

# Evaluation of a benchmark system for analyzing collaborative group performance as part of an educational online knowledge management system

Dr. Wolfgang Semar  
University of Konstanz  
Information Science  
Box 87  
D – 78457 Konstanz (Germany)  
wolfgang.semar@uni-konstanz.de

*Abstract - E-learning systems that support collaborative learning environments do not work without proactive motivation of their users. Users need to know what their benefits are when sharing knowledge and contributing actively in such forums. Therefore the collaborative knowledge management system “K3” which is used in academic education of Information Engineering students at the university of Konstanz in Germany has been developed within a benchmark system to motivate users. This paper, on one hand describes how the different benchmark means of quantifying work together in measuring and assessing users’ performance and thus stimulating their willing to cooperate in their collaborative work, and on the other hand this paper describes the result of the first evaluation of the benchmark system.*

**Keywords: collaborative knowledge management; CSCW; incentive system; motivation; measurement; benchmark system; evaluation**

## 1.0 Introduction

Well-established and popular groupware systems like Lotus Notes, MS Exchange, and many open source projects show the high standard of computer-based support of learning communities and work groups. The assumed supremacy of collaborative knowledge management (vs. the approach of individual learning) is based on the productive exchange and sharing of knowledge among virtual connected groups which balances knowledge asymmetries [5; 7, p. 196]. Moreover, an environment that requires frequent communication fosters constant feedback, both from instructor and group members. Feedback is a crucial issue for developing and maintaining students' motivation [9; 1]. Classic e-learning systems offer a broad range of functionality, but are no more than a technical platform for exchanging and hosting learning and teaching material. There, the actual process of learning is supported only by tools for organising and measuring learning success by assessment and examination. True feedback components are missing.

In collaborative learning environments mere measurement of students' performance is not the appropriate way. The evaluation of learning success is being done while the groups are collaborating. This permanent feedbacking ensures a successful cooperative process to which the result is a gain of knowledge for all participants.

Feedback and evaluation are processed with scores and benchmarks for the communicative and learning activities in the groups. They indicate how much interaction is going on and how the

discourse is controlled. They may also serve the lecturer in setting up marks. However, benchmarks give information about the quantity of students' work only, not on the quality. Though, benchmarks are of strongly motivational nature beside their evaluative quality.

It has been shown that the participants in collaborative electronic learning systems need to be proactively motivated and supported. Usually, this is done by personal addresses: A lecturer will ask the students for certain actions within the system. This forces performance in the learning environment, but it is more successful and sustainable to really motivate the students. That is where incentive and motivation systems come in [6, p. 23]. Users need to know what their benefits are when sharing knowledge and contributing actively in such forums. Therefore the collaborative knowledge management system "K3"<sup>1</sup> which is used in academic education of Information Engineering students at the university of Konstanz in Germany has been developed within a benchmark system to motivate users. It consists of various visualized benchmarks for every single user. So they get individual assessment and are motivated to collaborate and to take part in generating knowledge. This benchmark system is part of the "K3" software which enables collaborative creation of conceptual knowledge from heterogeneous resources and through electronic communication forums. K3 is an open software system that supports collaborative and distributed production of conceptual knowledge in academic learning environments by using heterogeneous resources and moderated electronic communication forums. Further information competency is to be gained by embedding external information resources. This knowledge is strongly linked, structured by context and semantics as well as visualized to ensure comfortable navigation. A rating feature is integral part of the K3 system and is the basis of the incentive system. Every entry a student makes to the system – be it a comment on a current thread or a reference link – is registered and credited as individual performance or as part of collaborative work. These contributions also generate certain scores and there is a visualized output. This is a permanent feedback function showing the students how they are performing. By comparing individual performance with other students' performance every participant can see their current standing within the community. Thus it is possible to have a dynamic and individual evaluation of learning success as well as an assessment of the group's collaboration activities. An incentive system with strong focus on reputational aspects has been established to support the whole process of generating knowledge. The underlying didactic idea of K3 is that of collaborative group work. A team is given a task by the instructor (on course level) and the team has to solve this task on their own (on group level). Each member of the group (on individual level) has to enrol to one of various given roles (i.e. presenter, researcher, moderator, summarizer) which they hold until the task is finished. The team decide on their own which role is taken by whom. This process of assigning roles (by being discussed in the system) has to be marked as entry type "organisational". A corresponding field type is provided. To ensure collaborative knowledge work, every participant has beside his role function to take part in the discussion. Each entry needs to be typed by its contributor. K3 provides the entry types: result, amendment, organisational entry, new topic, hypothesis, question, and q+a. Every entry is assigned a specific label, depending on the type of entry and the author. So every entry bears a special mark depending on its type and author, and it can be seen by any participant.

## **2.0 K3 benchmarks for measuring the performance of the participants**

To show the individual performance of users automatically generated benchmarks are taken. Different grades and levels of activity to measure the readiness for interaction and communication in electronic communication forums are described by [4, 50 pp.]. We use them as a basis for further measures to rate the activities of K3 users. However, one must not use too many benchmarks for they may cause information overload. To avoid this, the benchmarks are compacted in a benchmark system and are

---

<sup>1</sup> K3 is a system that is currently being developed at the university of Konstanz/Chair of Information Science. It is a project funded by the German Ministry of Science and Education (BMBF, DLR PT-NMB+F, Projectnumber: 08C5896). The acronym K3 for - collaboration, communication, and competence. For further information see the project's website: [www.k3forum.net](http://www.k3forum.net)

visualized in a second step. For setting up the incentive/motivational benchmark component in K3 it is not helpful to use a hierarchic method, for not all K3 benchmarks are mathematically related. The more useful approach is to have the measures in an order defined by subject and content criteria [3, p. 555]. [2, p. 50] suggest a benchmark system for LMSs (learning management systems) from which we borrow the K3 benchmarks: coverage, relation, and time range that are registered on four levels: system level, course level, group level, and individual level. Coverage is generated from measures like number of participants and entries and is given as absolute numbers (and sums). The combination of absolute numbers generates relation figures. They are shown as percentage or index numbers [8, p. 8]. Time range figures are derived from monitoring long-time user performance. By analyzing timelines changes in benchmarks can be identified then.

On team level and on individual level there are some benchmarks of organisational nature, but particularly there are didactic figures, for they are important for enhancing motivation, especially when showing and comparing performance of the different groups in relation to each other. Also the changing of a group's figures during time is important, because it shows the team's development. And it is these benchmarks on team level that are the most interesting, because they indicate the actual collaborative knowledge management. But also we need ways to measure the relative amounts of synthesis, independence, interaction, and participation of a group. For every team each of the four characteristics is taken and the "degree of collaboration" is set up. These "collaboration degree" will then allow us to compare groups for the amount of collaboration they exhibit [10, p. 57].

The benchmarks shown in tables 1 and 2 are mainly fix and relational measures. They are being stored for a period of 2-3 weeks, so time measures can be calculated. The lecturer can see how students and groups perform and s/he can intervene, if necessary. But not only to the lecturer, also to the students the benchmarks are a means of awareness. So anyone can see one's own and the others' state of performance and they can react accordingly.

<b>Benchmark on individual level</b>	<b>Calculating</b>	<b>Description</b>
Overall activity	Number of all articles of a participant.	Ranking of the most active participant.
Personal degree of reaction	Number of all replies by a participant to other articles.	Ranking of response frequency.
Degree of active reaction	Ratio of one's own replies to all one's own articles.	Value between 1 (participant did react) and 0 (participant did not react).
Degree of passive reaction	Ratio of group's replies to all one's own articles.	Value between infinite (participant received many reactions) and 0 (participant received no reactions)
Degree of reputation	Number of links (of all participants) to one's own articles.	Number of references given.
Personal degree of information	Discourse starting entries / all reactions	Indicates, if anyone reacted only or acted proactively.
Personal degree of referencing	Referential objects / all entries.	Shows the intensity of using external proof.
Degree of interaction of a group member	$1 - (\text{Number of stand-alone entries in a group} / \text{Number of entries by all students})$ .	If the result is close to 0, there is little interaction; if it is close to 1, there is a lot of interaction.
Degree of participation of a group member Pi	The ratio of (Number of entries by user i / Number of entries by all group members).	If this value is close to 0, the member has not done much group work, if the value is close to 1, this member has made all contributions.

Delta Pi (degree of deviation)	The deviation $\sigma$ from (default value)	Thus indicates the deviation between the standard value and the personal degree of participation.
--------------------------------	---	---

Tab. 1 A selection of the K3 benchmarks coverage, relation, and time range on individual level.

Benchmark on group level	Calculating	Description
Overall group activity	Number of all articles of the group	Ranking of the most active groups.
Degree of lecturer's correcting	Number of lecturer's correcting entries (within a group)	Number of lecturer's interventions.
Degree of underwriting	Ratio of given reference objects to all articles of the group.	Shows to which extent the group referred to external sources.
Degree of moderation	Number of moderation / Number of all entries (within a group)	Shows the intensity of moderation in a group.
Degree of organisation	Number of organisational entries / Number of group entries (within a group)	Shows how well-organised a group acts.
Degree of participation of a group	The degree is defined as identical with the normalised entropy $\ln(x)$ [10, p. 57].	If the result is close to 0, there is imbalanced participation of the single members; if it is close to 1, the participation of the members is fairly balanced.
Degree of interaction of a group	$1 - (\text{Number of stand-alone entries in a group} / \text{Number of entries by all students})$	If the result is close to 0, there is little team interaction; if it is close to 1, there is a lot of interaction.
The degree of independence of a group	$1 - (\text{Number of corrective instructor's entries} / \text{Number of all entries in the group (students' plus corrective instructor's entries)})$	If it is close to 0, there is little independence within the team; if close to 1, there is strong independence.
The degree of synthesis of a group	Every participant of the group has to consent and to rate the summary with a voting tool.	If all group members agree with the summary and each individual entry has been respected, the result is close to 1; if it is close to 0, there has been no collaborative group work.
Degree of collaboration within a group	This is a quadruple of the four degrees: "degree of participation", "degree of interaction", "degree of independence", and "degree of synthesis".	The "degree of collaboration of a group" shows if a group is really collaborating or if it is just cooperating, and how successful and effective the collaboration is.

Tab. 2 A selection of the K3 benchmarks coverage, relation, and time range on group level.

### 3.0 Evaluation of K3 benchmarks

In K3, benchmarks are presented in tables as well as by various visualised forms. They can be displayed on group level (group benchmarks only) and on individual level (for each participant). The temporal development of the data is displayed in a weekly chart.

It turned out quickly that the benchmark tables are more useful if put in relation to each other. F. ex., a measure like “overall group activity” provides more information when the highest and the lowest activity are known and how the group compares to other groups. So it is necessary to have comparative features.

Fig. 1 shows benchmarks in tables. The example given is a group of three. The upper frame displays fix measures (name, role, number of questions, hypotheses, new topics, amendments, q+a's, results, organisational, miscellaneous, hyperlinks, uploads, references, entries per role, overall entries (absolute, percentage)). The middle frame displays relational measures (personal degree of reaction, degree of active reaction, degree of passive reaction, degree of reputation, personal degree of information, Delta Pi (degree of deviation), personal degree of synthesis). The lower frame displays group degrees (degree of participation of the group, degree of interaction of a group, degree of independence of a group, degree of synthesis of a group, overall group activity, amount of lecturer's correcting, degree of moderation, degree of organisation, up-to-dateness of entries and degree of underwriting).



Fig. 1 Selected individual and group benchmarks of one group in table form.

To ensure user-friendliness and efficiency almost all benchmarks are given in graphic presentation. Fig. 2 shows the fix measures graphics for an individual member: The other group's members' scores are displayed as well, so it is possible to compare the members at one sight.

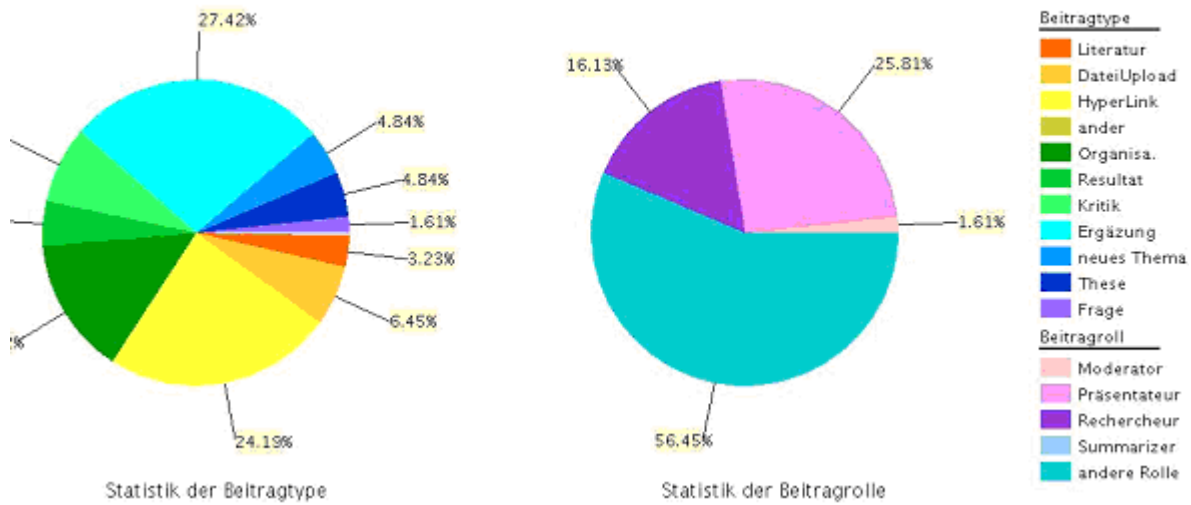


Fig. 2 Selected fix measures for an individual group member in graphical presentation.

As mentioned before, the most effective clues are to be taken from relational measures. They occur on individual and on group level. Fig. 3 shows the degree of deviation. The 0-line assigns a group's default value. The deviation from the line indicates, if a member contributed more (+ value) or less (-value) than average. The example shows that user "U2" contributed slightly more than twice as much than U1 and U3. Fig. 4 shows the degree of group collaboration for five different groups. One can see the similarity for participation, interaction, and independence, but also the difference in the degree of synthesis which is weaker. For reasons of readability the group collaboration benchmarks are displayed as columns. As shown in fig. 5 with the comparison of five different groups.

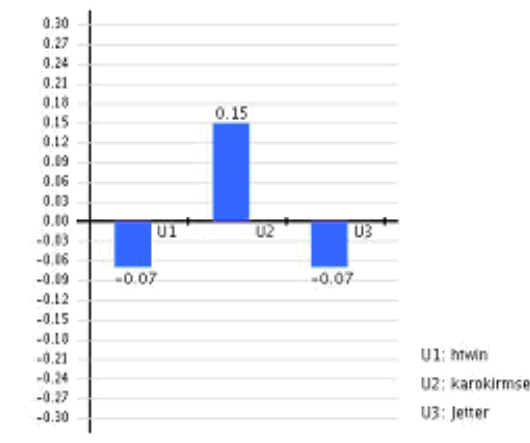


Fig. 3 "Degree of deviation"

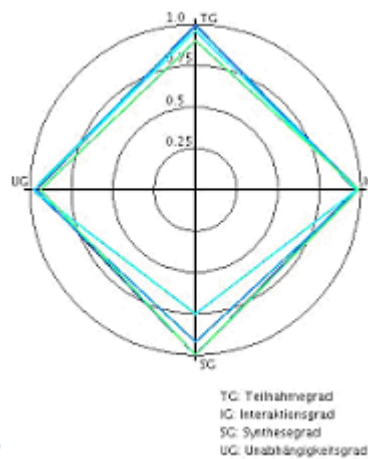


Fig. 4 Comparison of "degree of group collaboration" for 3 different groups

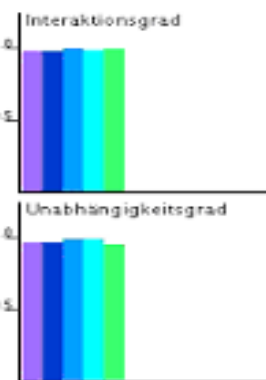


Fig. 5 Comparison of "degree of group interaction" and "degree of group independence" for 5 different groups

## 4.0 Summary

K3 benchmarks have been designed to evaluate and to rate the collaborative activities of the groups and members. The comparison of the individual scores and making it visible to every member is also a strongly motivational momentum. It is also a proof of discourse control (f.ex. to see whether the lecturer had to intervene or not). For the lecturer, it is a great help for assessing students. It has, however, to be kept in mind that benchmarks work on a quantity basis and do not reflect quality issues. To rate the quality of discourse objects it is necessary to analyse content (intellectually and/or automatically). The first evaluation of the benchmark system showed that the benchmarks have to be refined and that advanced visualisation will be helpful. All in all, we conclude that continuous assessment and displaying benchmarks have positive impacts on the work and motivation of K3 users.

## 5.0 References

- [1] Griesbaum, Joachim; Rittberger, Marc: A Collaborative Lecture in Information Retrieval for Students at Universities in Germany and Switzerland. In: Proceedings of the World Library and Information Congress: 71st IFLA General Conference and Council. "Libraries - A voyage of discovery". August 14th - 18th 2005, Oslo, Norway.
- [2] Grob, Heinz Lothar; Bensberg, Frank; Dewanto, Lofi; DÜppe, Ingo: Controlling von Learning Management-Systemen – ein kennzahlorientierter Ansatz. In: Carstensen, Doris; Barrios, Beate (eds.): Campus 2004. Kommen die digitalen Medien an der Hochschule in die Jahre? Münster: Waxmann, 2004, p. 46 - 56
- [3] Hummel, Thomas: Quellen und Elemente von Informationssystemen des Controlling. In: Steinle, Claus; Bruch, Heike (eds.): Controlling - Kompendium für Ausbildung und Praxis. Stuttgart: Schäffer-Poeschel, 2003
- [4] Kuhlen, Rainer: Mondlandung des Internet. Konstanz: UVK, 1998
- [5] Paechter, Manuela: Wissenskommunikation, Kooperation und Lernen in virtuellen Gruppen. Lengerich: Pabst Science Publishers, 2003
- [6] Schanz, Günther: Motivationale Grundlagen der Gestaltung von Anreizsystemen. In: Schanz, Günther (ed.): Handbuch Anreizsysteme in Wirtschaft und Verwaltung. Stuttgart: Poeschel, 1991, p. 3 - 30
- [7] Schulmeister, Rolf: Virtuelle Universität – Virtuelles Lernen. München: Oldenbourg, 2001
- [8] Schwickert, Axel C.; Wendt, Peter: Controlling-Kennzahlen für Web Sites. In: Arbeitspapiere WI, Nr. 8/2000, (ed): Lehrstuhl für Allg. BWL und Wirtschaftsinformatik, Johannes Gutenberg-Universität: Mainz, 2000
- [9] Semar, Wolfgang: Incentive Systems in Knowledge Management to Support Cooperative Distributed Forms of Creating and Acquiring Knowledge. In: Arabia, Hamid et al. (eds.): Proceedings of the International Conference on Information and Knowledge Engineering - IKE'04. Las Vegas: CSREA Press, 2004, p. 406 – 411
- [10] Semar, Wolfgang: Development of a benchmark system for analyzing collaborative group performance as part of an educational online knowledge management system. In: Arabia, Hamid et al. (eds.): Proceedings of the 2005 International Conference on Information and Knowledge Engineering - IKE'05. Las Vegas: CSREA Press, 2005. p. 53 – 59