

# Common Ground and Individual Accountability in Literature Selection of Groups: Three Different Group Learning Techniques

Michael Schubert, University of Tübingen, michael.schubert@uni-tuebingen.de

Jürgen Buder, Knowledge Media Research Center, j.buder@iwm-kmrc.de

Roman Rädle, University of Konstanz, roman.raedle@uni-konstanz.de

Friedrich W. Hesse, Knowledge Media Research Center, f.hesse@iwm-kmrc.de

**Abstract:** Libraries heavily changed over the last decades, supporting more and more the collaborative learner in providing adapted spaces and technologies. The transformation seems not surprising, as group learning is generally considered to be better than individual learning, although the mechanism behind is still unclear. In a single factor experiment with 180 students, 3-person groups jointly selected research literature for a Geology course at a Multitouch Tabletop. Based on the notions of common ground and individual accountability one third of the groups performed the task collaboratively. Another third had a preceding cooperative stage involving a division of labor, whereas the remaining third of groups had a preceding “parallel” stage. Results indicate a significant difference in learning achievements favoring the parallel condition. The results are discussed in the context of learning environments in libraries.

**Keywords:** library, problem-solving, group learning, multitouch tabletop, experiment

## Introduction

In the last decades, libraries experienced a tremendous transformation by providing more and more secluded rooms for groups to work and learn together using digital technologies (Gabbard, Kaiser, & Kaunelis, 2007). This transformation comes to no great surprise as the benefits of group learning has been shown for more than forty years, summarizing that learning in groups is better than learning alone (e.g. Springer, Stanne, & Donovan, 1999). There are many techniques to learn effectively in groups (e.g. collaboration, cooperation). Nonetheless, the way how group learning should be structured and fitted to real world scenarios is still unclear.

For the purpose of this paper, we build on socio-cognitive research, arguing for the combination of two factors: common ground as a social factor, and individual accountability as a cognitive factor. Based on this rationale, we report first results of an empirical study altering the two factors in three specific group learning techniques (collaboration, division of labor and parallel cooperation). The task involves computer-supported face-to-face problem-solving, which is situated in the context of literature screening and selecting in a library-based setting.

## Three techniques of group learning

Collaboration (C) can be regarded as the gold standard of group learning. Working permanently together, one of the many benefits of C is that over time learners gain a common ground about the task at hand. Common ground refers to shared knowledge, mutual beliefs and mutual assumptions to coordinate activities (Baker, Hansen, Joiner & Traum, 1999). By the means of grounding (Baker et al., 1999), groups not only overcome gaps in their shared understanding, they also build a group knowledge, which is especially important while performing problem solving tasks.

However, research on motivation losses also identified effects, such as social loafing, which hamper group achievement in collaboration (Kerr & Tindale, 2004). To bypass this problem, scholars studying cooperative learning techniques found that individual accountability plays an important role in motivating each group member to contribute to a group learning process (Johnson, Johnson, & Smith, 2007). Hence, instead of permanently working together, multi-stage approaches are introduced, including an individual stage and a collaborative stage. One way to initiate an individual stage, and therefore individual accountability, is to rely on techniques using some division of labor (DoL). In stages of DoL, mostly no or just part of the task material is being shared in the group. Accordingly, it can be assumed that group learners who first use a DoL are motivated to gain individual expertise as the group has to rely on it in a subsequent collaborative stage. However, the downside of DoL is that learners will not possess common ground, and therefore a shared understanding of the whole material, when entering the collaborative phase.

In order to induce both common ground and individual accountability at the same time, a two-stage technique of group learning is used to which we refer as the parallel cooperation (PC). In the first stage, learners individually work on an identical task (i.e. share the same material/no DoL), while in the second stage individual results can be discussed. Thus, a “divergence of labor” in the first phase results in heightened individual accountability producing idiosyncratic viewpoints with both varying and redundant contributions. At the same time, it creates a common ground as every learner needs to elaborate on the same material. We like to use the technique in a computer-mediated setting, which allows us to not only capture individual learning contributions systematically, but make them even comparable for the second stage, so that awareness about each group member's understanding is increased.

In order to address the possibility of comparing group member's learning activities, we decided to use a Multitouch Tabletop (MTT) technology. According to Jermann and Dillenbourg (2008), the MTT can either reflect individual and group activities, or transform these activities in a way that makes them more comparable and accessible for the group. While the former is important for collaborative and DoL groups, the latter characteristic might account especially for PC groups.

## The present study

The present study compares the three techniques of group learning in a problem-solving task that involves the joint activity of screening and selecting research literature in a computer-supported library setting. Three-person groups were asked to select four out of twelve journal articles, which they would further employ in a seminar report on “Water Scarcity in Sub-Saharan Africa”. To make the selection of four, participants read the abstracts of each article and rated them on a scale ranging from 0 (very irrelevant) to 6 (very relevant). Even though abstracts provide just a blink of information compared to full-text articles, we decided to use abstracts because in real-world literature selection tasks rarely all relevant articles are fully read before deepen into the material. After rating every article, the four highest rated articles were automatically selected if the group decided to finish the task. Groups in the C condition did the entire selection task in a one-stage approach as a group. In the DoL condition, a first stage had individuals rate four out of the twelve articles individually. In the second stage, individual ratings were transferred into a shared external representation, and learners had the opportunity to change these ratings collaboratively. In the PC condition, the first stage required learners to individually rate all twelve articles. For the second stage, individual ratings were also transferred into a shared external representation, and learners had the opportunity to change ratings collaboratively. In order to address the practical question of whether group learning outperforms individual learning, we employed a fourth condition as our base line with learners who did the entire task in a one-stage approach alone and without a group context.

Based on the considerations about common ground and individual accountability, it was expected that the C condition and the PC condition will be associated with common ground and shared understanding about the task at hand, reflected in a higher degree of post-interaction shared mental models (*Hypothesis 1*). Further, the DoL technique and the PC technique should benefit from heightened individual accountability, resulting in a higher motivation to contribute to the collaborative discussion in the second stage, reflected by a higher equality of participation (*Hypothesis 2*). Finally, we expected that the PC technique which benefits from both common ground and individual accountability would yield the highest problem-solving performance compared to the other group techniques, reflected in the similarity of the final selection compared to an expert solution (*Hypothesis 3*).

## Methods

### Participants and design

To conduct our experiment, we recruited a total of 180 students (119 women; 61 men; mean age = 24.39 years) from the University of Tübingen. Students were either paid 8 Euros, or received course credit. To reduce the effect of prior knowledge, we excluded students from Geography. The experiment had a single factor design with three different interventions (group learning techniques): collaboration, DoL and PC. Moreover, we also investigated an individual condition as our baseline. Participants were randomly assigned to the conditions.

### Materials

#### Articles

In a preliminary study, the twelve articles had been selected from a set of 21 articles, by having 30 participants individually rate article abstracts on readability and understandability. The items were presented as 5-point scales ranging from 1 (e.g. not readable at all) to 5 (e.g. very readable). Based on the results, we selected 12

articles with the highest ratings of readability ( $M = 3.89$ ;  $SD = 0.37$ ) and understandability ( $M = 3.81$ ;  $SD = 0.31$ ). After having the set of twelve articles, we additionally asked a subject matter expert (SME) to rate each article on a 7-point scale to conduct the final relevance ratings according to the task and the seminar theme. These relevance ratings have later been used to evaluate performance measures.

### Multi-touch tabletop (MTT)

As an MTT device for both the group conditions and for the individual condition, we used a Samsung SUR40. To make article reading and rating possible, we further developed a tool which allowed the handling of articles with multi-touch interactions.

In the collaboration condition and in the individual condition, the participants used the same interface for the whole time. In the DoL and PC condition, in order to allow individual working before discussing results, first, three distinct private spaces were presented, providing the functionality of reading and rating articles. Depending on condition, either four (DoL) or twelve (PC) articles were presented, and articles were color-coded to allow better distinction afterwards. Once all group members submitted their individual rating, private spaces disappeared, and the screen changed to the full-screen interface version. In the DoL condition, it showed the individual ratings as a cumulated representation, keeping the color-coding, so that participants knew which article was rated by whom. In the PC condition, it showed a transformation of the individual ratings to arithmetic means. Additionally, individual, color-coded ratings were preserved for each article, making it more comparable to distinguish which article was rated by whom to which degree.

## **Procedure**

On arrival at the laboratory, participants were assigned randomly to one of the group conditions. Participants first read task descriptions and filled out a pre-intervention questionnaire asking for demographic data and prior knowledge. After finishing the questionnaire, participants were invited to test all functionality (moving, opening and rating articles) in a demo version of the basic interface with three non-related articles for around five minutes. Afterwards, the group started the task as described above having no time limit to finish the task. The whole task procedure was videotaped. Once participants agreed upon their selection of four, they filled out a post-intervention questionnaire. Finally, participants were debriefed and rewarded for their participation.

## **Measures**

### Shared mental models

In the post-intervention questionnaire we measured shared mental models about selection criteria (Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000). Nine rating criteria have been presented to the participants (e.g., currentness of date or relatedness to the region) asking to indicate for each criterion whether it played an important role for the group during discussion or not (categorical: y/n). Frequencies and percentage of complete agreements (three times yes or no) upon one criterion were measured and transformed to odds ratio ( $-1 / 1$ , where 1 means perfect agreement upon all nine criteria).

### Equalization of participation

To calculate equalization of participation, we used the standard deviation of talking time (in seconds) across group members in the collaborative stage. The lower the standard deviation, the more equal the participation (whereas 0 would mean perfect equalization). Using the coding software ANVIL (Kipp, 2001), two coders double-coded 33% of the videos with an interrater-reliability of .765 (corrected Cohen's kappa), which indicates a substantial agreement.

### Performance

Before we started calculating the performance, data has been cleaned by dropping all ratings which a group set to a value between 0 and 1 (of a scale from 0 to 6) in the discussion phase, classifying them as dismissed (non relevant) articles. Then, we took the final rating of all not-dismissed articles from the log files and calculated differences to the corresponding subject matter expert ratings. Finally, the mean difference upon all articles (ranging between 0 and 6) of one group was computed—the smaller the mean difference, the better the group performance.

## **Findings**

Hypothesis 1 predicted that groups in which common ground is established (collaboration, PC) will have a greater degree of sharedness among post-intervention mental models about selection criteria. Using an odds ratio for calculating agreement, a one-way ANOVA of shared criteria mental models between the three group

conditions showed no significant effect,  $F < 1$ . There was no difference between the collaboration condition ( $M = .27$ ,  $SD = .82$ ), the DoL condition ( $M = .51$ ,  $SD = 1.07$ ), and the PC condition ( $M = .15$ ,  $SD = .75$ ). This means that hypothesis 1 has to be rejected.

For testing our hypothesis that individual accountability has a moderating effect on equality of participation (Hypothesis 2), we measured the total time of talk of each participant (in seconds) and looked on the standard deviation of each group. As a Levene's test of equality of error variance was significant ( $p < .05$ ), we ran a Kruskal-Wallis-Test. Results indicate a difference among the three group learning techniques ( $H(2) = 14.387$ ,  $p = .001$ ,  $w = .52$ ), with a mean rank of 38.72 for collaborative groups, 24 for DoL groups and 19.78 for PC groups. Further post-hoc analysis using the Mann-Whitney-U test, revealed significant differences between the collaborative condition ( $Mdn = 129.8$ ) and the DoL condition ( $Mdn = 67.6$ ),  $U(34) = 74$ ,  $Z = -2.78$ ,  $p = .005$ , and significant differences between the collaborative condition and the PC condition ( $Mdn = 39.1$ ),  $U(34) = 74$ ,  $Z = -3.607$ ,  $p = .000$ . No differences were found between the DoL condition and the PC condition ( $U(34) = 137$ ,  $Z = -0.79$ ,  $p = .429$ ). On the basis of these data, Hypothesis 2 can be confirmed.

Hypothesis 3 tested whether there were differences in group performance among all four conditions. Therefore, we performed a one-way ANOVA of selection performance (mean rating differences to SME ratings) between the four conditions to see which condition reached the lowest mean differences (best performance) at the end. Results indicated a significant difference of  $F(3, 68) = 5.576$ ,  $p = .002$ ,  $\eta^2 = .197$ . Further analyses using Tukey-HSD revealed that this effect was due to a better performance in the PC condition. This condition ( $M = 1.06$ ,  $SD = .29$ ) differed significantly from the collaborative condition ( $M = 1.49$ ,  $SD = .43$ ,  $p = .009$ ), from the DoL condition ( $M = 1.51$ ,  $SD = .44$ ,  $p = .006$ ), and also from the individual condition ( $M = 1.50$ ,  $SD = .43$ ,  $p = .007$ ). Thus, Hypothesis 3 could be confirmed.

The design of the experiment also allowed to further analyze whether groups in the two-stage conditions (DoL, PC) had different performance levels over time, as both the performance of collated ratings after the first, individual phase and of collated ratings after the second stage could be compared. A 2 (condition: DoL and PC)  $\times$  2 (time of measurement: before and after discussion phase) ANOVA with time as within-subject factor revealed a disordinal interaction effect  $F(3, 68) = 19.756$ ,  $p < .01$ ,  $\eta^2 = .368$ . Further paired t-tests comparing measurement times (pre/post discussion phase) revealed that performance increased for the PC condition,  $M_{pre} = 1.42$ ,  $SD_{pre} = .29$ ,  $M_{post} = 1.06$ ,  $SD_{post} = .29$ ,  $t(17) = 4.943$ ,  $p < .001$ . In contrast, performance in the DoL condition slightly decreased over time,  $M_{pre} = 1.29$ ,  $SD_{pre} = .22$ ,  $M_{post} = 1.51$ ,  $SD_{post} = .44$ ,  $t(17) = -2.028$ ,  $p = .059$ . In other words, these results lent further support to the idea that the PC technique was effective.

## Discussion

This study compared different techniques of group learning, based on the rational that they differ with regard to common ground and individual accountability. While our results underscore findings on cooperative learning emphasizing individual accountability as an important role in group problem solving, common ground seemed to have no effect in our setting. As the collaborative condition and the PC condition did not differ in regard to shared understanding of literature selection criteria, we need to question whether the advantages in group achievement of the PC condition were due to common ground. Calling for other possible factors, one candidate might be the increased comparability of group members' individual contributions. The external representations used in the PC condition did not only show collated rating averages after the first stage, but also made differences between individual ratings visible. In providing valuable cues for groups to compare and support negotiation, different instructions for evoking comparability seems a promising approach for future research. This accounts not only for learning scenarios involving libraries, but for learning environments in general.

## References

- Baker, M., Hansen, T. G. B., Joiner, R., & Traum, D. (1999). The role of grounding in collaborative learning tasks. In P. Dillenbourg (Ed.), *Collaborative Learning: Cognitive and Computational Approaches* (1st ed., pp. 31–63). Amsterdam ; New York: Elsevier Science.
- Gabbard, R. B., Kaiser, A., & Kaunelis, D. (2007). Redesigning a library space for collaborative learning. *Computers in Libraries*, 27(5), 6–11.
- Jermann, P., & Dillenbourg, P. (2008). Group mirrors to support interaction regulation in collaborative problem solving. *Computers & Education*, 51(1), 279–296. doi:10.1016/j.compedu.2007.05.012
- Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The State of Cooperative Learning in Postsecondary and Professional Settings. *Educational Psychology Review*, 19(1), 15–29. doi:10.1007/s10648-006-9038-8
- Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology*, 55, 623–55. doi:10.1146/annurev.psych.55.090902.142009

- Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *The Journal of Applied Psychology, 85*(2), 273–83. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10783543>
- Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of Small-Group Learning on Undergraduates in Science, Mathematics, Engineering, and Technology: A Meta-Analysis. *Review of Educational Research, 69*(1), 21–51. doi:10.3102/00346543069001021