} = P(B_{t+1} = i \mid B_t = s)$, where B_t is the random variable yielding the current badge holder at time t. It is assumed that the conditional probabilities $p_{s \to i}$ of passing the badge are constant in the model.*[16]

---

[14]In the following, it is assumed that students' choices whom they pass the badge to are only based on their appreciations of the students' contributions on the backchannel and do not depend on friendships, enmities, or any other personal or social aspects. This assumption simplifies the model in that the probabilities of passing the badge of two students can then be considered as independent probabilities.

[15]In practice, $n \gg 1$. The use of a classroom communication system such as *Backstage* for quite a small, let alone a one-student audience, is usually not necessary. Issues of large classes can be resolved more easily in small groups without requiring a classroom communication system.

[16]When the praise probability distributions change, a new model can be set up and the reputation values can be computed anew. This allows to assume the praise probability distributions to be constant.

Although it is assumed that a student holds praise probabilities it is not possible to directly measure these probabilities. Instead, these probabilities are estimated from the ratings (see Chapter 6) given by the students on *Backstage*, which to a certain degree reflect the praise probabilities – at least under the assumption above that the latter are constant. More specifically, the praise probability distribution of a student s is estimated by the positive ratings the student s has given.[17]

Definition 7.4 (Backchannel Praise Probability Distribution). *Let $u_{s \to i}$ denote the number of positive ratings ("up votes") a student s has given to a student i. Since students s cannot rate their own comments, set $u_{s \to s} = 0$. Let u_s denote the total number of ratings the student s has given to the other students: $u_s = \sum_{k=1}^{n} u_{s \to k}$. Define the praise probability distribution vector $\mathbf{b}_s = (b_{s \to 1}, \ldots, b_{s \to n})$ by distinguishing between the following two cases:*
***Case 1.*** *The student s has rated at least one comment: $u_s > 0$*

$$b_{s \to i} = \begin{cases} \dfrac{u_{s \to i}}{u_s} & \text{if } s \neq i \\ 0 & \text{otherwise} \end{cases}$$

Case 2. *The student s has not rated any comment yet: $u_s = 0$*

$$b_{s \to i} = \begin{cases} \dfrac{1}{n-1} & \text{if } s \neq i \\ 0 & \text{otherwise} \end{cases}$$

In case 1, the influence of a student s is distributed among the rated students in proportions given by the relative frequencies of s's ratings. In the event that s has not yet rated any other student (case 2), indifference of s to all other students is assumed and made explicit by pretending one rating for each other student and, thus, distributing s's influence uniformly among the other students. This approach avoids the technical problem associated with *sinks*.
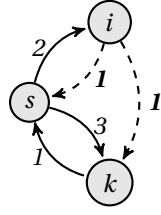
Example 7.5 (Backchannel Praise Probability Distribution). *Consider the following rating scenario between three students, s, i, and k. Accordingly, s has rated i twice and k three times, while k has rated s once.*
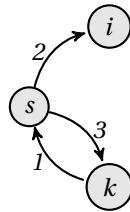
| ↗	s	i	k
s	0	2	3
i	0	0	0
k	1	0	0

Since the student i has not rated any other student yet, a pretended rating for each of the other students by i is added. Since s has not rated any other student, indifference of s is assumed. The indifference of s is made explicit in the rating graph by adding one rating for each of the students by s. This avoids technical issues associated with sinks in the sought for eigenvector centrality and is the same remedy for the the same problem as the handling of sinks ("dangling pages") with PageRank.

---

[17]In the following, "positively rating a student's backchannel comments" is simply referred to as "rating a student's comments", or simpler, "rating a student".

↗	s	i	k
s	0	2	3
i	**1**	0	**1**
k	1	0	0

The Backchannel Praise probability distribution is then determined by the relative rating frequencies. The Backchannel Praise probability distributions of each student correspond to the rows of the table below.
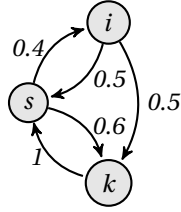
↗	s	i	k
s	0	0.4	0.6
i	0.5	0	0.5
k	1	0	0

Note the analogy of the approach with constructing from an adjacency matrix a stochastic hyperlink matrix for *PageRank*.

7.4.1 Considering the Rating Activity of a Student

The more a student engages in rating, i.e. the more ratings of a student are available in the determination of the backchannel praise probability distribution, the less is the backchannel praise probability distribution sensitive to outliers and the better is its performance as an estimator. The following two examples demonstrate undesired effects of low rating activities in the model.

Example 7.6 (Fluctuation of a Backchannel Praise Probability Distribution). *Let a community of three students, s, i and k be given. Assume that student s has rated exactly one comment, namely of the student i. Then $\mathbf{b_s} = (0, 1/1, 0)$. That is, s almost certainly passes the badge to i. At another point in time, assume that s rates for the second time, in this case a comment of the student k, hence $\mathbf{b_s} = (0, 1/2, 1/2)$. Now, the student s appears to be indifferent to the other students, i and k, meaning a significant drop-off for student i.*

It appears that a small data set with few ratings leads to backchannel praise probability distributions that are not good estimators for the praise probability distributions.

Example 7.7 (Differentiating between two Seemingly Equal Backchannel Praise Probability Distributions). *Let a community of three students, s, i and k be given. Let the student s's backchannel praise probability for the student i be $b_{s \to i} = 1/1 = 1$, and let the student k's backchannel praise probability for the student i be $b_{k \to i} = 20/20 = 1$. That is, both students s and k put all their influence on i. However, k appears to be more committed in her praising of i than s.*

As this example indicates, if a student exposes a high overall rating activity, backchannel praise probabilities intuitively seem to show a better picture compared to a student

with a less overall rating activity. However, because of the normalization to the relative frequencies this information is lost.

In order to allow to differentiate between two equal backchannel praise probability distributions it is necessary to consider the overall rating activity of each student and to attribute more significance to the backchannel praise probability distribution of the student with higher rating activity than of the student with lower rating activity.

Therefore, for each student s a rating activity coefficient $a_s \in (0;1)$ is introduced. The more a_s tends to 1 the more weight is attributed to s's ratings, and the more a_s tends to 0 the less weight is attributed to s's ratings. In the previous example, both relative frequencies $b_{s \to j}$ and $b_{k \to j}$ are equal; however, the student s has rated twenty times more often than the student j, hence, s's ratings should be given more significance than k's.

Postulate 7.8 (Rating Activity Coefficients). *The rating activity coefficient a_s of a student s measures that student s's overall rating activity on the backchannel. For the rating activity coefficient it holds that $0 < a_s < 1$.*

How exactly a_s can be defined and how it can be used in estimating the praise probability distribution on the basis of the backchannel praise probability distribution, will be discussed in the subsequent subsections.

7.4.2 Estimating the Rating Activity Coefficients

When statistics such as the rating activity of a student are determined on the basis of observations, care has to be taken so as to avoid over- and underestimation, which often occurs when only a small sample is available and maximum likelihood (ML) estimation is used. However, in the case of data sparsity smoothing can be used, the simplest of which is additive smoothing (see, for example, Chen and Goodman, 1996): A small sample is enlarged with artificial uniformly distributed observations before ML estimation is applied. More specifically, if one distinguishes between d different kinds of observations, e.g. different kinds of ratings, m artificial observations for each of the d kinds are added to the sample. That is, the sample is enlarged by $m \times d$ artificial data points. On the artificially enlarged sample, ML estimation can then be applied.

Definition 7.9 (Smoothed Estimator). *Let a sample of size N be given. Let $d > 1$ be the number of different kinds of possible observations. By adding $m > 0$ artificial observations for each of the possible kinds d, the smoothed estimator $\tilde{p}_{m,d}$ for making an observation y times can be obtained as follows.*

$$\tilde{p}_{m,d} = \frac{y + m}{N + m \times d} \tag{7.1}$$

Initially, when there are no observations available yet (i.e., $N = 0$ and $y = 0$), the smoothed estimator $\tilde{p}_{m,d}$ is $1/d$, which can be interpreted as the prior belief about the occurrence of an observation.

Example 7.10 (Smoothing the Estimator of a Rating Sample). *Assume a rating scheme, that allows approval, denoted as $\oplus$, as well as rejection, denoted as $\ominus$. Thus $d = 2$. Assume that the sample for the student i is the sequence $\Sigma = \langle \oplus \rangle$. According to ML estimation, the probability that the student i approves a backchannel comment is 1, and that i rejects a comment is 0. To avoid such extreme and rather uninformative values, the sample Σ is enlarged by uniformly adding $m = 2$ artificial data points for each of the possible ratings, $d = 2$. The sample becomes $\Sigma' = \langle \oplus, \underline{\oplus}, \underline{\oplus}, \underline{\ominus}, \underline{\ominus} \rangle$ (the underlined items are the artificial data*

points). According to the ML estimator applied to Σ', the probability that the student i approves a backchannel comment is now $p_\oplus = 3/5 = 0.6$, and that i rejects a comment is now $p_\ominus = 2/5 = 0.4$. Assert for observing $\oplus$ a prior belief of $1/d = 1/2$ (and a prior belief of $1/d = 1/2$ for observing $\ominus$, respectively).

With the considerations above, one possibility to define a smoothed rating activity coefficient a_i for a student i upon i's' rating activity (measured by u_i) is as follows:

Definition 7.11 (Smoothed Rating Activity Coefficients). *With smoothing parameters $m > 0$ and $d > 1$ define the smoothed activity coefficient a_s for a student s as*

$$a_s = \frac{u_s + m}{u_s + m \times d} \tag{7.2}$$

where u_s denotes the overall number of ratings given by s (as given in Definition 7.4).

The lower the rating activity of a student s, i.e. the lower u_s, the stronger is the activity coefficient pulled towards the prior belief $1/d$. In contrast, the greater u_s the weaker is the influence of the prior belief in the estimation.

The smoothed rating activity coefficients a_s as defined in Definition 7.4 satisfy the requirements given in Postulate 7.8, i.e. $0 < a_s < 1$. To see this, first consider $u_s = 0$. Then, a_s becomes $1/d$ (because $m > 0$) and with $d > 1$ it holds that $0 < a_s < 1$. If $u_s \neq 0$, the numerator is greater then zero and, because $m > 0$ and $d > 1$, the denominator is greater than the numerator.

There may be several ways to define (smoothed) rating activity coefficients, the definition above being but one possibility. Whether the smoothed rating activity coefficients as defined above make sense in practice needs to be thoroughly investigated in experiments.

Example 7.12 (Smoothing Activity Coefficients). *Let three students s, i, and k be given. Consider the two backchannel praise probability distributions to be $\mathbf{b_s} = (0, 1/1, 0)$ and $\mathbf{b_k} = (0, 20/20, 0)$. If the statistic is smoothed by adding $m = 1$ artificial data point and assume $d = 2$ different kinds of possible observations, the activity coefficient of student s is $a_s = 2/3 \approx 0.66$ and the activity coefficient of student k is $a_k = 21/22 \approx 0.95$. Comparing the two smoothed rating activity coefficients, the backchannel praise probability distribution of k is given way more significance than that of s.*

### 7.4.3 Estimating the Praise Probability Distribution

The activity coefficient specifies the significance that is given to a backchannel praise probability distribution in the estimation of the praise probability distribution. That is, to a degree as given by the activity coefficient, the backchannel praise probabilities are used as an estimate of a student's praise probabilities, and to the complementary degree it is assumed that the praise probabilities correspond to the uniform probability distribution.

Definition 7.13 (Estimated Praise Probability Distribution). *Let a student $s \in \mathcal{S}$ be given. Denote by $a_s \in (0;1)$ the rating activity coefficient for s. Define the estimated praise probability distribution of student s as the vector $\mathbf{r}_s = (r_{s \to 1}, \ldots, r_{s \to n})$ where*

$$r_{s \to j} = a_s \times b_{s \to j} + (1 - a_s) \times \frac{1}{n} \tag{7.3}$$

Note the analogy to the construction of the Google Matrix $\mathbf{G}$ from the stochastic hyperlink matrix $\mathbf{H}$ for *PageRank* (Section 7.3.1.2).

Example 7.14 (Example for the Estimated Praise Probability Distribution). *Consider the following rating scenario between three students, s, i, and k from Example 7.5.*
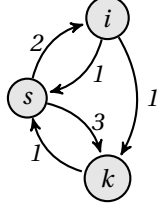
| ↗	s	i	k
s	0	2	3
i	1	0	1
k	1	0	0

According to the rating table above, $u_s = 5$, $u_i = 2$, and $u_k = 1$. For smoothing let $d = 2$ and $m = 1$. The smoothed rating activity coefficients are as follows.

Student	Rating Activity Coefficient
s	$a_s = \frac{u_s+1}{u_s+2} = \frac{6}{7} \approx 0.86$
i	$a_i = \frac{u_i+1}{u_i+2} = \frac{3}{4} = 0.75$
k	$a_k = \frac{u_k+1}{u_k+2} = \frac{2}{3} \approx 0.66$

The estimated praise probability distributions are as follows (BP-PD stands for backchannel praise probability distribution, and EP-PD stands for estimated praise probability distribution):

Stud.	BP-PD	EP-PD
s	$\mathbf{b_s} = (0, 2/5, 3/5)$	$\mathbf{r_s} = \left(a_s \times 0 + \frac{1-a_s}{3}, a_s \times \frac{2}{5} + \frac{1-a_s}{3}, a_s \times \frac{3}{5} + \frac{1-a_s}{3}\right) \approx (.05, .39, .56)$
i	$\mathbf{b_i} = (1/2, 0, 1/2)$	$\mathbf{r_i} = \left(a_i \times \frac{1}{2} + \frac{1-a_i}{3}, a_i \times 0 + \frac{1-a_i}{3}, a_i \times \frac{1}{2} + \frac{1-a_i}{3}\right) \approx (.46, .08, .46)$
k	$\mathbf{b_k} = (1, 0, 0)$	$\mathbf{r_k} = \left(a_k \times 1 + \frac{1-a_k}{3}, a_k \times 0 + \frac{1-a_k}{3}, a_k \times 0 + \frac{1-a_k}{3}\right) \approx (.78, .11, .11)$

The underlying graph, where s, i, and k are nodes and the elements of $\mathbf{r}_s$, $\mathbf{r}_i$, $\mathbf{r}_k$ represent the outgoing weighted edges of s, i, and k, respectively.
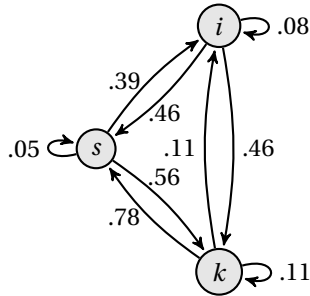
↗	s	i	k
s	0.05	0.39	0.56
i	0.46	0.08	0.46
k	0.78	0.11	0.11

The construction results in a strongly connected graph. That is, all nodes are directly accessible from each other. This property is required as becomes clear in the following sections, but not ensured by estimated backchannel praise probability distributions (see Example 7.5).

114

7.4.4 Determining the Probabilities of Holding the Badge

The estimated praise probabilities express the likelihood that the badge is passed by a student, if the student is the current badge holder. Recall that a student's probability of holding the badge in the badge-passing model is considered as that student's reputation. The probabilities of holding the badge can be determined on the basis of the estimated probability distribution by the law of total probabilities.

$$P(B_{t+1} = s) = \sum_{i=1}^{n} P(B_{t+1} = s \mid B_t = i) \times P(B_t = i) \tag{7.4}$$

According to Definition 7.3, the conditional probability $P(B_{t+1} = j \mid B_t = i)$ corresponds to the praise probabilities, which are estimated by the backchannel praise probabilities (cf. Definition 7.13). Thus, the probability of holding the badge in the next round is computed as follows:

$$P(B_{t+1} = s) = \sum_{i=1}^{n} r_{i \to s} \times P(B_t = i) \tag{7.5}$$

At the beginning the chance of a student s receiving the badge is given as $P(B_0 = s)$. This probability is specified by the initial probability distribution of holding the badge. For the subsequent units of time the probabilities of holding the badge is then computed using the Equation 7.5. Although the backchannel praise probabilities are constant in the model, the probabilities of holding the badge depend on the number of rounds played. Furthermore, the resulting probability distributions may be affected by the choice of the initial probability distribution. However, if the reputation model described so far corresponds to a so-called *Ergodic Markov Chain* the following holds (Schickinger and Steger, 2002):

- The process converges to a distinct stationary distribution as the number of time units approaches infinity.
- The stationary distribution is independent from the starting probability distribution. The initial probability distribution may thus be chosen arbitrarily.

If the stationary probabilities of holding the badge are used as a reputation index, the model becomes time-independent and, hence, the reputation index is well-defined.

7.4.4.1 The Reputation Model as a Markov Chain

Definition 7.15 (Finite-State Discrete-Time Markov Chain)**.** *A (finite-state, discrete-time) Markov chain over a state space $S = \{1, \dots, n\}$ is a collection of random variables $(X_t)_{t \in \mathbb{N}_0}$ ranging over S, and an initial probability distribution $\mathbf{p}_0 = (P(X_0 = i))_{i \in S}$. For the transitions between two states the Markov property holds:*

$$P(X_{t+1} = j \mid X_t = i_t, X_{t-1} = i_{t-1}, \dots, X_0 = i_0) = P(X_{t+1} = j \mid X_t = i_t).$$

That is, a transition only depends on the current state but not on how the current state has been reached. If the transition probabilities $p_{i \to j} = P(X_{t+1} = j \mid X_t = i)$ are time-independent the Markov chain is also called time-homogeneous. In this case it is possible to reformulate the Markov chain as a tuple $\langle \mathbf{P}, \mathbf{p}_0 \rangle$ where

- $\mathbf{P} = (p_{i \to j})$ *is the transition matrix, $p_{i \to j} = P(X_{t+1} = j \mid X_t = i)$*
- $\mathbf{p}_0$ *is the initial probability distribution.*

115

Since the transition matrix contains probability distributions, all entries have to be non-negative and each row sums to one. Matrices whose rows (columns) sum to one are called row (column) stochastic. Matrices that are both row and column stochastic are called doubly-stochastic.

Definition 7.16 (Estimated Praise Probability Matrix). *Let the backchannel praise probability distributions* $\mathbf{b}_s$ *be row vectors. Let* $\mathbb{1} = \mathrm{diag}(1,\dots,1)$ *denote the* $n \times n$ *identity matrix. Define the Uniform Praise Probability Matrix* $\mathbf{I} \in [0;1]^{n \times n}$ *as*

$$
\mathbf{I} = \begin{pmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \\ \vdots & \ddots & \vdots \\ \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix}
$$

and the Backchannel Praise Probability Matrix $\mathbf{B} \in (0;1)^{n \times n}$ *as*

$$
\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{pmatrix} = \begin{pmatrix} 0 & b_{1 \to 2} & \dots & b_{1 \to n} \\ b_{2 \to 1} & 0 & \dots & b_{2 \to n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n \to 1} & b_{n \to 2} & \dots & 0 \end{pmatrix}
$$

(for the entries $b_{i \to j}$, *see Definition 7.4), and the Rating Activity Coefficient Matrix* $\mathbf{A} \in (0;1)^{n \times n}$ *as*

$$
\mathbf{A} = \mathrm{diag}(a_1, a_2, \dots, a_n) = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix}
$$

Then, the Estimated Praise Probability Matrix $\mathbf{R} \in [0;1]^{n \times n}$ *can be formulated as*

$$
\mathbf{R} = \mathbf{A} \times \mathbf{B} + (\mathbb{1} - \mathbf{A}) \times \mathbf{I}
$$

Note the analogy with the construction of the Google Matrix $\mathbf{G}$ for *PageRank* (see Section 7.3.1.2). With the definition above it is possible to show that the reputation model corresponds to a time-homogeneous Markov chain.

Lemma 7.17. *The reputation model of* Backstage *corresponds to a time-homogeneous Markov chain* $\langle \mathbf{R}, \mathbf{p}_0 \rangle$ *where* $\mathbf{R}$ *is the estimated praise probability matrix given in Definition 7.16.*

Proof. By definition, $\mathbf{R}$ is constant during the game. Also by Definition 7.16, the Markov property holds. One has to to show that the transition matrix $\mathbf{R}$ is row-stochastic. This is done by confirming that the backchannel praise probability matrix $\mathbf{B}$ and the uniform praise probability matrix $\mathbf{I}$ are both row-stochastic. The convex combination of two row stochastic matrices again yields a row stochastic matrix. $\mathbf{I}$ clearly is row stochastic. As for $\mathbf{B}$ consider an arbitrary row j. According to Definition 7.4 two cases have to be checked. In each case the row has to sum to one.
Case 1. $u_s > 0$

$$
b_{s \to i} = \begin{cases} \dfrac{u_{s \to i}}{u_s} & \text{if } s \neq i \\ 0 & \text{otherwise} \end{cases} \qquad \sum_{k=1}^{n} b_{s \to k} = \sum_{s \neq k} \frac{u_{s \to k}}{u_s} = \frac{1}{u_s} \times \underbrace{\sum_{s \neq k} u_{s \to k}}_{= u_s \text{ by Def.}} = 1
$$

Case 2. $u_s = 0$

$$b_{s \to i} = \begin{cases} \dfrac{1}{n-1} & \text{if } s \neq i \\ 0 & \text{otherwise} \end{cases} \qquad \sum_{k=1}^{n} b_{s \to k} = \sum_{s \neq k} b_{s \to k} = (n-1) \times \frac{1}{n-1} = 1$$

$\square$

7.4.4.2 Convergence of the Reputation Model

As a special subclass of Markov chains, *ergodic* Markov chains turn out to expose convergent behavior if the chain is unbounded in time. As mentioned earlier, ergodic Markov chains have the desirable characteristic to converge to a distinct stationary probability distribution independent from the initial probability distribution. It can be shown that for a Markov chain to be ergodic it is sufficient that the transition matrix contains only strictly positive (non-zero) entries (Langville and Meyer, 2012).

Proposition 7.18. *The estimated praise probability matrix* **R** *is positive, i.e., all entries of* **R** *are greater than zero.*

Proof. Since $a_i \in (0; 1)$ it follows that $(1 - a_i) > 0$, and since **B** is non-negative, for an arbitrary row-index i and an arbitrary column-index j of **R** it holds that

$$r_{i \to j} = a_i \times b_{i \to j} + (1 - a_i) \times \frac{1}{n} \geq (1 - a_i) \times \frac{1}{n} > 0$$

$\square$

Thus, the reputation model of *Backstage* corresponds to an ergodic Markov chain that converges to a stationary distribution.

7.4.5 Reputation as the Eigenvector of the Estimated Praise Probability Matrix

Equation 7.5 can be rewritten in matrix notation as follows:

$$\mathbf{p}_{t+1} = \mathbf{p}_t \times \mathbf{R}$$

According to the considerations above, the repeated application of **R** to $\mathbf{p}_t$ eventually results in a vector $\mathbf{p}^\star$ such that $\mathbf{p}^\star = \lim_{t \to \infty} \mathbf{p}_t$. Moreover, by definition, applying **R** to the stationary vector $\mathbf{p}^\star$ leaves $\mathbf{p}^\star$ unchanged.

$$\mathbf{p}^\star = \mathbf{p}^\star \times \mathbf{R}$$

Thus $\mathbf{p}^\star$ can be characterized as the (left) eigenvector of **R** for the eigenvalue $\lambda = 1$. As to the existence and uniqueness of the eigenvector $\mathbf{p}^\star$ for the eigenvalue $\lambda = 1$, if **R** is a positive matrix, i.e. all entries are greater than zero, then $\lambda = 1$ is a simple and strictly dominant eigenvalue of **R** and a unique positive eigenvector $\mathbf{p}^\star$ for the eigenvalue $\lambda = 1$ exists (Langville and Meyer, 2012). Lemma 7.18 shows that the estimated praise probability matrix **R** is indeed a positive matrix. Hence, a unique eigenvector for the eigenvalue $\lambda = 1$ of **R** exists. The elements of the eigenvector are used as a measure of the students' reputation.

117

7.4.5.1 Initialization of the Model

In this section it is shown that at the beginning, when no ratings are available, the proposed reputation model yields the intuitive result of equally distributed reputation of the students.

Lemma 7.19 (Eigenvector of a Doubly-Stochastic Matrix). *For an $n \times n$ doubly-stochastic matrix $\mathbf{M}$, $\mathbf{p}^\star = \left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$ is an eigenvector for the eigenvalue $\lambda = 1$.*

Proof. Compute the i-th coefficient of the eigenvector (Schickinger and Steger, 2002).

$$(\mathbf{p}^\star \times \mathbf{M})_i = \sum_{k=1}^{n} p_k^\star \times m_{ik} = \sum_{k=1}^{n} \frac{1}{n} \times m_{ik} = \frac{1}{n} \times \underbrace{\sum_{k=1}^{n} m_{ik}}_{\text{col. sum} = 1} = \frac{1}{n}$$

$\square$

When a matrix is doubly-stochastic, the eigenvector for the eigenvalue of one corresponds to the uniform probability distribution. At the beginning, when no ratings have been given by students, the backchannel praise probability matrix has the value $1/(n-1)$ everywhere except for the diagonal elements, which are zero. The row and column sums amount to adding $1/(n-1)$ $(n-1)$ times, which yields 1. Thus, at the beginning, the backchannel praise probability matrix is doubly-stochastic. In Definition 7.16, the uniform praise probability matrix $\mathbf{I}$ is also doubly-stochastic. The convex combination of two doubly-stochastic matrices is again doubly-stochastic. Hence, the estimated praise probability matrix is doubly-stochastic and the uniform probability distribution is the eigenvector for $\lambda = 1$.

Proposition 7.20 (Uniformity of Initial Reputation Values). *If no ratings are available, the eigenvector $\mathbf{p}^\star$ of $\mathbf{R}$ for the eigenvalue $\lambda = 1$ corresponds to the uniform probability distribution. That is, initially, all students have equal reputation.*

7.4.6 Considering Quiz Performance in the Reputation Model

The participation in quizzes may comprise a significant proportion of a student's overall participation in *Backstage*. However, the proportion of quizzes in the lecture is not reflected in the reputation, that is, in the matrix $\mathbf{R}$. This section suggests an extension of the reputation model to also account for the participation in quizzes. A student is (automatically) praised by all the other students, when a quiz is answered correctly by that student. Taking into account a student's quiz performance in the determination of reputation adds objectivity, the degree of which is customizable.

In order to receive a praise for solving a quiz, it might be sensible to take a more relaxed standpoint as to the correctness of answers to quizzes. To account for the gradual nature of correctness, the notion of correctness threshold of a multiple choice quiz is introduced.

Definition 7.21 (Correctness Threshold of a Multiple Choice Quiz). *Define by $\alpha \in (0; 1]$ the percentage threshold of correct answers of a multiple choice quiz in order to be accounted for in the reputation model.*

On the basis of the correctness threshold the quizzes are determined for which praises can be earned.

Definition 7.22 (Quiz Performance Matrix). *Define for a given correctness threshold α the Quiz Praise Probability Matrix $\mathbf{Q}_\alpha = (q_{i \to j}^\alpha)$ where*

$$q_{i \to j}^\alpha = \begin{cases} \dfrac{n_j^\alpha}{\sum\limits_{k \in \mathcal{S}} n_{i \to k}^\alpha} & \textit{if } n_{i \to u}^\alpha \neq 0 \textit{ for some } u \\[4mm] \dfrac{1}{n} & \textit{otherwise (to avoid zero rows)} \end{cases}$$

where n_j^α denotes the number of multiple choice quizzes the student j has solved correctly at least to an α-degree, and $n_{i \to k}^\alpha$ denotes the number of praises given by i to k due to correctly solving multiple choice quizzes at least to an α-degree.

The quiz praise probability matrix also accounts for the difficulty of the quizzes. When many students have reached the α-threshold, the praises of a student s are shared among all these students. As a consequence, these students receive a smaller proportion of s's influence, resulting in a marginal increase of reputation.

Definition 7.23 (Estimated Praise Probability Matrix Including Quiz Performance). *Define the Estimated Praise Probability Matrix Including Quiz Performance as*

$$\mathbf{R}_\alpha' = \mathbf{A} \times \mathbf{Q}_a + (\mathbb{1} - \mathbf{A}) \times \mathbf{I}$$

More generally, the reputation model can be defined by accounting for both the backchannel and quiz performance expertise evaluations.

Definition 7.24 (Estimated Praise Probability Matrix Including Quiz Performance). *Define the Estimated Praise Probability Matrix extended by Quiz Performance as*

$$\mathbf{R}_\alpha = \mathbf{A}_1 \times \mathbf{B} + \mathbf{A}_2 \times \mathbf{Q}_\alpha + (\mathbb{1} - \mathbf{A}_1 - \mathbf{A}_2) \times \mathbf{I}$$

Although the convex combination in the previous definition works for any positive diagonal matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, it seems reasonable to require that $\mathbf{A}_1 + \mathbf{A}_2 < \mathbb{1}$. In the extended reputation model, the activity matrices $\mathbf{A}_i$ can be used to weigh the different components of the praise probability matrix. By increasing the entries of the quiz activity matrix $\mathbf{A}_2$ the students' participation in quizzes can be given more significance. As a consequence, reputation can be increasingly determined on objective evidence by quizzes rather than subjective assessment by student ratings.

7.5 Discussion

In this chapter a reputation system for incentivizing interactions and collectively regulating *Backstage* has been introduced. The reputation system is designed in a way that members of a high reputation have greater influence on the determination of reputation of others than members with relatively low reputation.[18] However, the realization and use of the reputation system in a learning environment such as *Backstage* raise challenges that makes further research necessary.

---

[18] Albeit, the reputation system permits that a large group of members with low reputation who achieved consensus can dominate the public opinion against a small group of well-reputed members.

7.5.1 Modeling and Instantiating Parameters of the Reputation Model

The reputation model described in the previous section is highly flexible and thus allows the model to be adjustable to the context of use to a great extent.

In Section 7.4.1 the need for considering a student's overall rating activity has been discussed as a way to assess the quality-measurety of the student's backchannel praise probability distribution as an estimator for the student's praise probability distribution. Therefore, in Section 7.4.2 a smoothed measure of a student's overall rating activity has been introduced. However, whether smoothing yields acceptable results needs to be investigated in detail. The investigation has to address not only the determination of appropriate smoothing parameters m and d, but also how the smoothing parameters affect the reputation values.

From a modeling perspective, taking the smoothed overall rating activity as a quality-measure seems to be an acceptable simplification. Recall that the idea behind this approach is that the ratings of students who frequently engage in ratings reflect the praise probabilities to a greater extent compared to students who are less active in the rating of backchannel comments. However, whether the smoothed overall rating activity performs well in practice needs to be investigated in system-related evaluations.

An alternative to smoothing is the use of statistical models for the activity coefficients. Statistical models are capable of handling factors that are not or only marginally reflected in the smoothed overall rating activities. For example, the activity coefficients may no longer be considered as independent from each other, which is more realistic. Furthermore, the commenting activity may as well be accounted for in the determination of reputation because the backchannel comments are rated, not the authors. Thus, in order to receive ratings, students have to participate in the backchannel discourse. However, the consideration of statistical models may just as well lead to models that do not take any kind of the students' backchannel activities into account anymore. Rather, the models may be based on expert knowledge or experimental findings. Whether the development of alternative models for the student activity is required needs to be experimentally investigated.

Aside from modeling the activity coefficients, the uniform praise probability matrix **I** in Definition 7.16 (page 116) may also be adjusted to the context of use. Recall that the uniform praise probability reflects that students have no preference of one another, which is the case at the beginning, e.g., when no ratings are available. However, one can surmise that a student who actively engages in the backchannel is more likely to receive ratings for her posts in contrast to students posting only few comments. The assumption that students have equal probabilities of receiving a praise, which is modeled by the uniform praise probability matrix **I**, may not reflect the dependencies of receiving praises from posting comments. If this interdependency should be taken into account to a greater extent it is sensible to replace the uniform praise probability matrix **I** by another row-stochastic matrix that models the students' probabilities of receiving a praise due to their presence in the backchannel.

7.5.2 Considering Negative Ratings in the Reputation Model

It may be worth considering negative ratings in the reputation model as well. However, as it turns out, the proposed model does not seem to be amenable to handling negative ratings in the same way as positive ratings. The application of the same algorithmic principle for the determination of a kind of "negative" reputation is difficult because of the reinforcement of praising built in the model. The reinforcement entails that students with a high negative reputation would exert greater influence in the determination of

the others' negative reputation. One attempt would be to try using the complement of negative reputation, which can be thought of as a "non-negative" reputation. Using this attempt, however, "non-negative" reputation makes the estimated praise probability matrix sub-stochastic. The same problem can be found, for example, when negative ratings are accounted for in the backchannel praise probabilities, e.g. by computing the fraction of positive ratings in all ratings a student has given.[19] Recall that stochasticity of the estimated praise probability matrix ensures that the desired eigenvector exists and that it is unique; furthermore it ensures that an iterative method for obtaining the eigenvector converges independently from the chosen initial vector. All these indispensable properties would be endangered when giving up stochasticity.

Apart from the difficulties associated with explicitly considering negative ratings in the reputation model, one may ask whether considering negative ratings for reputation as proposed in this chapter is necessary in the first place. As the former paragraph suggests, negative ratings are likely to be handled differently than positive ratings. A conceptual justification for it may be difficult to find. It can be surmised that negative ratings are useful for damping reputation and enabling reputation to deteriorate. Both, however, can also be realized without negative ratings. Thus, damping may be realized by appropriately weighting the backchannel praise probability matrix; deterioration of reputation may be realized by establishing a shifting time window and disregarding ratings in the computation that fall outside the window. In particular, shifting time windows also implicitly account for negative ratings of backchannel comments. Since a student may rate a comment only once, an increase of negative ratings may lead to a decrease of positive ratings. A decrease of positive ratings is accounted for in the activity coefficients which results in a low value for reputation.

7.5.3 Integration of Reputation in Backstage

In social media, reputation serves the purposes of reducing uncertainty, incentivizing behavior beneficial for the community and collectively regulating the interactions. Although these purposes are in line with constructivist learning theory, which suggests that students should have control over the learning environment, the effectiveness of reputation in a learning environment such as *Backstage* remains unclear and deserves thorough evaluation. Educational practices may differ from Web 2.0 practices (Bennett et al., 2012), users may interact differently on general-purpose than on educational social media.

While reputation may serve similar purposes in learning environments, both in terms of incentivization and social regulation, the assumption needs to be confirmed in evaluations of the model. To confirm the usefulness of reputation several aspects are likely to play a role. For example, the availability of so-called *cheap pseudonyms* enable students to behave badly without bearing the social costs for it (Friedman and Resnick, 2001). Cheap pseudonyms are pseudonyms that can easily be created by users, e.g. simply by registering anew, in order to escape bad reputation resulting from bad behavior. If a community is aware of cheap pseudonyms, interactions would hardly be considered to be reciprocal, and the shadow of the future would not apply. To avoid cheap pseudonyms, ideally, social media that employ reputation should ensure that a user can register at most once.[20]

---

[19]That is, $b_{s \to i} = u_{s \to i}/(u_s + d_s)$, where u_s denotes the overall number of positive ratings and d_s the overall number of negative ratings given by s.

[20]On *Backstage* this could be achieved by requesting and validating enrollment numbers or some other kind of student identification tokens such as institutionally administered email addresses during registration.

Another factor to be considered is the way reputation is presented and distributed in the learning environment. For instance, it is plausible (and common in other social media) that each backchannel comment displays its author's reputation value, making reputation visible to all students. However, the approach may contribute to the so-called *Matthew Effect* that refers to the accumulation of advantage, where "the rich get richer at a rate that makes the poor become relatively poorer" (Merton, 1968, p. 62). Accordingly, students may be tempted to approve comments without proper reflection, solely for the reason that the author is well-reputed. In this way, well-reputed students gain potentially unjustified approvals, which contributes to a better reputation. It seems, however, that this is an issue that foremost needs to be addressed in the visual presentation of reputation.

Reputation can contribute to a higher acceptance and use of *Backstage* and can enhance the students' awareness of what others expected from them and thus makes social influence visible. Also having the lecturer rate comments may increase the importance of reputation for students as a means of guidance. Furthermore, students should be made aware of the fact that reputation is also visible to the lecturer. Moreover, making reputation more objective, e.g. by considering quiz performance in a student's reputation as proposed in Section 7.4.6 may help to make reputation a useful means of guidance and contribute to engagement in *Backstage*.

7.6 Conclusion

This chapter introduces a model of reputation of students on *Backstage* based on eigenvector centrality. Eigenvector centrality is chosen to allow students who valuably engage on *Backstage* to gain more influence in the determination of reputation than less active students on *Backstage*. Therefore, a hypothetical "game" of passing a badge among *Backstage* participants is developed. The model allows an approach quite similar to *PageRank* that has proven to work very well in practice. This chapter also proposes extensions of the reputation model. Considering a student's quiz performance in her reputation not only allows taking into account a student's activity more thoroughly but also incorporates an objective aspect in reputation. The model is highly customizable as it allows reputation to be defined more on subjective opinions of peers by rating of comments or rather on objective performance-related measurements of quiz performances. Although being conceived for *Backstage*, the reputation model is likely to be applicable in other classroom communication systems and learning environments as well. It is planned to implement and thoroughly investigate the reputation model. Possibly, an investigation may be based on multi-user simulations, since in real settings too many variables are involved that cannot be controlled for (Vassileva, 2012). Nonetheless, a survey-based evaluation on *Backstage* and other learning environments are required to find out whether the system is accepted and yields useful measurements.

Part IV

Conclusion

CONCLUSION AND PERSPECTIVES

This thesis reported about the conception, development and evaluation of the classroom communication system (CCS) *Backstage* that aims at promoting active participation of students in large-class lectures. *Backstage* consists of a microblog-based backchannel and an Audience Response System. The backchannel enables students to initiate communication with others and the lecturer, e.g., for asking and answering questions or receiving social support. The Audience Response System provides quizzes that can be used by the lecturer to promote active involvement of students in the lecture. Audience Response Systems are a valuable feedback source for both the lecturer and the students and give insights into "how well the students are doing". Both components, the backchannel and the Audience Response System, contribute to gathering feedback and enhancing social awareness.

The conception and the development of *Backstage* was informed by a preliminary usability study that investigated usability aspects and analyzed the backchannel communication, in particular question asking by the participants, during a presentation. Among others, the study revealed that a conventional display of backchannel posts in a linear order without any references to the lecture negatively affects comprehensibility and, hence, the effectiveness of a CCS. This finding is also in line with CMC research. As a remedy, slides were integrated into the backchannel as the central structuring artifact. Only those posts are displayed that belong to the currently selected slide. Furthermore, the backchannel was reconsidered to facilitate a collaborative annotation of slides. That is, posts need to be placed at regions of slides relevant for the posts, thus establishing explicit references between the backchannel and the slides. Placing posts on slides allows participants to easily recognize whether a post might be of interest or not. Additionally, a few fixed post categories (e.g., *Question* and *Answer*) were introduced into the backchannel. Post categories mitigate the problem of information entropy (Hiltz and Turoff, 1985). The necessity to find a suitable location on a slide and selecting a proper category for publishing a post introduces a certain extent of effort that aims at preventing irrelevant communication.

In the revision of the first prototype the Audience Response System is integrated with the backchannel by using slides for the display of quiz questions and quiz-related graphics. Answers to the quizzes are submitted using a separate view. During a running

quiz the lecturer is provided with a real-time overview of the submitted responses. When the quiz is stopped the lecturer can choose to add an automatically created summary slide to the lecture slides, e.g., a histogram that displays the students' answers to the quizzes. As ordinary slides quiz summaries can be annotated by participants to discuss the quizzes and their results.

To facilitate filtering of the backchannel and establish self-regulation on *Backstage*, rating and ranking of posts were introduced. This thesis investigated two ranking methods: Ranking by the Bayesian Average Rating seems to put emphasis on the number of a ratings a post received; ranking by the lower bound of Wilson's Confidence Interval seems to weight agreement among raters over the number of ratings. Also, ratings are used for an extensible and customizable eigenvector-based reputation system of the posts' authors similar to *Google*'s PageRank. The eigenvector-based reputation system gives well-reputed students more influence in determining the other participants' reputation than less-reputed students. The system is highly customizable. For example, it can be configured to also consider quiz participation in the computation of reputation.

An extensive evaluation of the prototype in four Computer Science courses showed that *Backstage* contributes to more learning-related activities of students in lectures. The result is underpinned by Gehlen-Baum, Weinberger, et al. (2014) who investigate the students' use of mobile devices brought to conventional and to technology-supported lectures. In conventional lectures (without technology support) students mostly used their mobile devices for lecture-unrelated activities. In contrast, students in *Backstage*-supported lectures used their mobile devices to a notable extent for lecture-related activities.

Backstage was also successfully used for the implementation of new didactic approaches. For example, in one of the evaluated courses, mock exams were conducted. Students corrected and graded their own solutions according to a predefined grading scheme. *Backstage*'s Audience Response System was used for reporting and aggregating the results of the exams. This approach was well received by students. Further studies should be undertaken to investigate possible uses of *Backstage* for more active and interactive teaching in lectures. Besides, more pedagogical studies are necessary in order to find out how *Backstage* affects learning in lectures. However, they may be difficult to conduct considering the number of variables to be taken into account.

8.1 Perspectives

The following sections provide an outlook of possible technological improvements and extensions of *Backstage*. Some of these advancements may be indispensable for the practical use of *Backstage*. For instance, the ways slides are integrated in *Backstage* certainly have an impact on whether lecturers accept the tool and consider it for use in lectures. As pointed out previously, making *Backstage* available for public use requires more focus on stability, availability, and usability of the software. Other advancements of *Backstage*, in turn, are more premature and require further conceptual work.

8.1.1 Using Proper HTML as Input Format for Slides

Using raster graphics for the display of slides in *Backstage* was quite beneficial, since (1) it enabled to quickly realize a prototype based on standard Web technologies that can be used for evaluation purposes and (2) it made it possible to quickly use the prototype with existing lecture slides for experimenting. Yet, the approach has numerous limitations. For example, raster graphics do not permit the search and selection of text. Also, raster graphics are insufficient for embedding multimedia such as audio recordings or videos.

Using proper HTML as an input format would resolve these issues. In fact, HTML subsumes raster graphics. To simplify creation of slides in HTML, it is reasonable to integrate an easy-to-use editor that provides simpler formalisms than HTML such as Wikitext[1] or Markdown[2] both of which can be "transpiled" to HTML.[3] The use of proper HTML also simplifies the Audience Response System in numerous aspects. For example, the input view of a quiz can be integrated into the quiz slide, thus avoiding redundancy and simplifying the usability of *Backstage*.

8.1.2 Identifying Relevant Parts of Slides by Filter Boxes

As a way to enhance awareness of relevant communication, parts of slides may be particularly distinguished by filter boxes, i.e., rectangular areas that encompass meaningful content. Filter boxes can be attributed by titles and tags. Post filters may be attached as well, thus enabling a filtering of the backchannel communication by location on slides and by post-related attributes such as post categories (e.g., *Question*), keywords or ratings that characterize the posts of interest. For example, the lecturer could use filter boxes to filter the backchannel communication to only show "all *Question*s that relate to Proof 1.4".

The filter box approach also provides an alternative to the graphical, two-dimensional annotation of slides by posts, which is especially tedious on small screen devices such as smart phones. With filter boxes, smart phone users can use the titles of the filter boxes to designate the parts to which the comments refer. The actual positioning on the slide can then be accomplished (semi-) automatically. In this way, a one-dimensional annotation of slides, which is much more convenient on small screens, can be realized. A qualitative evaluation of a user interface concept for smart phones has shown that using filter boxes for referencing relevant parts of a slide makes *Backstage* also usable on small screen devices (Hadersberger, 2012).

Several issues are still open. For example, slides can as well be considered to be filter boxes themselves. The resulting hierarchical structure of boxes requires an intuitive navigation. Strategies for a (semi-) automatic positioning of posts using filter boxes needs to be developed. Also, filter boxes create new possibilities for social interactions such as allowing attachments of supplementary material to slides. Apart from the lecturer, it may be interesting to also allow students to add and maintain these attachments (e.g., categorize, rate, comment). That is, by introducing attachments, *Backstage* may become a social sharing platform for topic-related educational resources profitable for reworking lectures.

8.1.3 Developing and Evaluating Further Quiz Types

It may be interesting to realize quizzes other than the conventional types, e.g., multiple choice quizzes or polls. Thus, a special kind of (multiple choice) quiz may invite active participation in a more playful manner by rewarding quick answering of the quiz. For this purpose, simple factual quiz questions seem suitable. For rewarding correct answers, the respondents can be divided into three groups: the "top respondents" who answered the fastest and earn the most points, the "runner ups" who answered the second fastest thus earn fewer points and a given number of "particpants" who earn only a few points. The summary of such a quiz may not only show which answer is correct but also the

---

[1]See Wikipedia, `http://bit.ly/1FmFOzL` (last visited on April 30th, 2015).

[2]See Wikipedia, `http://bit.ly/1d32oOa` (last visited on April 30th, 2015).

[3]For a description on transcompilers see Wikipedia, `http://bit.ly/1All6hv` (last visited on April 30th, 2015).

scoring list with the participants' user names. Furthermore, a course-wide "best scorer" list may be maintained on *Backstage*.

Another kind of quiz may be particularly interesting for teaching and learning programming languages. In this type of quiz students answer to a quiz with executable code fragments. The students' responses may be automatically tested for correctness by tests provided along with the quiz. The degree of correctness of a student's answer is determined by the number of tests passed. The summary may be a histogram showing how the students' answers distribute across the number of tests passed. A realization of this type of quiz is more demanding, both on the lecturer's side as they would have to develop these tests for the quizzes, as well as on the technological side. For example, the compilation of code, if required, needs to be robust and must also be able to handle load peaks, since in live quizzes answers usually come in bursts. The code fragments should also be tested in sandboxed environments[4] so as to prevent system breakdowns due to erroneous answers (e.g., non-terminating loops or unwanted access to system resources).

Further research is necessary to fully understand how quizzes should be used on *Backstage*. In particular, a more holistic approach to investigating quizzes may be beneficial. For example, difficulty levels of (multiple choice) questions are often not only influenced by the questions but also by the answer options provided. Timing can be another issue. Thus, research on question asking by the lecturer has shown that timing, e.g. wait time, is a crucial factor (Rowe, 1986). Thus, in the frontchannel, i.e., the face-to-face exchange in the lecture hall, lecturers tend to wait not long enough for students to properly think of the question and to formulate an answer. This finding may as well be relevant for quizzes. However, a prolonged wait time may have adverse effects on students' attention as well, e.g., by contributing to fatigue. Also, running quizzes in the last minutes of a lecture may negatively affect acceptance by students and likewise reduce the utility of feedback gathered by quizzes.

Other uses or misuses of quizzes are worth investigating. *Backstage* can be a valuable tool for experimenting with various quiz types and for evaluating proper uses of an Audience Response System in lectures as *Backstage* facilitates collection of data on the students' as well as the lecturers' behavior in various contexts.

8.1.4 Conceiving Further Awareness and Social Visualization Components

More elaborate concepts of stimulating social awareness on *Backstage* are conceivable. Displaying a summary of awareness metrics is referred to as social visualization (Vassileva, 2012). According to Vassileva, social visualization can be a means to motivate desirable behavior and to sustain social norms that can yield engagement and participation. A first concept of social visualization on *Backstage* has been conceived by (Pohl, Gehlen-Baum, and Bry, 2012; see Figure 8.1).

It borrows from the following metaphor: Often, more engaging students tend to sit at the front rows, while less active students rather prefer sitting at the back rows of a lecture hall. The concept makes use of this metaphor by displaying a lecture hall in which participants on *Backstage* are represented as discs of different sizes and colors at different locations in the lecture hall according to the participants' activities. The more active a participant is the further is the participant's disc displayed at the front of the lecture hall. The two discs at the center of the lecture hall represent the average student and the current participant.

---

[4]See Wikipedia, `http://bit.ly/1ncki50` (last visited on April 30th, 2015).

The center is flanked by a decent number of discs that represent other participants. Instead of displaying the entire audience, which may quickly become confusing, the concept proposes to provide follower-relationships, akin to *Twitter*. A recommender system can be used to help establish follower-relationships with others. The size of a disc may indicate the participant's reputation. Likewise the color of a disc can be used to indicate the ratings received. Accordingly, the more negative ratings a participant receives the more the participant's disc is displayed in red.

When a disc is selected, further information about the selected participant is shown below the depicted lecture hall. The profile information of the selected participant may comprise a picture and a summary of the participant's activities on *Backstage*, e.g., presence data, number of backchannel posts and ratings received. Of course, all participants should be able to decide which profile information is accessible by others.
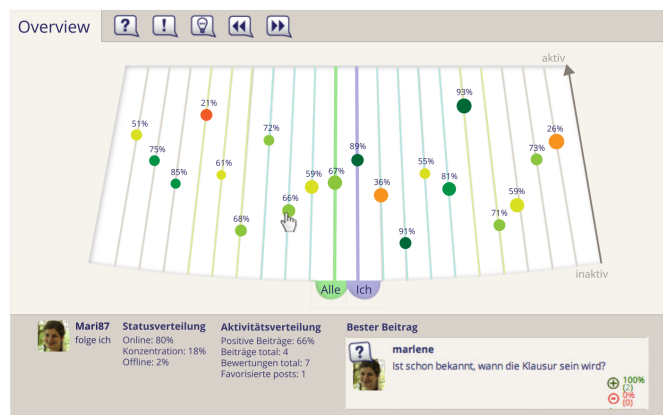
Figure 8.1: Concept Drawing of a Social Visualization Component for *Backstage*. The concept borrows the metaphor of a lecture hall in which active students tend to sit at the front and less active students rather at the back of a lecture hall.

Furthermore, the Audience Response System can be extended to track a student's quiz performance. An analysis of the data would raise a student's awareness of difficult lecture topics. Intelligent tutoring systems make use of a probabilistic method called Bayesian Knowledge Tracing (BKT; Corbett and Anderson, 1995). BKT is a Hidden Markov Model (HMM; Rabiner, 1989) that models learning progress as a probability of a student's grasp on the basis of the answers given to multiple choice quizzes. If the probability of a student's grasp is sufficiently large (say, greater than 0.5), the student is granted access to the next topic. The model explicitly takes into account slips (a wrong answer has been accidentally chosen) and guesses (a correct answer has been chosen by chance).

In a similar way, *Backstage* could track a student's learning progress inferred from the quiz performances. Using a HMM similar to BKT for this purpose seems reasonable, since HMMs support various kinds of queries. A HMM can be used not only to find explanations of a student's development, but also to predict future development based on the quiz performance observed so far. Thus, the model can provide valuable information for more adaptive guidance of students on *Backstage*, e.g., by providing adaptive educational scripts (Weinberger, 2011). However, the instantiation of such a model, i.e., finding appropriate values for the parameters, can be quite sophisticated. Nevertheless, the assessment of a student over time using a HMM can be reported back in a comprehensible form, e.g., as a traffic light system, indicating whether the student

seems to do well, or whether greater effort is needed, e.g., by reworking a given topic more thoroughly.

8.1.5 Considering a More Structured Approach Towards Interactive Lectures

In joint work with Vera-Gehlen Baum a concept about how lecturers can be supported in arranging and running interactive lectures with *Backstage* was initiated. This section considers the model as a basis for a recommender system for lectures.[5]

To summarize, the model assumes that lecturers, particularly when inexperienced, are likely to profit from a more structured approach towards interactive lectures. To account for this, a lecture is considered as a sequence of so-called *topics*, each of which takes about 20 to 30 minutes. Limiting time ensures periodic activity switching by students which has been shown to conduce to students' attention (see Stuart and Rutherford, 1978; Bligh, 1998; Young, Robinson, and Alberts, 2009). According to the model, a *topic* is associated with a *purpose*, e.g., knowledge building or problem solving, and a *goal* defined by the lecturer. Usually, the *purpose* of a topic determines the elements of the *topic*, i.e., *presentation* of information, *instruction* such as question asking or repetition by the lecturer, and *group* activities such as discussions. A *topic* ends with a *closure*, which can be a short break for reflection, a summary by the lecturer or a quiz.

In future versions of *Backstage*, a guiding system based on the proposed model could be integrated. The guiding system intends to provide support during the preparation and the execution phase of lectures on *Backstage*. Thus, the preparation of a lecture on *Backstage* would require the creation of a set of *topic*s. Each topic would require the provision of a learning *goal* and the assignment of lecture slides to the *topic*'s elements, i.e., *presentation*, *instruction*, and *group* activities. Note, that the preparation does not necessitate that the lecturer states the *purpose* of a *topic*. It may be sensible to let *Backstage* derive on the basis of a *topic*'s configuration what *purpose*s the lecturer seems to pursue. In this way, no particular *purpose* may be prescribed but instead feedback about whether the *topic*s match the lecturer's intentions may be provided. This approach also allows a mixture of *purpose*s, e.g., 60% problem solving and 40% presentation are conceivable.

During the execution phase of a lecture, *Backstage* may record, analyze and display metrics that are supposed to help the lecturer. The time elapsed during *topic*s may be displayed at the lecturer's user interface. Such measurements may take into account other similar courses of the same lecturer or the same course held by similar lecturers as a basis for predictions. On the basis of these predictions an early warning system may notify the lecturer whether there is a risk that not all *topic*s can be worked off and suggest alternatives, e.g., which *topic*s to leave out. Timers can be provided to help the lecturer to keep an eye on time, e.g., to uphold time restrictions of running quizzes. Whether audio and vibration cues should be used for signaling needs to be investigated.

For arranging and conducting *group* activities, *Backstage* could be extended to allow for formation of ad hoc groups (see Figure 8.2) according to relevant parameters. During a *group* activity, a group may be provided with a private *Backstage* instance equipped with slides and quizzes relevant for the group. The lecturer may be provided with a dashboard that gives an overview of the groups' activities. The lecturer can then join a private *Backstage* instance so as to skim over the group-specific backchannel or to run quizzes within a group, for example. Beside private *Backstage* instances, groups

---

[5]An in-depth description of the model, and the pedagogical rationale behind the model, can be found in Gehlen-Baum (to appear).
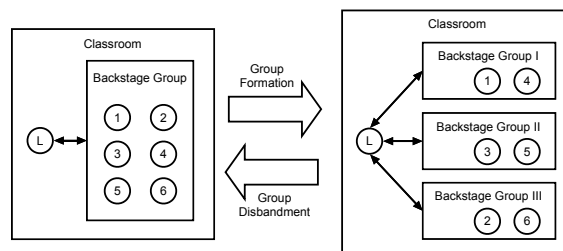
Figure 8.2: Group Formation in and the Group-wise Provision of *Backstage* (L = Lecturer, 1-6 = Students).

may also be enabled to share one global instance of *Backstage*. Thus, the lecturer may ask members of a group to collectively create backchannel posts. Similarly, answers to quizzes may be submitted group-wise. When the lecturer terminates *group* activities the groups are disbanded. If the groups worked on private instances of *Backstage* the outcomes of the group activities, e.g., group-specific backchannel communication and quiz answers, may be merged into the public *Backstage* instance. Several merging strategies are conceivable and need to be worked out.

BIBLIOGRAPHY

Abrahamson, A. L. (1999). "Teaching with Classroom Communication System - What it Involves and Why it Works." In: *Invited Paper presented at the 7th International Workshop "New Trends in Physics Teaching".* Puebla, Mexico (cit. on p. 19).

Agichtein, Eugene, Carlos Castillo, Debora Donato, Aristides Gionis, and Gilad Mishne (2008). "Finding High-Quality Content in Social Media." In: *Proceedings of the International Conference on Web Search and Web Data Mining.* WSDM '08. New York, NY, USA: ACM, pp. 183–194 (cit. on pp. 90 sq.).

Aitken, Joan E. and Michael R. Neer (1991). "Variables Associated with Question-Asking in the College Classroom." In: Atlanta, USA (cit. on pp. 16 sq.).

Aleven, Vincent, Elmar Stahl, Silke Schworm, Frank Fischer, and Raven Wallace (2003). "Help Seeking and Help Design in Interactive Learning Environments." In: *Review of Educational Research* 73.3, pp. 277–320 (cit. on p. 16).

Altman, Alon and Moshe Tennenholtz (2005). "Ranking Systems: The PageRank Axioms." In: *Proceedings of the 6th ACM conference on Electronic commerce.* Vancouver, BC, Canada: ACM, pp. 1–8 (cit. on p. 108).

Anderson, R., T. Vandegrift, S. Wolfman, and K. Yasuhara (2003). "Designing for Change in Networked Learning Environments." In: *Designing for Change in Networked Learning Environments.* Springer Netherlands, pp. 119–123 (cit. on p. 19).

Apel, Hans Jürgen (1999). ""Das Abenteuer auf dem Katheder". Zur Vorlesung als rhetorische Lehrform." In German. In: *Zeitschrift für Pädagogik* 45.1, pp. 61–79 (cit. on p. 3).

Bauer, Henry H. and William E. Snizek (1989). "Encouraging Students in Large Classes to Ask Questions: Some Promising Results from Classes in Chemistry and Sociology." In: *Teaching Sociology* 17.3. ArticleType: research-article / Full publication date: Jul., 1989 / Copyright © 1989 American Sociological Association, pp. 337–340 (cit. on p. 17).

Baumgart, Daniel, Alexander Pohl, Vera Gehlen-Baum, and François Bry (2012). "Providing Guidance on Backstage, a Novel Digital Backchannel for Large Class Teaching." In: *Education in a Technological World: Communicating Current and Emerging Research and Technological Efforts.* Badajoz, Spain: Formatex, pp. 364–371 (cit. on p. 41).

Beatty, Ian (2004). "Transforming Student Learning with Classroom Communication Systems." In: *EDUCAUSE Research Bulletin* 2004.3, pp. 1–13 (cit. on pp. 10 sq.).

Bennett, Sue, Andrea Bishop, Barney Dalgarno, Jenny Waycott, and Gregor Kennedy (2012). "Implementing Web 2.0 Technologies in Higher Education: A Collective Case Study." In: *Computers & Education* 59.2, pp. 524–534 (cit. on p. 121).

Bian, Jiang, Yandong Liu, Eugene Agichtein, and Hongyuan Zha (2008a). "A Few Bad Votes too Many? Towards Robust Ranking in Social Media." In: *Proceedings of the 4th International Workshop on Adversarial Information Retrieval on the Web*. AIRWeb '08. New York, NY, USA: ACM, pp. 53–60 (cit. on p. 90).

— (2008b). "Finding the Right Facts in the Crowd: Factoid Question Answering over Social Media." In: *Proceedings of the 17th International Conference on World Wide Web*. WWW '08. New York, NY, USA: ACM, pp. 467–476 (cit. on pp. 90 sq.).

Bligh, Donald A. (1998). *What's the Use of Lectures?* Exeter, England: Intellect Books (cit. on pp. 3, 5, 19, 130).

BMBF (2013). *Studierende nach Hochschularten, Ländern und Geschlecht*. In German. Time Series. German Federal Ministry of Education and Research (cit. on p. 3).

Bolton, Gary, Elena Katok, and Axel Ockenfels (2002). *How Effective are Online Reputation Mechanisms?* Papers on Strategic Interaction 2002-25. Max Planck Institute of Economics, Strategic Interaction Group (cit. on p. 104).

Bonacich, Phillip (1987). "Power and Centrality: A Family of Measures." In: *American Journal of Sociology* 92.5, pp. 1170–1182 (cit. on p. 106).

— (2007). "Some Unique Properties of Eigenvector Centrality." In: *Social Networks* 29.4, pp. 555–564 (cit. on p. 106).

Bonacich, Phillip and Paulette Lloyd (2004). "Calculating Status with Negative Relations." In: *Social Networks* 26.4, pp. 331–338 (cit. on p. 106).

Bortz, Jürgen and Nicola Döring (2009). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. In German. 4th. Heidelberg: Springer (cit. on pp. 65 sqq.).

Boyle, James T. and David J. Nicol (2003). "Using Classroom Communication Systems to Support Interaction and Discussion in Large Class Settings." In: *Association for Learning Technology Journal* 11.3, pp. 43–57 (cit. on p. 19).

Brin, Sergey and Lawrence Page (1998). "The Anatomy of a Large-Scale Hypertextual Web Search Engine." In: *Proceedings of the Seventh International World Wide Web Conference* 30.1–7, pp. 107–117 (cit. on pp. 93, 107).

Brown Lawrence, D., T. Cai Tony, and Anirban DasGupta (2001). "Interval Estimation for a Binomial Proportion." In: *Statistical Science* 16.2, pp. 101–17 (cit. on p. 95).

Brugger, Pia and Miriam Wolters (2012). "Von der Hochschulreife zum Studienabschluss." In German. In: *Wirtschaft und Statistik* 8. (In German), pp. 655–663 (cit. on p. 3).

Bry, François (2013). "Human Computation-Enabled Network Analysis for a Systemic Credit Risk Rating." In: *Handbook of Human Computation.* Ed. by Pietro Michelucci. Springer-Verlag (cit. on p. 108).

Bry, François, Vera Gehlen-Baum, and Alexander Pohl (2011). "Promoting Awareness and Participation in Large Class Lectures: The Digital Backchannel Backstage." In: *Proceedings of the IADIS International Conference e-society.* Spain, Avila, pp. 27–34 (cit. on p. 7).

Bry, François and Alexander Pohl (2014). "Backstage: A Social Medium for Large Classes." In: *Campus Transformation - Education, Qualification and Digitalization.* Ed. by Frank Keuper and Heinrich Arnold. Berlin, Germany: Logos, pp. 255–280 (cit. on pp. 9 sq., 41).

Caldwell, Jane E. (2007). "Clickers in the Large Classroom: Current Research and Best-Practice Tips." In: *CBE-Life Sciences Education* 6.1, pp. 9–20 (cit. on p. 18).

Carbone, Elisa (1999). "Students Behaving Badly in Large Classes." In: *New Directions for Teaching and Learning* 1999.77, pp. 35–43 (cit. on p. 5).

Chen, Stanley F. and Joshua Goodman (1996). "An Empirical Study of Smoothing Techniques for Language Modeling." In: *Proceedings of the 34th Annual Meeting on Association for Computational Linguistics.* ACL '96. Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 310–318 (cit. on p. 112).

Chin, Christine and Jonathan Osborne (2008). "Student's Questions: A Potential Resource for Teaching and Learning Science." In: *Studies in Science Education* 44.1, pp. 1–39 (cit. on pp. 12, 16 sq.).

Clark, Richard E. (1994). "Media will Never Influence Learning." In: *Educational Technology Research and Development* 42.2, pp. 21–29 (cit. on pp. 6, 27 sq.).

Cogdill, Sharon and Judith Kilborn (2001). "Backchannel: Whispering in Digital Conversation." In: *Proc. of the 34th Hawaii International Conference on System Sciences (HICSS 2001).* Hawaii, pp. 8–15 (cit. on pp. 18, 65).

Cook, Karen S., Richard M. Emerson, Mary R. Gillmore, and Toshio Yamagishi (1983). "The Distribution of Power in Exchange Networks: Theory and Experimental Results." In: *American Journal of Sociology* 89.2, pp. 275–305 (cit. on p. 106).

Cooper, James L. and Pamela Robinson (2000). "The Argument for Making Large Classes Seem Small." In: *New Directions for Teaching and Learning* 2000.81, pp. 5–16 (cit. on pp. 3, 5).

Corbett, Albert T. and John R. Anderson (1995). "Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge." In: *User Modeling and User-Adapted Interaction* 4, pp. 253–278 (cit. on p. 129).

Costa, Cristina, Guenter Beham, Wolfgang Reinhardt, and Martin Sillaots (2008). "Microblogging In Technology Enhanced Learning: A Use-Case Inspection of PPE Summer School." In: *Proceedings of the 2nd SIRTEL Workshop on Social Information Retrieval for Technology Enhanced Learning.* Ed. by Riina Vuorikari, Barbara Kieslinger, Ralf Klamma, and Erik Duval. Maastricht, pp. 1–9 (cit. on p. 29).

Costin, Frank (1972). "Lecturing Versus Other Methods of Teaching: A Review of Research." In: *British Journal of Educational Technology* 3.1, pp. 4–31 (cit. on p. 3).

Crouch, Catherine H. and Eric Mazur (2001). "Peer Instruction: Ten Years of Experience and Results." In: *American Journal of Physics* 69.9, pp. 970–977 (cit. on p. 19).

Dancer, Diane and Patty Kamvounias (2005). "Student Involvement in Assessment: A Project Designed to Assess Class Participation Fairly and Reliably." In: *Assessment & Evaluation in Higher Education* 30.4, pp. 445–454 (cit. on p. 9).

Das Sarma, Anish, Atish Das Sarma, Sreenivas Gollapudi, and Rina Panigrahy (2010). "Ranking Mechanisms in Twitter-Like Forums." In: *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining*. WSDM '10. New York, NY, USA: ACM, pp. 21–30 (cit. on pp. 90 sq.).

Deci, Edward L., M. Ryan Richard, and Richard Koestner (1999). "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation." In: *Psychological Bulletin* 125.6, pp. 627–668 (cit. on p. 92).

Deci, Edward L. and Richard M. Ryan (1993). "Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik." In German. In: *Zeitschrift für Pädagogik* 39.2, pp. 223–238 (cit. on p. 38).

Dellarocas, Chrysanthos, Ming Fan, and Charles A. Wood (2004). *Self-Interest, Reciprocity, and Participation in Online Reputation Systems*. SSRN Scholarly Paper ID 585402. Rochester, NY: Social Science Research Network (cit. on p. 104).

DeStatis (2012). *Number of Students in Germany at Record High*. Press Release. Statistical Office of Germany (cit. on p. 3).

Ding, Yi and Xue Li (2005). "Time Weight Collaborative Filtering." In: *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*. CIKM '05. New York, NY, USA: ACM, pp. 485–492 (cit. on p. 98).

Döring, Nicola (2003). *Sozialpsychologie des Internet*. In German. Göttingen, Germany: Hogrefe (cit. on pp. 13 sqq., 47 sq.).

Dourish, Paul and Victoria Bellotti (1992). "Awareness and Coordination in Shared Workspaces." In: *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*. CSCW '92. New York, NY, USA: ACM, pp. 107–114 (cit. on p. 12).

Dourish, Paul and Sara Bly (1992). "Portholes: Supporting Awareness in a Distributed Work Group." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '92. New York, NY, USA: ACM, pp. 541–547 (cit. on p. 12).

Draper, S. W. and M. I. Brown (2004). "Increasing Interactivity in Lectures Using an Electronic Voting System." In: *Journal of Computer Assisted Learning* 20.2, pp. 81–94 (cit. on pp. 18 sq., 27, 52).

Dufresne, R. J., W. J. Gerace, W. J. Leonard, J. P. Mestre, and L. Wenk (1996). "Classtalk: A Classroom Communication System for Active Learning." In: *Journal of Computing in Higher Education* 7, pp. 3–47 (cit. on p. 19).

Dunbar, R. I., N. D. Duncan, and D. Nettle (1995). "Size and Structure of Freely Forming Conversational Groups." In: *Human nature (Hawthorne, N.Y.)* 6.1, pp. 67–78 (cit. on p. 5).

Earle, Paul S., Daniel C. Bowden, and Michelle Guy (2011). "Twitter Earthquake Detection: Earthquake Monitoring in a Social World." In: *Annals of Geophysics* 54.6, pp. 708–715 (cit. on p. 26).

Ebersbach, Anja, Markus Glaser, and Richard Heigl (2011). *Social Web*. In German. 2nd Edition. Stuttgart: UTB (cit. on pp. 6, 22–25).

Ebner, Martin (2009). "Introducing Live Microblogging: How Single Presentations can be Enhanced by the Mass." In: *Journal of Research in Innovative Teaching* 2.1, pp. 108–111 (cit. on pp. 17, 27, 29).

— (2011). "Is Twitter a Tool for Mass-Education?" In: *Proceedings of the 4th International Conference on Student Mobilty and ICT*. Vienna, Austria, pp. 1–6 (cit. on pp. 21, 28 sqq., 101).

Ebner, Martin, Christian Haintz, Karin Pichler, and Sandra Schön (2014). "Technologiegestützte Echtzeit-Interaktion in Massenvorlesungen im Hörsaal – Entwicklung und Erprobung eines digitalen Backchannels während der Vorlesung." In German. In: *Lernräume gestalten – Bildungskontexte vielfältig denken.* Ed. by Klaus Rummler. Münster: Waxmann, pp. 567–578 (cit. on pp. 18, 22).

Ebner, Martin, Conrad Lienhardt, Matthias Rohs, and Iris Meyer (2010). "Microblogs in Higher Education – A Chance to Facilitate Informal and Process-Oriented Learning?" In: *Computers & Education* 55.1, pp. 92–100 (cit. on p. 27).

Ebner, Martin and Wolfgang Reinhardt (2009). "Social Networking in Scientific Conferences–Twitter as Tool for Strengthen a Scientific Community." In: *Workshop Science 2.0 for TEL, ECTEL 2009, Nice, Italy* (cit. on p. 27).

Ebner, Martin and Mandy Schiefner (2008). "Microblogging - More than Fun?" In: *Proceedings of the IADIS Mobile Learning Conference 2008.* Portugal, pp. 155–159 (cit. on p. 27).

Ertl, B. and H. Mandl (2004). *Kooperationsskripts als Lernstrategie.* Research Report 172. Department of Psychology, University of Munich, p. 16 (cit. on p. 16).

Falk, Armin and Urs Fischbacher (2006). "A Theory of Reciprocity." In: *Games and Economic Behavior* 54.2, pp. 293–315 (cit. on p. 105).

Fassinger, Polly A. (1995). "Understanding Classroom Interaction: Students' and Professors' Contributions to Students' Silence." In: *The Journal of Higher Education* 66.1, pp. 82–96 (cit. on p. 17).

Fies, Carmen and Jill Marshall (2006). "Classroom Response Systems: A Review of the Literature." In: *Journal of Science Education and Technology* 15.1, pp. 101–109 (cit. on p. 18).

Friedman, Eric J. and Paul Resnick (2001). "The Social Cost of Cheap Pseudonyms." In: *Journal of Economics & Management Strategy* 10.2, pp. 173–199 (cit. on p. 121).

Friesen, N. and S. Lowe (2012). "The Questionable Promise of Social Media for Education: Connective Learning and the Commercial Imperative." In: *Journal of Computer Assisted Learning* 28.3, pp. 183–194 (cit. on p. 6).

Fritschner, Linda Marie (2000). "Inside the Undergraduate College Classroom: Faculty and Students Differ on the Meaning of Student Participation." In: *The Journal of Higher Education* 71.3, pp. 342–362 (cit. on p. 17).

Gaertner, Wulf (2009). *A Primer in Social Choice Theory.* Revised Edition. LSE Perspectives in Economic Analysis. New York, NY, USA: Oxford University Press (cit. on p. 90).

Garcin, Florent, Boi Faltings, and Radu Jurca (2009). "Aggregating Reputation Feedback." In: *Proceedings of the First International Conference on Reputation: Theory and Technology.* Vol. 1. 1. Gargonza, Italy, pp. 62–67 (cit. on p. 68).

Gehlen-Baum, Vera (to appear). "Mobile Geräte in der Präsenzlehre: Ablenkung oder Lernchance? – Von der unstrukturierten Nutzung von Smartphone & Co. hin zu einem orchestrierten Modell für Vorlesungen." In German. Dissertation. Munich: Ludwig-Maximilian University Munich (cit. on pp. 83, 130).

Gehlen-Baum, Vera, Alexander Pohl, and François Bry (2011). "Assessing Backstage–A Backchannel for Collaborative Learning in Large Classes." In: *Proceedings of the 14th International Conference on Interactive Collaborative Learning (ICL 2011).* Piestany, Slovakia, pp. 154–160 (cit. on p. 29).

Gehlen-Baum, Vera, Armin Weinberger, Alexander Pohl, and François Bry (2014). "Technology Use in Lectures to Enhance Students's Attention." In: *Proceedings of the 9th International Conference on Technology Enhanced Learning.* Graz, Austria, pp. 125–137 (cit. on pp. 83, 85, 126).

Gibbs, Graham and Alan Jenkins, eds. (1992). *Teaching Large Classes in Higher Education.* London: Kogan Page (cit. on pp. 3, 5).

Gutwin, Carl, Gwen Stark, and Saul Greenberg (1995). "Support for Workspace Awareness in Educational Groupware." In: *Proceedings of the First International Conference on Computer Support for Collaborative Learning.* Bloomington, USA, pp. 147–156 (cit. on p. 13).

Gysbers, Vanessa, Jill Johnston, Dale Hancock, and Gareth Denyer (2011). "Why do Students Still Bother Coming to Lectures, When Everything is Available Online?" In: *International Journal of Innovation in Science and Mathematics Education* 19.2, pp. 20–36 (cit. on p. 10).

Hadersberger, Julia (2012). "Backstage Mobile – Merging Usability Guidelines with Educational Requirements." Supervised by François Bry and Alexander Pohl. Diploma Thesis. Munich: Ludwig-Maximilian University (cit. on p. 127).

Hadersberger, Julia, Alexander Pohl, and François Bry (2012). "Discerning Actuality in Backstage – Comprehensible Contextual Aging." In: *Proceedings of the European Conference on Technology Enhanced Learning.* Saarbrücken, Germany, pp. 126–139 (cit. on p. 89).

Haintz, Christian, Karin Pichler, and Martin Ebner (2014). "Developing a Web-Based Question-Driven Audience Response System Supporting BYOD." In: *Journal of Universal Computer Science* 20.1, pp. 39–56 (cit. on pp. 18, 22).

Harry, Drew, Joshua Green, and Judith Donath (2009). "Backchan.nl: Integrating Backchannels in Physical Space." In: *Proceedings of the 27th CHI Conference on Human Factors in Computing Systems.* Boston, USA, pp. 1361–1370 (cit. on p. 21).

Harry, Drew, Dan Gutierrez, Joshua Green, and Judith Donath (2008). "Backchan.nl: Integrating Backchannels with Physical Space." In: *CHI '08 Extended Abstracts on Human Factors in Computing Systems.* New York, NY, USA: ACM, pp. 2751–2756 (cit. on pp. 90, 98).

Hartley, James and Ivor K. Davies (1978). "Note-Taking: A Critical Review." In: *Programmed Learning and Educational Technology* 15.3, pp. 207–24 (cit. on p. 16).

Hartmann, Michael (2007). *Eliten und Macht in Europa: Ein Internationaler Vergleich.* In German. Frankfurt/Main; New York: Campus (cit. on p. 3).

Hattie, John (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement.* London: Routledge (cit. on p. 11).

Hattie, John and Helen Timperley (2007). "The Power of Feedback." In: *Review of Educational Research* 77.1, pp. 81–112 (cit. on p. 11).

Haveliwala, Taher H. (2002). "Topic-sensitive PageRank." In: *Proceedings of the 11th international conference on World Wide Web.* Honolulu, Hawaii, USA: ACM, pp. 517–526 (cit. on p. 108).

Herring, Susan (1999). "Interactional Coherence in CMC." In: *Journal of Computer-Mediated Communication* 4.4 (cit. on pp. 13, 15, 18, 45, 48).

Hieronymi, Pamela (2012). "Don't Confuse Technology With College Teaching." In: *The Chronicle of Higher Education* (cit. on p. 27).

Hiltz, Starr Roxanne and Murray Turoff (1985). "Structuring Computer-Mediated Communication Systems to Avoid Information Overload." In: *Communications of the ACM* 28.7, pp. 680–690 (cit. on pp. 42 sq., 84, 92, 125).

Huber, George (1982). "Organizational Information Systems: Determinants of their Performance and Behavior." In: *Management Science* 28.2, pp. 138–153 (cit. on p. 42).

Huo, Wenyu and Vassilis J. Tsotras (2010). "Temporal Top-k Search in Social Tagging Sites using Multiple Social Networks." In: *Proceedings of the 15th International Conference on Database Systems for Advanced Applications - Volume Part I.* DASFAA'10. Berlin, Heidelberg: Springer-Verlag, pp. 498–504 (cit. on p. 98).

Java, Akshay, Xiaodan Song, Tim Finin, and Belle Tseng (2007). "Why We Twitter: Understanding Microblogging Usage and Communities." In: *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis.* New York, NY, USA: ACM, pp. 56–65 (cit. on p. 25).

Jensen, Carlos, John Davis, and Shelly Farnham (2002). "Finding Others Online: Reputation Systems for Social Online Spaces." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '02. New York, NY, USA: ACM, pp. 447–454 (cit. on p. 104).

Jones, Norah, Haydn Blackey, Karen Fitzgibbon, and Esyin Chew (2010). "Get Out of MySpace!" In: *Computers & Education* 54.3, pp. 776–782 (cit. on p. 6).

Jøsang, Audun, Roslan Ismail, and Colin Boyd (2007). "A Survey of Trust and Reputation Systems for Online Service Provision." In: *Decision Support Systems* 43.2, pp. 618–644 (cit. on pp. 104 sq.).

Juhn, Chinhui, Dae Il Kim, and Francis Vella (2005). "The Expansion of College Education in the United States: Is there Evidence of Declining Cohort Quality?" In: *Economic Inquiry* 43.2, pp. 303–315 (cit. on p. 3).

Junco, R., G. Heiberger, and E. Loken (2011). "The Effect of Twitter on College Student Engagement and Grades." In: *Journal of Computer Assisted Learning* 27.2, pp. 119–132 (cit. on p. 27).

Kam, Matthew, Jingtao Wang, Alastair Iles, Eric Tse, Jane Chiu, Daniel Glaser, Orna Tarshish, and John Canny (2005). "Livenotes: A System for Cooperative and Augmented Note-Taking in Lectures." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Portland, Oregon, USA: ACM, pp. 531–540 (cit. on p. 20).

Kamvar, Sepandar D., Mario T. Schlosser, and Hector Garcia-Molina (2003). "The Eigentrust Algorithm for Reputation Management in P2P Networks." In: *Proceedings of the 12th International Conference on World Wide Web*. WWW '03. New York, NY, USA: ACM, pp. 640–651 (cit. on p. 90).

Kaplan, Andreas M. and Michael Haenlein (2010). "Users of the World, Unite! The Challenges and Opportunities of Social Media." In: *Business Horizons* 53.1, pp. 59–68 (cit. on pp. 6, 23).

— (2011). "The Early Bird Catches the News: Nine Things You Should Know About Micro-Blogging." In: *Business Horizons* 54.2, pp. 105–113 (cit. on p. 25).

Karp, David and William Yoels (1976). "The College Classroom: Some Observations on the Meaning of Student Participation." In: *Sociology and Social Research* 60, pp. 421–439 (cit. on pp. 4, 17, 28).

Kassens-Noor, Eva (2012). "Twitter as a Teaching Practice to Enhance Active and Informal Learning in Higher Education: The Case of Sustainable Tweets." In: *Active Learning in Higher Education* 13.1, pp. 9–21 (cit. on p. 27).

Katz, Leo (1953). "A New Status Index Derived from Sociometric Analysis." In: *Psychometrika* 18.1, pp. 39–43 (cit. on p. 106).

Kay, Robin H. and Ann LeSage (2009). "Examining the Benefits and Challenges of Using Audience Response Systems: A Review of the Literature." In: *Computers & Education* 53.3, pp. 819–827 (cit. on pp. 18, 83).

Kellogg, Wendy A., Thomas Erickson, Tracee Vetting Wolf, Stephen Levy, Jim Christensen, Jeremy Sussman, and William E. Bennett (2006). "Leveraging Digital Backchannels to Enhance User Experience in Electronically Mediated Communication." In: ACM Request Permissions, pp. 451–454 (cit. on pp. 6, 17).

Kietzmann, Jan H., Kristopher Hermkens, Ian P. McCarthy, and Bruno S. Silvestre (2011). "Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media." In: *Business Horizons* 54.3, pp. 241–251 (cit. on pp. 23 sq.).

King, Alison (1992). "Facilitating Elaborative Learning through Guided Student-Generated Questioning." In: *Educational Psychologist* 27.1, pp. 111–126 (cit. on p. 48).

— (1994). "Guiding Knowledge Construction in the Classroom: Effects of Teaching Children How to Question and How to Explain." In: *American Educational Research Journal* 31.2, pp. 338–368 (cit. on pp. 17, 34, 36 sq.).

— (2002). "Structuring Peer Interaction to Promote High-Level Cognitive Processing." In: *Theory into Practice* 41.1, pp. 33–39 (cit. on p. 47).

— (2007). "Scripting Collaborative Learning Processes: A Cognitive Perspective." In: *Scripting Computer-Supported Collaborative Learning.* Ed. by F. Fischer, I. Kollar, H. Mandl, and Jörg M. Haake. Boston MA/USA: Springer US, pp. 13–37 (cit. on p. 17).

Kleinberg, Jon M. (1999). "Authoritative Sources in a Hyperlinked Environment." In: *Journal of the ACM* 46.5, pp. 604–632 (cit. on p. 106).

Kollar, Ingo, Frank Fischer, and James D. Slotta (2005). "Internal and External Collaboration Scripts in Web-based Science Learning at Schools." In: *Proceedings of the International Conference on Computer Support for Collaborative Learning: The Next 10 Years! International Society of the Learning Sciences,* pp. 331–340 (cit. on pp. 47 sq., 85).

Kowalski, B. (1987). "Teaching Less and Learning More? - A Personal Experience." In: *Innovations in Education & Training International* 24.3, pp. 174–186 (cit. on p. 5).

Kraushaar, James M. and David C. Novak (2010). "Examining the Affects of Student Multitasking with Laptops during the Lecture." In: *Journal of Information Systems Education* 21.2, pp. 241–251 (cit. on p. 28).

Kwak, Haewoon, Changhyun Lee, Hosung Park, and Sue Moon (2010). "What is Twitter, a Social Network or a News Media?" In: *Proceedings of the 19th International Conference on World Wide Web.* WWW '10. New York, NY, USA: ACM, pp. 591–600 (cit. on pp. 26, 105).

Lambertz, Kathrin (2011). "Back-Channelling: The Use of Yeah and Mm to Portray Engaged Listenership." In: *Griffith Working Papers in Pragmatics and Intercultural Communication* 4, pp. 11–18 (cit. on pp. 15, 17).

Lambropoulos, Niki, Xristine Faulkner, and Fintan Culwin (2012). "Supporting Social Awareness in Collaborative E-Learning." In: *British Journal of Educational Technology* 43.2, pp. 295–306 (cit. on p. 13).

Lampos, V. and N. Cristianini (2010). "Tracking the Flu Pandemic by Monitoring the Social Web." In: *Cognitive Information Processing (CIP), 2010 2nd International Workshop on*, pp. 411–416 (cit. on p. 26).

Langville, Amy N. and Carl D. Meyer (2012). *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton and Oxford: Princeton University Press (cit. on pp. 107, 117).

Lantz, Michael E. (2010). "The Use of 'Clickers' in the Classroom: Teaching Innovation or merely an Amusing Novelty?" In: *Computers in Human Behavior* 26.4, pp. 556–561 (cit. on pp. 18, 83).

Lempel, R. and S. Moran (2001). "SALSA: The Stochastic Approach for Link Structure Analysis." In: *Proceedings of the 22nd International Conference on World Wide Web* 19.2, pp. 131–160 (cit. on p. 107).

Lerman, Kristina (2009). "Advances in Web-Mining and Web Usage Analysis." In: *Advances in Web Mining and Web Usage Analysis*. Berlin, Heidelberg: Springer, pp. 77–96 (cit. on pp. 90 sq.).

Marmolowski, Mateusz (2008). *Real-Life Rating Algorithm*. Technical Report. DERI (cit. on pp. 90, 93).

Marshall, Catherine C. (1997). "Annotation: From Paper Books to the Digital Library." In: *Proceedings of the second ACM international conference on Digital libraries*. DL '97. New York, NY, USA: ACM, pp. 131–140 (cit. on p. 42).

Mazur, Eric (1997). *Peer Instruction – A User's Manual*. Upper Saddle River, New Jersey, USA: Prentice Hall (cit. on p. 22).

— (2009). "Farewell, Lecture?" In: *Science* 323.5910, pp. 50–51 (cit. on p. 27).

McCarthy, Dennis and Chris Crane (2008). *Comet and Reverse Ajax: The Next-Generation Ajax 2.0*. English. 2009 edition. Berkeley, CA; New York, NY: Apress (cit. on p. 54).

Merton, Robert K. (1968). "The Matthew Effect in Science." In: *Science* 159.3810, pp. 56–63 (cit. on p. 122).

Miyazoe, Terumi and Terry Anderson (2011). "Anonymity in Blended Learning: Who Would You Like to Be?" In: *Educational Technology and Society* 14.2, pp. 175–187 (cit. on p. 14).

Nassaji, H. and G. Wells (2000). "What's the Use of 'Triadic Dialogue'?: An Investigation of Teacher-Student Interaction." In: *Applied Linguistics* 21.3, pp. 376–406 (cit. on p. 5).

Nichols, David M. (1998). "Implicit Rating and Filtering." In: *Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering*, pp. 31–36 (cit. on p. 91).

Nokelainen, Petri, Jaakko Kurhila, Miikka Miettinen, Patrik Floréen, and Henry Tirri (2003). "Evaluating the Role of a Shared Document-Based Annotation Tool in Learner-Centered Collaborative Learning." In: *Proceedings of the 3rd International*

Conference on Advanced Learning Technologies. IEEE Computer Society Press, pp. 200–203 (cit. on p. 42).

Nokelainen, Petri, Miikka Miettinen, Jaakko Kurhila, Patrik Floréen, and Henry Tirri (2005). "A Shared Document-Based Annotation Tool to Support Learner-Centred Collaborative Learning." In: *British Journal of Educational Technology* 36.5, pp. 757–770 (cit. on p. 42).

Oblinger, Diana G. (2004). "The Next Generation of Educational Engagement." In: *Journal of Interactive Media in Education* 2004.1, p. 10 (cit. on p. 17).

O'Donnell, Angela and Donald Dansereau (1992). "Scripted Cooperation in Student Dyads: A Method for Analyzing and Enhancing Academic Learning and Performance." In: *Interactions in Cooperative Groups. The Theoretical Anatomy of Group Learning.* Ed. by Rachel Hertz-Lazarowitz and Norman Miller. New York: Cambridge University Press, pp. 120–141 (cit. on p. 17).

OECD (2007). *Participative Web and User-Created Content: Web 2.0, Wikis, and Social Networking.* English. Tech. rep. Paris, France: Organization for Economic Cooperation and Development (cit. on p. 22).

O'Reilly, Tim (2005). *What is Web 2.0? Design Patterns and Business Models for the Next Generation Software* (cit. on p. 22).

Ostrom, Elinor (1998). "A Behavioral Approach to the Rational Choice Theory of Collective Action." In: *The American Political Science Review* 92.1, p. 1 (cit. on p. 105).

Palincsar, A. S. and A. L. Brown (1984). "Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities." In: *Cognition and Instruction* 1.2, pp. 117–175 (cit. on p. 17).

Paulsen, Erik, Edvin Bru, and Terje A. Murberg (2006). "Passive Students in Junior High School: The Associations with Shyness, Perceived Competence and Social Support." In: *Social Psychology of Education* 9.1, pp. 67–81 (cit. on p. 4).

Pavlou, Paul and Sulin Ba (2000). "Does Online Reputation Matter? An Empirical Investigation of Reputation and Trust in Online Auction Markets." In: *AMCIS 2000 Proceedings* (cit. on p. 104).

Pearson, Judy and Richard West (1990). "An Initial Investigation of Gender on Student Questions in the Classroom: Developing a Descriptive Base." In: Chicago, USA, p. 32 (cit. on p. 17).

Peters, Dorian (2014). *Interface Design for Learning - Design Strategies for Learning Experiences.* New Riders (cit. on pp. 43, 46 sq., 83).

Pohl, Alexander, François Bry, Jeannette Schwarz, and Marlene Gottstein (2012). "Sensing the Classroom: Improving Awareness and Self-Awareness of Students with Backstage." In: *Proceedings of the International Conference on Interactive and Collaborative Learning.* Villach, Austria, pp. 1–8 (cit. on pp. 9, 41).

Pohl, Alexander, Vera Gehlen-Baum, and François Bry (2012). "Enhancing the Digital Backchannel Backstage on the Basis of a Formative User Study." In: *International*

Journal of Emerging Technologies in Learning (iJet) 7.1. Extended Version of "Assessing Backstage - A Backchannel for Collaborative Learning in Large Classes", pp. 33–41 (cit. on pp. 29, 128).

Rabiner, Lawrence R. (1989). "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition." In: *Proceedings of the IEEE* 77.2, pp. 257–286 (cit. on p. 129).

Ratto, M., R. B. Shapiro, M. T. Truong, and W. G. Griswold (2003). "The ActiveClass Project: Experiments in Encouraging Classroom Participation." In: *Proceedings of the Fifth International Conference on Computer Supported Collaborative Learning*. Bergen, Norway (cit. on p. 20).

Redecker, C., K. Ala-Mutka, and Y. Punie (2010). "Learning 2.0 – The Impact of Social Media on Learning in Europe." In: *JRC European Commission* (cit. on p. 5).

Reid, Fraser J. M., Vlastimil Malinek, Clifford J. T. Stott, and Jonathan ST. B. T. Evans (1996). "The Messaging Threshold in Computer-Mediated Communication." In: *Ergonomics* 39.8, pp. 1017–1037 (cit. on p. 47).

Reinhardt, Wolfgang, Michael Sievers, Johannes Magenheim, Dennis Kundisch, Philipp Herrmann, Marc Beutner, and Andrea Zoyke (2012). "PINGO: Peer Instruction for Very Large Groups." In: *21st Century Learning for 21st Century Skills*. Ed. by Andrew Ravenscroft, Stefanie Lindstaedt, Carlos Delgado Kloos, and Davinia Hernández-Leo. Vol. 7563. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 507–512 (cit. on p. 22).

Resnick, Paul, Ko Kuwabara, Richard Zeckhauser, and Eric Friedman (2000). "Reputation Systems." In: *Communications of the ACM* 43.12, pp. 45–48 (cit. on p. 104).

Richardson, John T. E., Lesley Dawson, Gaynor Sadlo, Virginia Jenkings, and Janet McInnes (2007). "Perceived Academic Quality and Approaches to Studying in the Health Professions." In: *Medical Teaching* 29.5, e108–e116 (cit. on p. 58).

Rittenbruch, Markus, Tim Mansfield, and Stephen Viller (2009). "Design and Evaluation of Intentionally Enriched Awareness." In: *Awareness Systems*. Ed. by Panos Markopoulos, Boris De Ruyter, and Wendy Mackay. Human-Computer Interaction Series. Springer London, pp. 367–395 (cit. on p. 12).

Rittenbruch, Markus and Gregor McEwan (2009). "An Historical Reflection of Awareness in Collaboration." In: *Awareness Systems*. Ed. by Panos Markopoulos, Boris De Ruyter, and Wendy Mackay. Human-Computer Interaction Series. Springer London, pp. 3–48 (cit. on p. 12).

Rocca, Kelly A. (2010). "Student Participation in the College Classroom: An Extended Multidisciplinary Literature Review." In: *Communication Education* 59.2, pp. 185–213 (cit. on p. 9).

Rosenshine, Barak and Carla Meister (1994). "Reciprocal Teaching: A Review of Research." In: *Review of Educational Research* 64.4, pp. 479–530 (cit. on p. 16).

Ross, C., M. Terras, C. Warwick, and A. Welsh (2011). "Enabled Backchannel: Conference Twitter Use by Digital Humanists." In: *Journal of Documentation* 67.2, pp. 214–237 (cit. on p. 30).

Rowe, Mary Budd (1986). "Wait Time: Slowing Down May be a Way of Speeding Up!" In: *Journal of Teacher Education* 37.43, pp. 43–51 (cit. on pp. 12, 128).

Sachs, Lothar and Jürgen Hedderich (2009). *Angewandte Statistik: Methodensammlung mit R*. In German. 13th edition. Berlin: Springer (cit. on pp. 94, 96).

Sadalage, Pramodkumar J. and Martin Fowler (2012). *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence*. Upper Saddle River, NJ: Addison Wesley (cit. on p. 58).

Saunders, Neil, Pedro Beltrão, Lars Jensen, Daniel Jurczak, Roland Krause, Michael Kuhn, and Shirley Wu (2009). "Microblogging the ISMB: A New Approach to Conference Reporting." In: *PLoS Computational Biology* 5.1, e1000263 (cit. on p. 30).

Schank, R., Robert Abelson, and Roger C. Schank (1977). *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures*. Englisch. Hillsdale, N.J.; New York: Lawrence Erlbaum Assoc Inc (cit. on p. 47).

Schickinger, Thomas and Angelika Steger (2002). *Diskrete Strukturen - Band 2: Wahrscheinlichkeitstheorie und Statistik*. In German. Berlin, Heidelberg Germany: Springer (cit. on pp. 109, 115, 118).

Schmidt, Jan-Hinrik (2013). *Social Media (Medienwissen kompakt)*. In German. Berlin, Heidelberg, Germany: Springer VS (cit. on p. 84).

Schmidt, Kjeld (2002). "The Problem with 'Awareness': Introductory Remarks on 'Awareness in CSCW'." In: *Comput. Supported Coop. Work* 11.3, pp. 285–298 (cit. on p. 13).

Schworm, Silke and Frank Fischer (2006). "Academic Help Seeking." In German. In: *Handbuch Lernstrategien*. Göttingen: Hogrefe Verlag, pp. 282–297 (cit. on p. 4).

Al-Sharrah, Ghanima (2010). "Ranking Using the Copeland Score: A Comparison with the Hasse Diagram." In: *Journal of Chemical Information and Modeling* 50.5, pp. 785–791 (cit. on pp. 92 sq.).

Simon, Beth, Krista Davis, William G. Griswold, Michael Kelly, and Roshni Malani (2008). "Noteblogging: Taking Note Taking Public." In: *Proceedings of the 39th SIGCSE Technical Symposium on Computer Science Education*. Portland, OR, USA: ACM, pp. 417–421 (cit. on p. 20).

Smith, Peter (2013). *Professional Website Performance*. Indianapolis, Indiana, USA: John Wiley & Sons Inc. (cit. on p. 56).

Sommer, Robert (1961). "Leadership and Group Geography." In: *Sociometry* 24.1, pp. 99–110 (cit. on p. 5).

Sprenger, Timm (2011). "TweetTrader.net: Leveraging Crowd Wisdom in a Stock Microblogging Forum." In: *Proceedings of the 5th International Conference of Weblogs and Social Media*. Barcelona, Spain (cit. on p. 101).

Stuart, J. and R. J. D. Rutherford (1978). "Medical Student Concentration During Lectures." In: *The Lancet* (cit. on pp. 4, 130).

Surowiecki, James (2005). *The Wisdom of Crowds*. Reprint edition. New York: Anchor (cit. on p. 104).

Suthers, D. (2007). "Roles of Computational Scripts." In: *Scripting Computer-Supported Collaborative Learning*. Ed. by F. Fischer, I. Kollar, H. Mandl, and Jörg M. Haake, pp. 177–187 (cit. on p. 48).

Suthers, Daniel and Jun Xu (2002). "Kukakuka: An Online Environment for Artifact-Centered Discourse." In: *Education Track of the 11th World Wide Web Conference (WWW 2002)*. Hawaii, pp. 7–11 (cit. on p. 42).

Sutton, J., L. Palen, and I. Shklovski (2008). "Backchannels on the Front Lines: Emergent Uses of Social Media in the 2007 Southern California Wildfires." In: Washington D.C., USA (cit. on p. 26).

Tess, Paul A. (2013). "The Role of Social Media in Higher Education Classes (Real and Virtual) – A Literature Review." In: *Computers in Human Behavior* 29.5, A60–A68 (cit. on pp. 5, 27).

Thaler, Richard H. and Cass R. Sunstein (2003). "Libertarian Paternalism." In: *The American Economic Review* 93.2, pp. 175–179 (cit. on p. 104).

Tullis, Tom and Bill Albert (2008). *Measuring the User Experience*. Burlington MA, USA: Morgan Kaufmann (cit. on pp. 33, 38).

Valtonen, T., S. Havu-Nuutinen, P. Dillon, and M. Vesisenaho (2011). "Facilitating Collaboration in Lecture-Based Learning Through Shared Notes Using Wireless Technologies." In: *Journal of Computer Assisted Learning* 27.6, pp. 575–586 (cit. on p. 16).

Vassileva, Julita (2012). "Motivating Participation in Social Computing Applications: A User Modeling Perspective." In: *User Modeling and User-Adapted Interaction* 22.1-2, pp. 177–201 (cit. on pp. 84, 91, 103 sq., 122, 128).

Verner, Coolie and Gary Dickinson (1967). "The Lecture, An Analysis and Review of Research." In: *Adult Education Quarterly* 17.2, pp. 85–100 (cit. on p. 4).

Wallis, Sean A. (2013). "Binomial Confidence Intervals and Contingency Tests: Mathematical Fundamentals and the Evaluation of Alternative Methods." In: *Quantitative Linguistics* 20.3, pp. 178–208 (cit. on p. 96).

Wang, G. Alan, Jian Jiao, Alan S. Abrahams, Weiguo Fan, and Zhongju Zhang (2013). "ExpertRank: A Topic-Aware Expert Finding Algorithm for Online Knowledge Communities." In: *Decision Support Systems* 54.3, pp. 1442–1451 (cit. on p. 108).

Wasserman, Stanley and Katherine Faust (1994). *Social Network Analysis: Methods and Applications.* New York, NY, USA: Cambridge University Press (cit. on p. 105).

Weiand, Klara, Fabian Kneißl, Tim Furche, and François Bry (2012). "PEST: Fast Approximate Keyword Search in Semantic Data using Eigenvector-based Term Propagation." In: *Information Systems* 37.4, pp. 372–290 (cit. on p. 108).

Weinberger, Armin (2011). "Principles of Transactive Computer-Supported Collaboration Scripts." In: *Nordic Journal of Digital Literacy* 6.3, pp. 189–202 (cit. on pp. 48, 85, 129).

Weinberger, Armin, Frank Fischer, and Karsten Stegmann (2005). "Computer-Supported Collaborative Learning in Higher Education: Scripts for Argumentative Knowledge Construction in Distributed Groups." In: *Proceedings of the 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: The Next 10 Years!* CSCL '05. Taipei, Taiwan, pp. 717–726 (cit. on p. 48).

Wessels, Anja, Stefan Fries, Holger Horz, Nicolai Scheele, and Wolfgang Effelsberg (2007). "Interactive Lectures: Effective Teaching and Learning in Lectures using Wireless Networks." In: *Computers in Human Behavior* 23.5, pp. 2524–2537 (cit. on p. 21).

Wilkerson, Michelle, William G. Griswold, and Beth Simon (2005). "Ubiquitous Presenter: Increasing Student Access and Control in a Digital Lecturing Environment." In: *Proceedings of the 36th SIGCSE Technical Symposium on Computer Science Education.* St. Louis, Missouri, USA: ACM, pp. 116–120 (cit. on p. 20).

Wilson Edwin, B. (1927). "Probable Inference, the Law of Succession, and Statistical Inference." In: *Journal of the American Statistical Association* 22.158, pp. 209–212 (cit. on p. 95).

Yardi, Sarita (2006). "The Role of the Backchannel in Collaborative Learning Environments." In: *Proceedings of the 7th International Conference on Learning Sciences.* ICLS '06. Bloomington, Indiana: International Society of the Learning Sciences, pp. 852–858 (cit. on p. 6).

— (2008). "Whispers in the Classroom." In: *The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning.* Ed. by Tara McPherson. Cambridge, MA: MIT Press, pp. 143–164 (cit. on pp. 17 sq.).

Young, Mark S., Stephanie Robinson, and Phil Alberts (2009). "Students Pay Attention!: Combating the Vigilance Decrement to Improve Learning During Lectures." In: *Active Learning in Higher Education* 10.1, pp. 41–55 (cit. on pp. 4, 19, 83, 130).

Zhou, Ming, Martin Dresner, and Robert J. Windle (2008). "Online Reputation Systems: Design and Strategic Practices." In: *Decision Support Systems* 44.4, pp. 785–797 (cit. on p. 104).

APPENDIX

APPENDIX

A.1 List of Acronyms

Ajax	Asynchronous Javascript and XML
ARS	Audience Response System
BKT	Bayesian Knowledge Tracing
CCS	Classroom Communication System
CMC	Computer-Mediated Communication
CSCL	Computer-Supported Collaborative Learning
EURIBOR	Euro Interbank Offered Rate
GPR	Grace Period Reward
HMM	Hidden Markov Model
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
JEE	Java Enterprise Edition
JSON	Javascript Object Notation
IRE/F	Initiation-Response-Evaluation/Feedback Interaction Pattern
LIBOR	London Interbank Offered Rate
LMU	Ludwig-Maximilian University of Munich
ML	Maximum Likelihood
MOOC	Massive Open Online Course
PDA	Personal Digital Assistant
Q&A	Question and Answer
REST	Representational State Transfer
RDF	Resource Description Framework

RFC	Request for Comments
XHR	XML HTTP Request

A.2 List of Course Abbreviations

L "Logic and Discrete Mathematics" at LMU (Summer Term 2014)

P "Introduction to Programming and Modeling" at LMU (Summer Term 2013)

P1 "Programming 1" at Saarland University (Winter Term 2014/2015)

W "Web Information Systems" at LMU (Winter Term 2013/2014)

A.3 List of Symbols used Throughout this Thesis

M	Mean
MAD	Median Absolute Deviation
Mdn	Median
Min	Minimum Number
Max	Maximum Number
N	Sample Size / Number of Participants
SD	Standard Deviation
$\mathcal{N}(\mu, \sigma)$	Normal Distribution with Mean μ and Standard Deviation σ

A.4 List of Symbols used in Chapter 5

α	Cronbach's α, Reliability Index for Survey Items
κ	Cohen's κ, Inter-Rater Reliability Index
N_Q	Number of Backchannel Posts of Category *Question*
N_A	Number of Backchannel Posts of Category *Answer*
N_R	Number of Backchannel Posts of Category *Remark*
N_F	Number of Backchannel Posts of Category *Too Fast*
N_S	Number of Backchannel Posts of Category *Too Slow*
N_{co}	Number of Backchannel Posts Coded as *content-oriented*
N_i	Number of Backchannel Posts Coded as *independent*
N_o	Number of Backchannel Posts Coded as *organizational*
N_{po}	Number of Backchannel Posts Coded as *process-oriented*
N_{pe}	Number of Backchannel Posts Coded as *participation-enabling*
N_{pos}	Number of Positive Ratings
N_{neg}	Number of Negative Ratings
N_{off}	Number of Off-Topic Ratings
N_{Rel}	Number of Relevant Backchannel Posts
N_{Irrel}	Number of Irrelevant Backchannel Posts
N_{stud}	Number of Students

A.5 List of Symbols used in Chapter 6

$r_b(x)$	Rating Score for an Item x Based on Bayesian Average
$r_w(x)$	Rating Score for an Item x Based on Wilson's Confidence Interval
$C_i(\alpha, p_i)$	Confidence Interval for Parameter p_i and Confidence Level α

A.6 List of Symbols used in Chapter 7

$\mathbb{1}$	Diagonal Matrix of Ones
a_s	Coefficient that Measures the Rating Activity of a Student s
$\mathbf{A}$	Rating Activity Coefficient Matrix
$b_{s \to i}$	Backchannel Praise Probability (Probability of a Student s to Praise Student i on the Backchannel)
$\mathbf{B}$	Backchannel Praise Probability Matrix
$\mathbf{I}$	Uniform Praise Probability Matrix
$\mathbf{b}_s$	Backchannel Praise Probability Distribution of a Student s
$p_{s \to i}$	Praise Probability (Probability of a Student s to Praise Student i)
$\mathbf{p}_s$	Praise Probability Distribution of a Student s
$q_{s \to j}^{\alpha}$	Quiz Praise Probability (Probability of a Student s to Praise Student i at Correctness Threshold α)
$\mathbf{Q}_\alpha$	Quiz Praise Probability Matrix
$r_{s \to i}$	Estimated Praise Probability (Estimated Probability of a Student s to Praise Student i)
$\mathbf{r}_s$	Estimated Praise Probability Distribution of a Student s
$\mathbf{R}$	Estimated Praise Probability Matrix
$\mathbf{R}_\alpha$	Estimated Praise Probability Matrix Extended by Quiz Performance
$\mathcal{S}$	Set of Students

A.7 Survey of the Formative Usability Study in December 2010

Fragebogen "Backchanneling"

Alter: _____

Männlich: □   **Weiblich:** □

Semester: _____

Benutzerkennung: _____

Fachrichtung: _____

_____

Folgendes würde zum Verständnis des Vortrags beitragen:

Besonders gut gefallen hat mir:

Vorschläge, Anregungen, sonstige Anmerkungen:

	trifft nicht zu				trifft zu
Der Vortrag regte mich dazu an, Fragen zu stellen					
Der Vortrag regte mich dazu an, auf Fragen zu antworten					
Durch die gestellten Fragen erhöhte sich mein Verständnis des Vortrags					
Meine Bereitschaft während des Vortrags mit anderen zu kommunizieren ist hoch					
Ich stelle gerne Fragen während Vorträgen					
Ich würde dem Dozenten gerne während des Vortrags Feedback geben					
Es fällt mir leicht während Vorträgen aufmerksam zuzuhören					
Ich würde mich gerne bei Vorträgen aktiver einbringen					
Ich würde dem Dozenten gerne mitteilen, wenn mir etwas zu schnell geht					
Ich beantworte gerne Fragen von anderen					
Das Fragenstellen war schwierig					
Ich war mir oft unsicher, wie ich Fragen richtig beantworten soll					
Ich hätte gerne während des Vortrags mit anderen diskutiert					
Mein Verständnis vom Vortrag wäre durch mehr Fragen positiv beeinflusst worden					
Ich würde gerne in einer Vorlesung mit anderen zu kommunizieren					
Ich fühlte mich sicher beim Beantworten von Fragen					
Ich hatte hohe Unsicherheit beim Stellen von Fragen					
Ich wüsste gerne, welche Fragen die Mithörer gestellt haben					
Auf Fragen der Mithörer wusste ich nicht, wie ich antworten soll					
Die Beantwortung der Fragen war wichtig für mein Verständnis					
Ich konnte mir ein Bild über meine Mithörer und deren Verständnis zum Vortrag machen					

A.8 Course Outline "Introduction to Programming and Modeling" at LMU in the Summer Term 2013

Identifier	Title	Keywords
P-01	Introduction to SML	First steps, expressions, values, types and polymorphic type checking, precedence and associativity
P-02	Introduction to SML, Evaluation, Substitution Model	names, bindings, recursion, declarations, conditions, re-declarations, static and dynamic binding, evaluation order, substitution
P-03	Functions	Procedure vs. functions, local declarations, block structure and shadowing, mutual recursion, recursion and processes
P-04	Predefined Data Types	Base types, composite types, lists, functions on lists
P-05	Type Checking	Definition of type checking, polymorphism, type variables, type constants, type constructors and type expressions, type inference
P-06	Higher-Order Functions	Functions as arguments and return values of functions, currying, function composition, fundamental higher-order functions
P-07	Custom Data Types	Type alias, datatype declarations, recursive types, fundamental functions on binary trees
P-08	Pattern Matching	Purpose, principles, programming with pattern matching, linearity constraint for patterns, pattern matching algorithm
P-09	Exceptions	Errors as values, exceptions vs. values, exception handling, evaluation of raise and handle, criticism of the exception approach
P-10	Abstract Data Types and Modules	Information hiding, abstract types for sets, modules
P-11	Imperative Programming in SML	References, iteration, arrays, basic IO

A.9 Course Outline "Web Information Systems" at LMU in the Winter Term 2013/14

Identifier	Title	Keywords
W-01	HTML and Information Retrieval	Markup, HTML, XHTML
W-02	HTML and Information Retrieval	Hypertext Models, CSS, Basic Information Retrieval
W-03	Search Engines and Network Analysis	Components of a search engine, PageRank
W-04	Search Engines and Network Analysis	HITS, Keyword Query Languages
W-05	BigTable, MapReduce	BigTable, MapReduce, PageRank with MapReduce
W-06	XML, Data Models, DTD, XMLS	Introduction to XML
W-07	XML, Data Models, DTD, XMLS	Data Models (relational *vs.* semistructured), DTD, XMLS
W-08	XML, Data Models, DTD, XMLS	XMLS
W-09	Query Languages: XPath and XQuery	Autocompletion Search
W-10	Query Languages: XPath and XQuery	XPath, XQuery
W-11	Query Languages: XPath and XQuery	Classification of XPath Queries, Navigational XPath, Evaluating XPath Queries, Structural Joins, Relational Evaluation of XPath
W-12	RDF/S	Motivation, Concepts of RDF, RDF Schema
W-13	RDF/S	SPARQL

A.10 Course Outline "Logics and Discrete Structures" at LMU in the Summer Term 2014

Identifier	Title	Keywords
L-01	Introduction	Logics as formalizations of rational thought, logics vs. classical logics, necessity of a formal language, deduction as computation
L-02	Propositional Logics pt. I	Enumerable and countable sets, syntax, structural induction
L-03	Propositional Logics pt. II	Uniqueness Theorem, semantics, Replacement Theorem
L-04	Predicate Logics pt. I	syntax, semantics
L-05	Predicate Logics pt. II	normal forms
L-06	Proof Methods	Resolution, Natural Deduction
L-07	1st Mock Exam	60 minutes examination time, presentation of sample solutions, self-grading
L-08	Natural and Integer Numbers	Natural Numbers, Integer Numbers, Prime Numbers
L-09	Modular Arithmetics	Congruence relations, Euclidean Division, Euclidean Algorithm, Chinese Remainder Theorem, Fermat's Little Theorem
L-10	Combinatorics	Drawing elements from a set, fundamental formulas
L-11	2nd Mock Exam	60 minutes examination time, presentation of sample solutions, self-grading

A.11 Course Outline "Programming 1" at Saarland University in the Summer Term 2014

Identifier	Title
S-01	Introduction, Crash Course pt. 1
S-02	Crash Course pt. 2
S-03	Crash Course pt. 3, Phrases, Phases of Interpretation
S-04	Semantic Admissibility, Execution
S-05	Higher-Order Procedures pt. 1
S-06	Higher-order Procedures pt. 2
S-07	Lists and Strings pt. 1
S-08	Lists and Strings pt. 2
S-09	Strings, Sorting
S-10	Sorting pt. 2, Sets, Constructors
S-11	Exceptions, Constructor Types
S-12	Trees
S-13	Mock Exam
S-14	Trees pt. 2
S-15	Trees pt. 3
S-16	Set Theory

157

Identifier	Title
S-17	Set Theory pt. 2
S-18	Mathematical Procedures
S-19	Mathematical Procedures pt. 2
S-20	Inductive Correctness Proofs
S-21	Inductive Correctness Proofs
S-22	Inductive Correctness Proofs, Time Complexity of Recursive Procedures
S-23	Time Complexity of Recursive Procedures
S-24	Syntax and Semantics
S-25	Interpreter for "F": Parsing
S-26	Interpreter for "F": Lexer+Parser for Arithmetic Expressions; Data Structures: Structures and Signatures, Vectors, Searching
S-27	Data Structures: Queues, Functors; Memory and Mutable Objects
S-28	Memory and Mutable Objects
S-29	Memory and Mutable Objects pt. 3; Stack Machines and Interpreters
S-31	Stack Machines

A.12 Original Student Responses to Open Items

A.12.1 Responses in Course P mentioned in Section 5.3.6.1

„... Fragen und Kommentare zur aktuellen Folie stellen zu können."

„... es motiviert, in der Vorlesung 'dabei' zu bleiben."

„... es die Quizfragen gab. Die Vorlesung wurde durch Backstage interaktiver."

„... die Fragen der Komilitonen zu eigenen Verständnis beigetragen haben und das Quiz."

„... Es gestaltet die Vorlesung um einiges interaktiver und es fällt einem leichter Fragen zur Vorlesung zu stellen, welche man sonst wohl nicht gestellt hätte."

„... man die Folien zur Vorlesung sieht und oft jemand Kommentare und Fragen reinschreibt. Die Autonavigation immer zur aktuellen Folie springt."

„... die Folien kurz und übersichtlich gestaltet sind und das Verstandene durch die Quizfrage gefestigt wird."

„Hatte am Anfang der Vorlesungszeit nie geglaubt wie toll Backstage ist. Großes Lob und weiter so!"

„Eine super Sache, die in keinem Fach fehlen sollte!"

A.12.2 Responses in Course W mentioned in Section 5.3.6.2

„[...] ich kann sehr gut durch Zuhören lernen, aber wenn ich den Redner nicht sehe (weil ein digitales Gerät da ist) dann verpass ich viel mehr. ich glaube für den Einsatz von Backstage sind Erstsemester-Vorlesungen besser geeignet, weil viel mehr Studenten anwesend sind und es bei der Menge an zuhörern auch schwerer ist Fragen zu stellen bzw. eher auf Fragen verzichtet wird, weil der ein oder andere sich nicht traut. Da halte ich Backstage eher für sinnvoll, als in einer 5. Semester Vorlesung in der 30-40 Leute anwesend sind. Das ist Seminargröße und nicht die richtige Umgebung."

A.12.3 Responses in Course L mentioned in Section 5.3.6.3

„... die Teilnahme an der Vorlesung aktiv gefördert wurde."

„– Wenn man mal nicht mitkommt zurückblättern kann. – Bei den Quizes mal wieder wach wird falls man mal nicht aufpasst. – Von Zuhause die VL auch zu einem gewissen Grad miterleben kann."

„... Die Interaktivität gefördert wurde und ein Schritt entgegen der Standard Vorlesung (Dozent redet) enstanden ist."

„... Man (anonym) Fragen stellen und beantworten kann."

„... Fragen der Studenten schriftlich fixiert waren und den Vorbereitungsprozess auf die Klausur unterstützt haben."

„... man selbst mit einem Klick eine folie zurückgehen konnte, falls mal etwas zu schnell ging; dort die Probeklausuren abgehalten wurden; man die Quizzes auch vom Smartphone aus beantworten konnte."

„Ich möchte mich erstmal für das Programm bedanken. Trotz einiger Kritik, die wahrscheinlich aufkommen wird, verändert die Anwendung den Vorlesungsalltag jedoch eindeutig in eine positive Richtung. Ganz reibungslos hat es in der Vorlesung noch nicht funktioniert, aber es hat als jetzige Beta-Version noch einiges an Potential und es würde mich freuen, wenn mich Backstage in den nächsten Semestern in der einen oder anderen Vorlesung wieder begleitet."

„auf jeden Fall positiven Effekt für die vorlesung, hat mir sehr gefallen."

A.12.4   Responses in Course P1 mentioned in Section 5.3.6.4

„Insgesamt eine schöne Spielerei, die nicht wirklich nützlich für Vorlesungen ist."

„Ist leider noch ein bisschen buggy, aber tolle Idee."

„Interessantes System, verlangt anständiges Verhalten der Studenten sonst lenkt es eher ab als zu helfen."

„Gute Ergänzung!"

„Gutes System für Erstsemesterveranstaltungen, wenn sehr viele Studenten im Hörsaal sind und somit die Kommunikation zwischen Studenten und Dozenten sehr schwierig ist, jedoch nach wie vor Verbesserungswürdig."

„Fand das ganze System wirklich gut! Macht die Vorlesung gleich mehr Spaß! :)"

„Gefällt mir sehr gut! Praktisch fand ich es, dass die Seite auch einwandfrei mit Kommentarleiste auf dem Tablet funktionieren, dann muss man nicht immer ein Laptop mit dabei haben! [...]"

„Funktionierte schon erstaunlich gut allerdings hab ich es später nur noch genutzt um die Quizfragen zu beantworten und damit die Folien automatisch auf dem Laptop mitliefen.“

„Sehr hilfreich, aufgrund der vorhandenen Nutzung von Dozenten, Tutoren und Komillitonen. Auf aufkommende Fragen wurde geantwortet und somit der Vorlesungsstoff verständlicher. Quizfragen waren genügend vorhanden und nicht zuviele, sodass die Vorlesung gestört würde. Vorallem die Erklärung der korrekten Lösung und warum andere falsch waren, beseitigten schnell mögliche verbleibende Fragen. Allgemein also sehr nützlich und auch in weiteren Vorlesungen erwünschenswert.“

A.13   Surveys in the Courses P, W, L, and P1

Nutzung von Backstage in der Vorlesung "Einführung in die Programmierung und Modellierung im Sommersemester 2013"

Dieser Fragebogen soll uns Einblicke gewähren,

- wie Du die Funktionalitäten von Backstage einschätzt,
- wie und für was Du Backstage verwendet hast und
- welche Verbesserungen Du für Backstage siehst.

Bitte investiere die wenigen Minuten, die nötig sind, um den Fragebogen vollständig zu beantworten. Vielen Dank!

There are 18 questions in this survey

Fragebogen

1 [Block1Likert]Funktionalitäten von Backstage *

Please choose the appropriate response for each item:

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Mir gefiel, in der Vorlesung öffentliche Kommentare mit Backstage erstellen zu können	○	○	○	○	○	○
Mir gefiel, in der Vorlesung anonyme Kommentare mit Backstage erstellen zu können	○	○	○	○	○	○
Mir gefiel, in der Vorlesung eigene Notizen mit Backstage erstellen zu können	○	○	○	○	○	○
Mir gefiel, Nachrichten auf Backstage bewerten zu können	○	○	○	○	○	○
Ich fand die Länge der Nachrichten von 140 Zeichen ausreichend	○	○	○	○	○	○
Die Bewertungen der Nachrichten war geeignet, um relevante Nachrichten zu erkennen	○	○	○	○	○	○
Es macht mir Spaß Nachrichten in Backstage zu bewerten	○	○	○	○	○	○
Es macht mir Spaß die Quizfragen in Backstage zu beantworten	○	○	○	○	○	○
Die Quizfragen sind ein geeignetes Mittel, Aktivität in der Vorlesung zu fördern	○	○	○	○	○	○
Es hat mir Spaß gemacht, die Vorlesung zu besuchen und mit Backstage wurde die Vorlesung für mich	○	○	○	○	○	○

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
angenehmer Ich finde es gut, dass ich auf der Übersichtsseite Informationen über mein Abschneiden in den Quizfragen bekomme	○	○	○	○	○	○
Ich finde es gut, dass ich auf der Übersichtsseite Informationen über das Abschneiden meiner Kommilitonen in den Quizfragen bekomme	○	○	○	○	○	○
Mir würde es ausreichen wenn die Quizfragen mündlich gestellt und per Handmeldung beantwortet werden	○	○	○	○	○	○

2 [Block1OI1]Mir hat besonders gut an Backstage gefallen, dass …

Please write your answer here:

3 [Block1OI2]

Der Einsatz von Backstage wäre für die folgenden Vorlesungen wünschenswert:

Please write your answer here:

4 [Block1OI3]Ich würde mir wünschen, dass zukünftige Versionen von Backstage noch folgende Funktionen für die Nutzung während der Vorlesung bereitstellen

Please write your answer here:

7 [Block2Likert]Nutzung von Backstage *

Please choose the appropriate response for each item:

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Ich habe Backstage genutzt, um zu erfahren welche Fragen meine Kommilitonen haben	○	○	○	○	○	○
Ich habe Backstage genutzt, um die Antworten meiner Kommilitonen zu lesen	○	○	○	○	○	○
Ich habe Backstage genutzt, um die Antworten der Tutoren und des Dozenten zu lesen	○	○	○	○	○	○
Ich habe am Austausch von Nachrichten in Backstage gerne teilgenommen	○	○	○	○	○	○
Backstage ist für die Nachbereitung des Stoffs für die Vorlesungssitzungen nützlich gewesen	○	○	○	○	○	○
Backstage ist für die Nachbereitung des Stoffs für die Übungen nützlich gewesen	○	○	○	○	○	○
Backstage wird mir für die die Nachbereitung des Stoffs für die Klausur hilfreich sein	○	○	○	○	○	○
Mithilfe von Backstage habe ich viele Inhalte bereits in der Vorlesung verstanden	○	○	○	○	○	○
Backstage hat mich während der Vorlesung abgelenkt	○	○	○	○	○	○
Die Möglichkeit, Notizen auf Backstage zu erstellen	○	○	○	○	○	○

5 [Block1OI4]Ich würde mir wünschen, dass zukünftige Versionen von Backstage noch folgende Funktionen für die Nutzung für die Nachbereitung bereitstellen:

Please write your answer here:

6 [Block1OI5]Auf folgende Funktionen von Backstage könnte ich während der Vorlesung verzichten:

Please write your answer here:

9 [Block2OI2]

Ich habe Backstage für die Nachbereitung genutzt, weil...

*

Please choose **only one** of the following:

○ Ja
○ Nein

Make a comment on your choice here:

| |
| |

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
unterscheidet sich nicht von handschriftlichen Notizen	○	○	○	○	○	○
Mit Backstage konnte ich meine Unterlagen besser verwalten	○	○	○	○	○	○
Beim Wiederholen des Vorlesungsstoffs war Backstage hilfreich	○	○	○	○	○	○
Die Quizfragen haben mir geholfen, den Lernstoff besser zu verstehen	○	○	○	○	○	○
Die Quizfragen haben mir geholfen, zu erkennen, wo ich Probleme hatte	○	○	○	○	○	○
Es gibt für mich nur wenige Anreize am Austausch in Backstage teilzunehmen	○	○	○	○	○	○

8 [Block2OI1]

Ich hätte mehr am Austausch in Backstage teilgenommen, wenn folgende Bedingungen vorhanden gewesen wären

Please write your answer here:

| |
| |

10 [Block3Likert]Nutzung von Backstage durch Kommilitonen und Lehrkörper *

Please choose the appropriate response for each item:

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Die Vorlesung war für die Nutzung von Backstage geeignet	○	○	○	○	○	○
Es war gut, dass sich Tutoren während der Vorlesung in Backstage beteiligt haben	○	○	○	○	○	○
Es war gut, dass sich Tutoren nach der Vorlesung in Backstage beteiligt haben	○	○	○	○	○	○
Der Dozent hat in der Vorlesung ausreichend Bezug auf den studentischen Austausch in Backstage genommen	○	○	○	○	○	○
Ohne die Beiträge der Tutoren in Backstage wäre Backstage weniger nützlich	○	○	○	○	○	○
Die Anzahl der Quizfragen in den Vorlesungen war mir zu viel	○	○	○	○	○	○
Es hätten auch mehr Quizfragen in einer Vorlesung gestellt werden können	○	○	○	○	○	○
Die Quizfragen waren für das Verständnis des Vorlesungsstoffs geeignet	○	○	○	○	○	○
Die Quizfragen waren nützlich für die Übungen	○	○	○	○	○	○
Ich glaube nicht, dass die Quizfragen für die Klausur hilfreich sein werden	○	○	○	○	○	○

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Das Stellen von Quizfragen in der Vorlesung finde ich gut, aber die Fragen selbst schienen mir ungeeignet	○	○	○	○	○	○

11 [Block3OI1]Zur Nutzung von Backstage durch den Dozenten würde ich mir wünschen, dass …

Please write your answer here:

12 [Block3OI2]Zur Nutzung von Backstage durch die Tutoren würde ich mir wünschen, dass …

Please write your answer here:

13 [Block3OI3]Zur Nutzung von Backstage durch die Kommilitonen würde ich mir wünschen, dass...

Please write your answer here:

14 [Block3OI4]Von den gestellten Quizfragen würde ich mir wünschen, dass ...

Please write your answer here:

15 [Generic1]Allgemeine Anmerkungen zum Erstellen von Nachrichten und Notizen in Backstage:

Please write your answer here:

16 [Generic2]Allgemeine Anmerkungen zu den Quizfragen:…

Please write your answer here:

17 [Generic3]Allgemeine Anmerkungen zu Backstage:

Please write your answer here:

18 [Generic4]Allgemeine Anmerkungen zur Vorlesung:

Please write your answer here:

Vielen Dank für die Beantwortung der Fragen!
01.01.1970 – 01:00

Submit your survey.
Thank you for completing this survey.

eLearning WIS

Der Fragebogen liefert für uns wertvolle Informationen darüber, wie wir die Software, als auch deren Einsatzbedingungen verbessern können.

Der Fragebogen teilt sich in 3 Gruppen auf:

- Backstage (7 kurze Fragen)
- Termina (8 kurze Fragen)
- WIS Markt (10 Fragen)

Diese Umfrage enthält 25 Fragen.

Backstage

1 [B1]Ich besitze ein...

Bitte wählen Sie alle zutreffenden Antworten aus:

- [ ] ein Notebook
- [ ] ein Tablet (Android, iPad)
- [ ] ein Smartphone (Android, iPhone)
- [ ] keines der genannten Sorten

2 [B2]Ich bringe kein Notebook in die Vorlesung mit, weil...

Bitte wählen Sie alle zutreffenden Antworten aus:

- [ ] es mich zu sehr von der Vorlesung abhält
- [ ] die Vorlesung für die Arbeit mit Notebooks ungeeignet ist
- [ ] mein Notebook zu unhandlich für eine Mitnahme ist
- [ ] Sonstiges:

3 [B3]Falls ich eines der oben genannten webfähigen Geräte mitbringe, aber nicht am Backchannel teilnehme: Ich tausche mich nicht auf Backstage aus, weil....

Bitte wählen Sie alle zutreffenden Antworten aus:

- [ ] es zu zeitaufwendig ist
- [ ] ich den Austausch im Backchannel ablenkend finde
- [ ] ich selten Beiträge während der Vorlesung beisteuern kann
- [ ] ich lieber lese, was meine Mitstudenten schreiben
- [ ] Sonstiges:

4 [B4]Ich würde am Backchannel teilnehmen, wenn....

Bitte wählen Sie alle zutreffenden Antworten aus:

- [ ] es mit Smart Phones/Tablets möglich wäre
- [ ] dafür mehr Zeit während der Vorlesung gegeben würde
- [ ] die Beiträge stärker durch den Dozenten berücksichtigt würden
- [ ] es mehr Anleitung gäbe, wie ich mit dem Backchannel lernen soll
- [ ] Sonstiges:

5 [B5]Für die Nachbereitung der Vorlesung (z.B. Übungsblätter, Klausur)...

Bitte wählen Sie alle zutreffenden Antworten aus:

- [ ] sind die Backchannel-Beiträge hilfreich
- [ ] weiß ich nicht, ob die Beiträge hilfreich sind
- [ ] Sonstiges:

6 [B6]Meine Anmerkungen zu Backstage-Software:

Bitte geben Sie Ihre Antwort hier ein:

7 [B7]Meine Anmerkungen zu den Einsatzbedingungen von Backstage

Bitte geben Sie Ihre Antwort hier ein:

Termina

8 [T1]Bitte beurteilen Sie Ihre Motivation fürs Spielen von Termina

Bitte wählen Sie alle zutreffenden Antworten aus:

- ☐ Mich fasziniert der Spiel-Aspekt
- ☐ Ich wollte die Vorlesungsinhalte vertiefen
- ☐ Ich wollte die Forschung unterstützen
- ☐ Ich habe nicht teilgenommen
- ☐ Sonstiges: [ ]

9 [T2]Welches Vorwissen hatten Sie zu E-Learning-Spielen?

Bitte wählen Sie nur eine der folgenden Antworten aus:

- ○ Nie gehört
- ○ Gelesen/gehört
- ○ Schon mal verwendet

10 [T3]Ich habe Termina aufgerufen

Bitte wählen Sie nur eine der folgenden Antworten aus:

- ○ Selten
- ○ Einmal pro Woche
- ○ Mehrmals pro Woche

Nutzung von Backstage in der Vorlesung LDS im Sommersemester 2014

Dieser Fragebogen soll uns Einblicke gewähren,

- wie Du die Funktionalitäten von Backstage einschätzt,
- wie und für was Du Backstage verwendet hast und
- welche Verbesserungen Du für Backstage siehst.

Bitte investiere die wenigen Minuten, die nötig sind, um den Fragebogen vollständig zu beantworten. Vielen Dank!

There are 18 questions in this survey

Fragebogen

---

1 [Block1Likert]Funktionalitäten von Backstage *

Please choose the appropriate response for each item:

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Mir gefiel, in der Vorlesung öffentliche Kommentare mit Backstage erstellen zu können	○	○	○	○	○	○
Mir gefiel, in der Vorlesung anonyme Kommentare mit Backstage erstellen zu können	○	○	○	○	○	○
Mir gefiel, in der Vorlesung eigene Notizen mit Backstage erstellen zu können	○	○	○	○	○	○
Mir gefiel, Nachrichten auf Backstage bewerten zu können	○	○	○	○	○	○
Ich fand die Länge der Nachrichten von 200 Zeichen ausreichend	○	○	○	○	○	○
Die Bewertungen der Nachrichten war geeignet, um relevante Nachrichten zu erkennen	○	○	○	○	○	○
Es macht mir Spaß Nachrichten in Backstage zu bewerten	○	○	○	○	○	○
Es macht mir Spaß die Quizfragen in Backstage zu beantworten	○	○	○	○	○	○
Die Quizfragen sind ein geeignetes Mittel, Aktivität in der Vorlesung zu fördern	○	○	○	○	○	○
Es hat mir Spaß gemacht, die Vorlesung zu besuchen und mit Backstage wurde die Vorlesung für mich	○	○	○	○	○	○

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
angenehmer Ich finde es gut, dass ich auf der Übersichtsseite Informationen über mein Abschneiden in den Quizfragen bekomme	○	○	○	○	○	○
Ich finde es gut, dass ich auf der Übersichtsseite Informationen über das Abschneiden meiner Kommilitonen in den Quizfragen bekomme	○	○	○	○	○	○
Mir würde es ausreichen wenn die Quizfragen mündlich gestellt und per Handmeldung beantwortet werden	○	○	○	○	○	○

2 [Block1OI1]Mir hat besonders gut an Backstage gefallen, dass …

Please write your answer here:

3 [Block1OI2]

Der Einsatz von Backstage wäre für die folgenden Vorlesungen wünschenswert:

Please write your answer here:

4 [Block1OI3]Ich würde mir wünschen, dass zukünftige Versionen von Backstage noch folgende Funktionen für die Nutzung während der Vorlesung bereitstellen

Please write your answer here:

5 [Block1OI4]Ich würde mir wünschen, dass zukünftige Versionen von Backstage noch folgende Funktionen für die Nutzung für die Nachbereitung bereitstellen:

Please write your answer here:

6 [Block1OI5]Auf folgende Funktionen von Backstage könnte ich während der Vorlesung verzichten:

Please write your answer here:

---

7 [Block2Likert]Nutzung von Backstage *

Please choose the appropriate response for each item:

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Ich habe Backstage genutzt, um zu erfahren welche Fragen meine Kommilitonen haben	○	○	○	○	○	○
Ich habe Backstage genutzt, um die Antworten meiner Kommilitonen zu lesen	○	○	○	○	○	○
Ich habe Backstage genutzt, um die Antworten der Tutoren und des Dozenten zu lesen	○	○	○	○	○	○
Ich habe am Austausch von Nachrichten in Backstage gerne teilgenommen	○	○	○	○	○	○
Backstage ist für die Nachbereitung des Stoffs für die Vorlesungssitzungen nützlich gewesen	○	○	○	○	○	○
Backstage ist für die Nachbereitung des Stoffs für die Übungen nützlich gewesen	○	○	○	○	○	○
Backstage wird mir für die die Nachbereitung des Stoffs für die Klausur hilfreich sein	○	○	○	○	○	○
Mithilfe von Backstage habe ich viele Inhalte bereits in der Vorlesung verstanden	○	○	○	○	○	○
Backstage hat mich während der Vorlesung abgelenkt	○	○	○	○	○	○
Die Möglichkeit, Notizen auf Backstage zu erstellen	○	○	○	○	○	○

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
unterscheidet sich nicht von handschriftlichen Notizen	○	○	○	○	○	○
Mit Backstage konnte ich meine Unterlagen besser verwalten	○	○	○	○	○	○
Beim Wiederholen des Vorlesungsstoffs war Backstage hilfreich	○	○	○	○	○	○
Die Quizfragen haben mir geholfen, den Lernstoff besser zu verstehen	○	○	○	○	○	○
Die Quizfragen haben mir geholfen, zu erkennen, wo ich Probleme hatte	○	○	○	○	○	○
Es gibt für mich nur wenige Anreize am Austausch in Backstage teilzunehmen	○	○	○	○	○	○

8 [Block2OI1]

Ich hätte mehr am Austausch in Backstage teilgenommen, wenn folgende Bedingungen vorhanden gewesen wären

Please write your answer here:

9 [Block2OI2]

Ich habe Backstage für die Nachbereitung genutzt, weil… *

Please choose **only one** of the following:

○ Ja
○ Nein

Make a comment on your choice here:

10 [Block3Likert]Nutzung von Backstage durch Kommilitonen und Lehrkörper *

Please choose the appropriate response for each item:

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Die Vorlesung war für die Nutzung von Backstage geeignet	○	○	○	○	○	○
Es war gut, dass sich Tutoren während der Vorlesung in Backstage beteiligt haben	○	○	○	○	○	○
Es war gut, dass sich Tutoren nach der Vorlesung in Backstage beteiligt haben	○	○	○	○	○	○
Der Dozent hat in der Vorlesung ausreichend Bezug auf den studentischen Austausch in Backstage genommen	○	○	○	○	○	○
Ohne die Beiträge der Tutoren in Backstage wäre Backstage weniger nützlich	○	○	○	○	○	○
Die Anzahl der Quizfragen in den Vorlesungen war mir zu viel	○	○	○	○	○	○
Es hätten auch mehr Quizfragen in einer Vorlesung gestellt werden können	○	○	○	○	○	○
Die Quizfragen waren für das Verständnis des Vorlesungsstoffs geeignet	○	○	○	○	○	○
Die Quizfragen waren nützlich für die Übungen	○	○	○	○	○	○
Ich glaube nicht, dass die Quizfragen für die Klausur hilfreich sein werden	○	○	○	○	○	○

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Das Stellen von Quizfragen in der Vorlesung finde ich gut, aber die Fragen selbst schienen mir ungeeignet	○	○	○	○	○	○

11 [Block3OI1]Zur Nutzung von Backstage durch den Dozenten würde ich mir wünschen, dass ...

Please write your answer here:

12 [Block3OI2]Zur Nutzung von Backstage durch die Tutoren würde ich mir wünschen, dass ...

Please write your answer here:

13 [Block3OI3]Zur Nutzung von Backstage durch die Kommilitonen würde ich mir wünschen, dass...

Please write your answer here:

14 [Block3OI4]Von den gestellten Quizfragen würde ich mir wünschen, dass ...

Please write your answer here:

15 [Generic1]Allgemeine Anmerkungen zum Erstellen von Nachrichten und Notizen in Backstage:

Please write your answer here:

16 [Generic2]Allgemeine Anmerkungen zu den Quizfragen:...

Please write your answer here:

17 [Generic3]Allgemeine Anmerkungen zu Backstage:

Please write your answer here:

18 [Generic4]Allgemeine Anmerkungen zur Vorlesung:

Please write your answer here:

Vielen Dank für die Beantwortung der Fragen!
01.01.1970 – 01:00

Submit your survey.
Thank you for completing this survey.

Nutzung von Backstage in der Vorlesung Programmierung 1 im Wintersemester 14/15

Dieser Fragebogen soll uns Einblicke gewähren,

- wie Du die Funktionalitäten von Backstage einschätzt,
- wie und für was Du Backstage verwendet hast und
- welche Verbesserungen Du für Backstage siehst.

Bitte investiere die wenigen Minuten, die nötig sind, um den Fragebogen vollständig zu beantworten. Vielen Dank!

Diese Umfrage enthält 18 Fragen.

Fragebogen

---

1 [Block1Likert]Funktionalitäten von Backstage *

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Mir gefiel, in der Vorlesung öffentliche Kommentare mit Backstage erstellen zu können	○	○	○	○	○	○
Mir gefiel, in der Vorlesung anonyme Kommentare mit Backstage erstellen zu können	○	○	○	○	○	○
Mir gefiel, in der Vorlesung eigene Notizen mit Backstage erstellen zu können	○	○	○	○	○	○
Mir gefiel, Nachrichten auf Backstage bewerten zu können	○	○	○	○	○	○
Ich fand die Länge der Nachrichten von 200 Zeichen ausreichend	○	○	○	○	○	○
Die Bewertungen der Nachrichten war geeignet, um relevante Nachrichten zu erkennen	○	○	○	○	○	○
Es macht mir Spaß Nachrichten in Backstage zu bewerten	○	○	○	○	○	○
Es macht mir Spaß die Quizfragen in Backstage zu beantworten	○	○	○	○	○	○
Die Quizfragen sind ein geeignetes Mittel, Aktivität in der Vorlesung zu fördern	○	○	○	○	○	○
Es hat mir Spaß gemacht, die Vorlesung zu besuchen und mit Backstage wurde die Vorlesung für mich	○	○	○	○	○	○

3 [Block1OI2]

Der Einsatz von Backstage wäre für die folgenden Vorlesungen wünschenswert:

Bitte geben Sie Ihre Antwort hier ein:

4 [Block1OI3]Ich würde mir wünschen, dass zukünftige Versionen von Backstage noch folgende Funktionen für die Nutzung während der Vorlesung bereitstellen

Bitte geben Sie Ihre Antwort hier ein:

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
angenehmer						
Ich finde es gut, dass ich auf der Übersichtsseite Informationen über mein Abschneiden in den Quizfragen bekomme	○	○	○	○	○	○
Ich finde es gut, dass ich auf der Übersichtsseite Informationen über das Abschneiden meiner Kommilitonen in den Quizfragen bekomme	○	○	○	○	○	○
Mir würde es ausreichen wenn die Quizfragen mündlich gestellt und per Handmeldung beantwortet werden	○	○	○	○	○	○

2 [Block1OI1]Mir hat besonders gut an Backstage gefallen, dass …

Bitte geben Sie Ihre Antwort hier ein:

5 [Block1OI4]Ich würde mir wünschen, dass zukünftige Versionen von Backstage noch folgende Funktionen für die Nutzung für die Nachbereitung bereitstellen:

Bitte geben Sie Ihre Antwort hier ein:

6 [Block1OI5]Auf folgende Funktionen von Backstage könnte ich während der Vorlesung verzichten:

Bitte geben Sie Ihre Antwort hier ein:

---

7 [Block2Likert]Nutzung von Backstage *

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Ich habe Backstage genutzt, um zu erfahren welche Fragen meine Kommilitonen haben	○	○	○	○	○	○
Ich habe Backstage genutzt, um die Antworten meiner Kommilitonen zu lesen	○	○	○	○	○	○
Ich habe Backstage genutzt, um die Antworten der Tutoren und des Dozenten zu lesen	○	○	○	○	○	○
Ich habe am Austausch von Nachrichten in Backstage gerne teilgenommen	○	○	○	○	○	○
Backstage ist für die Nachbereitung des Stoffs für die Vorlesungssitzungen nützlich gewesen	○	○	○	○	○	○
Backstage ist für die Nachbereitung des Stoffs für die Übungen nützlich gewesen	○	○	○	○	○	○
Backstage wird mir für die die Nachbereitung des Stoffs für die Klausur hilfreich sein	○	○	○	○	○	○
Mithilfe von Backstage habe ich viele Inhalte bereits in der Vorlesung verstanden	○	○	○	○	○	○
Backstage hat mich während der Vorlesung abgelenkt	○	○	○	○	○	○
Die Möglichkeit, Notizen auf Backstage zu erstellen	○	○	○	○	○	○

9 [Block2OI2]

Ich habe Backstage für die Nachbereitung genutzt, weil... *

Bitte wählen Sie nur eine der folgenden Antworten aus:

○ Ja
○ Nein

Bitte schreiben Sie einen Kommentar zu Ihrer Auswahl

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
unterscheidet sich nicht von handschriftlichen Notizen	○	○	○	○	○	○
Mit Backstage konnte ich meine Unterlagen besser verwalten	○	○	○	○	○	○
Beim Wiederholen des Vorlesungsstoffs war Backstage hilfreich	○	○	○	○	○	○
Die Quizfragen haben mir geholfen, den Lernstoff besser zu verstehen	○	○	○	○	○	○
Die Quizfragen haben mir geholfen, zu erkennen, wo ich Probleme hatte	○	○	○	○	○	○
Es gibt für mich nur wenige Anreize am Austausch in Backstage teilzunehmen	○	○	○	○	○	○

8 [Block2OI1]

Ich hätte mehr am Austausch in Backstage teilgenommen, wenn folgende Bedingungen vorhanden gewesen wären

Bitte geben Sie Ihre Antwort hier ein:

10 [Block3Likert]Nutzung von Backstage durch Kommilitonen und Lehrkörper *

Bitte wählen Sie die zutreffende Antwort für jeden Punkt aus:

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Die Vorlesung war für die Nutzung von Backstage geeignet	○	○	○	○	○	○
Es war gut, dass sich Tutoren während der Vorlesung in Backstage beteiligt haben	○	○	○	○	○	○
Es war gut, dass sich Tutoren nach der Vorlesung in Backstage beteiligt haben	○	○	○	○	○	○
Der Dozent hat in der Vorlesung ausreichend Bezug auf den studentischen Austausch in Backstage genommen	○	○	○	○	○	○
Ohne die Beiträge der Tutoren in Backstage wäre Backstage weniger nützlich	○	○	○	○	○	○
Die Anzahl der Quizfragen in den Vorlesungen war mir zu viel	○	○	○	○	○	○
Es hätten auch mehr Quizfragen in einer Vorlesung gestellt werden können	○	○	○	○	○	○
Die Quizfragen waren für das Verständnis des Vorlesungsstoffs geeignet	○	○	○	○	○	○
Die Quizfragen waren nützlich für die Übungen	○	○	○	○	○	○
Ich glaube nicht, dass die Quizfragen für die Klausur hilfreich sein werden	○	○	○	○	○	○

	Stimme voll zu	Stimme zu	Stimme eher zu	Stimme eher nicht zu	Stimme nicht zu	Stimme gar nicht zu
Das Stellen von Quizfragen in der Vorlesung finde ich gut, aber die Fragen selbst schienen mir ungeeignet	○	○	○	○	○	○

11 [Block3OI1]Zur Nutzung von Backstage durch den Dozenten würde ich mir wünschen, dass …

Bitte geben Sie Ihre Antwort hier ein:

12 [Block3OI2]Zur Nutzung von Backstage durch die Tutoren würde ich mir wünschen, dass …

Bitte geben Sie Ihre Antwort hier ein:

15 [Generic1]Allgemeine Anmerkungen zum Erstellen von Nachrichten und Notizen in Backstage:

Bitte geben Sie Ihre Antwort hier ein:

16 [Generic2]Allgemeine Anmerkungen zu den Quizfragen:…

Bitte geben Sie Ihre Antwort hier ein:

17 [Generic3]Allgemeine Anmerkungen zu Backstage:

Bitte geben Sie Ihre Antwort hier ein:

13 [Block3OI3]Zur Nutzung von Backstage durch die Kommilitonen würde ich mir wünschen, dass…

Bitte geben Sie Ihre Antwort hier ein:

14 [Block3OI4]Von den gestellten Quizfragen würde ich mir wünschen, dass …

Bitte geben Sie Ihre Antwort hier ein:

18 [Generic4]Allgemeine Anmerkungen zur Vorlesung:

Bitte geben Sie Ihre Antwort hier ein:

Vielen Dank für die Beantwortung der Fragen!
01.01.1970 – 01:00

Übermittlung Ihres ausgefüllten Fragebogens:
Vielen Dank für die Beantwortung des Fragebogens.