Task Complexity as a Driver for Collaborative Learning Efficiency: The Collective Working-Memory Effect

Femke Kirschner\textsuperscript{1,2*}, Fred Paas\textsuperscript{1,2}, and Paul A. Kirschner\textsuperscript{1}

\textsuperscript{1}Centre for Learning Sciences and Technologies, Open University of the Netherlands, P.O. Box 2960, 6401 DL Heerlen, the Netherlands.

\textsuperscript{2}Institute of Psychology, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands.

kirschner@fsw.eur.nl / paas@fsw.eur.nl / paul.kirschner@ou.nl

*Correspondence concerning this article can be sent to Femke Kirschner, Institute of Psychology, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands.
Telephone: +31 10 4088845, Fax: +31 10 408 9009, E-mail address: kirschner@fsw.eur.nl
Abstract

This study investigated the differential effects of learning task complexity on both learning process and outcome efficiency of 83 individual and group learners in the domain of biology. Based upon cognitive load theory, it was expected that for high-complexity tasks, group members would learn in a more efficient way than individual learners, while for low-complexity tasks, individual learning would be more efficient. This interaction hypothesis was confirmed, supporting our premise that the learning efficiency of group members and individuals is determined by a trade-off between the group’s advantage of dividing information processing amongst the collective working memories of the group members and its disadvantage in terms of associated costs of information communication and action coordination.

Keywords: Collaboration, Learning efficiency, Task complexity, Instructional design, Cognitive load.
Task Complexity and Collaborative Learning Efficiency

Task Complexity as a Driver for Collaborative Learning Efficiency: The Collective Working-Memory Effect

Although collaborative learning is a popular and widely used educational approach, research on its effects on learning has been inconclusive (Kester & Paas, 2005). Research stressing the potential of collaborative learning shows that collaborative learning environments can stimulate and enable learners to engage in activities that are valuable for learning (Beers, Boshuizen, & Kirschner, 2007; P. A. Kirschner, Beers, Boshuizen, & Gijselaers, 2008; Munneke, Andriessen, Kanselaar, & Kirschner, 2007). However, looking at the studies showing mixed and even negative findings regarding the learning process (e.g., Gregor & Cuskelly 1994; Heath, 1998), group forming, and group dynamics (e.g., Hughes & Hewson 1998; Taha & Caldwell, 1993), collaborative learning environments seem to be no guarantee for positive learning outcomes (Soller, 2001).

In a review of research comparing the effectiveness of individual learning approaches to collaborative learning approaches, F. Kirschner, Paas, and Kirschner (2009a) argued that one possible cause for these inconclusive results may be that the structures constituting cognitive architecture have not been systematically considered when designing and carrying out research on collaborative learning (see also, P. A. Kirschner, Sweller, & Clark, 2006). Specifically, the differing complexities of learning tasks used and the concomitant load imposed by these tasks on the learner’s cognitive architecture could be an important factor affecting the efficiency of individual and group learning, which is defined here as the amount of mental effort invested by a learner to reach a certain level of performance; high efficiency is associated with a relatively high test performance in combination with a relatively low mental-effort investment. In contrast, low efficiency is associated with a relatively low test performance in combination with a relatively high mental-effort investment.
Two recent studies investigated the efficiency of individual versus group learning as a function of task complexity (i.e., determined by the number of elements in a learning task and the interaction between those elements; Sweller & Chandler, 1994). In a study of high school students learning from solving biology problems, F. Kirschner, Paas, and Kirschner (2009b) investigated the hypothesis that, in contrast to individual learners, group members would be able to share the high cognitive load imposed by complex learning tasks, leaving more processing capacity available for constructing high quality cognitive schemas. This hypothesis was confirmed when the results showed a more favorable relationship between mental effort and performance on a transfer test for students who learned in groups than for students who learned individually. Another study by F. Kirschner, Paas, and Kirschner (2009c) which looked at high school students learning from solving mathematics problems, found a more favorable relationship between effort and performance in the learning phase for students who worked in a group than for students who worked individually with high-complexity tasks, but not with low-complexity tasks. The former study focused solely on the learning process and outcome of high-complexity learning tasks and consequently did not allow conclusions about the effects of low-complexity learning tasks. The latter study focused solely on the learning process while studying low and high-complexity learning tasks and consequently did not allow conclusions about the effects of task complexity on learning outcomes. The present study was designed to provide a more complete picture by investigating the effects of low- and high-complexity learning tasks on the efficiency of both the individual and group learning process and learning outcome.

The theoretical framework of cognitive load theory (CLT: P. A. Kirschner, 2002; Paas, Renkl, & Sweller, 2003, 2004; Sweller, 1988; Sweller, Van Merriënboer, & Paas, 1998) can be used to analyze how the complexity of a task can inform the design of efficient group-based learning environments. According to CLT, individual learning depends on the limited processing
capacity of the learner’s cognitive architecture and the cognitive load imposed by a task. The cognitive architecture consists of an effectively unlimited long-term memory (LTM) which interacts with a working memory (WM) that is very limited in both capacity (Baddeley & Hitch, 1974; Miller, 1956) and duration (Peterson & Peterson, 1959). For new, yet to be learned information, processing capacity is limited to only 4±1 information elements, and if not rehearsed, the information is lost within 30 seconds (Cowan, 2001). LTM contains cognitive schemas that store and organize knowledge by incorporating multiple elements of information into a single element (also referred to as chunking; Chase & Simon, 1973; Miller, 1956; Simon, 1974) with a specific function (i.e., learning). If the learning process has occurred over a long period of time, eventual schemas may consist of huge amounts of information. Because a schema can be treated by WM as a single element or even bypass WM if it has become sufficiently automated after long and consistent practice, the limitations of WM disappear for more knowledgeable learners dealing with previously learned information stored in LTM.

The cognitive load experienced by learners working on a learning task can be caused by the intrinsic nature of the task or the way that the information in the task is presented (Sweller et al., 1998). Intrinsic load is imposed by the number of interactive information elements in a task and the interactions between the elements. The more elements within a learning task and the more interactions between them, the higher the experienced intrinsic cognitive load will be. The manner in which the information is presented to learners can impose a load which is either ineffective for learning (i.e., extraneous cognitive load) or effective for learning (i.e., germane cognitive load). Intrinsic, extraneous, and germane load are additive, thus for learning to occur (i.e., schema construction and automation) it is important to take into account that the total cognitive load associated with an instructional design (i.e., the sum of the three separate loads) should stay within WM limits (Paas, Tuovinen, Tabbers, & Van Gerven, 2003).
To facilitate changes in LTM associated with learning, the focus of CLT has been on overcoming individual WM-limitations by introducing instructional manipulations in individual learning settings that are compatible with the human cognitive architecture (Ayres & Paas, 2009). Recently, group or collaborative learning has become recognized as an alternative way of overcoming individual WM limitations (F. Kirschner et al., 2009a, 2009b), in the sense that groups of collaborative learners can be considered as information processing systems (Hinsz, Tindale, & Vollrath, 1997; Tindale & Kameda, 2000; Ickes & Gonzalez, 1994), consisting of multiple limited WMs which can create a collective working space. Within these systems, valuable task-relevant information and knowledge held by each group member is consciously and actively shared (i.e., retrieving and explicating information), discussed (i.e., encoding and elaborating information) and remembered (i.e., personalizing and storing information) (Hinsz et al., 1997; Tindale & Kameda, 2000; Tindale & Sheffy, 2002). As long as the information is communicated between the group members and their actions are coordinated, the information elements within the task and the associated cognitive load caused by the intrinsic nature of the task can be divided across a larger reservoir of cognitive capacity (F. Kirschner et al., 2009a; Ohtsubo, 2005; Stasser, Stewart, & Wittenbaum, 1995). In terms of CLT, this has two conflicting consequences. First, collaborating individuals can invest less cognitive effort as compared to individual learners, because of a *distribution advantage*. Second, collaborating individuals need to invest cognitive effort in communicating information with each other and the coordination of their actions (i.e., transaction costs: Ciborra & Olson, 1988; F. Kirschner et al., 2009b; Yamane, 1996), which individuals working alone do not have to exert.

Previous research comparing collaborative learning to individual learning provides some support for the idea that the complexity of a task and the associated cognitive load are important factors for determining learning efficiency. More specifically, the research suggests that group
learning is superior to individual learning for relatively complex problem-solving tasks (e.g., F. Kirschner et al., 2009b; Laughlin, Bonner, & Miner, 2002; Laughlin, Hatch, Silver, & Boh, 2006), and that individual learning is superior to group learning for relatively simple recall tasks (e.g., Andersson & Rönnberg, 1995; Meudell, Hitch, & Kirby, 1992; Weldon & Bellinger, 1997).

For high-complexity tasks, research has shown that when learners have to work with the information elements relevant for carrying out the task, relate them to each other, and by doing so come up with a solution to a problem, groups perform better than individuals or nominal groups (i.e., fictitious groups formed by pooling the non-redundant performances of individuals) (Laughlin et al., 2006; Ohtsubo, 2005). Under such conditions, participating in a group facilitates the performance of individual group members. Based on CLT, one can argue that by dividing the high intrinsic cognitive load imposed by the complex learning task across learners, the risk of exceeding the limits of the WMs of the individual group members is reduced (i.e., distribution advantage). Although the additional cognitive load imposed by communicating information and coordinating actions (i.e., transaction costs) must be taken into account, this load can be considered to be relatively low compared to the distribution advantage for complex tasks. Consequently, the learning process will be more efficient for group members, allowing them to construct higher quality schemas in LTM than individual learners who must process all of the information individually. This collective working-memory effect was confirmed in an experiment comparing the effects of group learning and individual learning from complex cognitive tasks on the efficiency of cognitive schema construction (F. Kirschner et al., 2009b). The collective WM effect holds that group members can make use of each others’ WM capacity by sharing the cognitive load imposed by a task, to process information elements deeply and construct higher quality schemas in their LTM than learners working individually.
For low-complexity tasks, research on fairly simple recall tasks has shown that groups performed worse than nominal groups, indicating that collaborative learning was detrimental to the performance of individual group members (e.g., Meudell et al., 1992; Stephenson, Clark, & Wade, 1986; Weldon & Bellinger, 1997). Based on CLT one can argue that with low-complexity tasks, individual learners have sufficient capacity to carry out the tasks alone and no advantage of learning together is expected. Dividing the information among the group members means that they must communicate information and coordinate their actions, which for low-complexity tasks imposes a relatively high load (in relation to the benefits that will be accrued) thereby negating the distribution advantage. If the transaction costs are minimized, the learning process for groups will be as efficient as that for individuals while if the transaction costs are high, groups can be expected to learn less efficiently than individuals. This was confirmed in an experiment comparing the effects of group and individual learning from low-complexity and high-complexity learning tasks on the efficiency of the learning process (F. Kirschner et al., 2009c). For low-complexity tasks, the limited WM capacity of the individual learner and the expanded processing capacity of the group were equally effective for learning. For high-complexity tasks the group members could profit from distributing the high intrinsic cognitive load, as indicated by lower mental effort and higher learning performance (i.e., more efficient learning process) than individuals. Although the results on transfer-test efficiency showed a significant advantage of learning in a group above individual learning, the results did not allow the drawing of unequivocal conclusions about the differential effect of task complexity.

The present study attempted to determine the effects of low- and high-complexity learning tasks on the efficiency of the process and outcome of learning individually and in a group. With regard to learning from high-complexity tasks, the advantage of being able to divide the processing of information among group members was expected to be larger than the
disadvantage of investing additional effort related to the transaction costs. Therefore, the learning process was expected to be more efficient (i.e., higher learning performance with lower mental effort) for participants who learned in a group than for participants who learned individually. Consequently, group members were expected to be able to develop higher quality schemas than individual learners, which would be manifested in more efficient learning outcome (i.e., higher transfer performance with lower mental effort) for those who learn in a group.

With regard to learning from low-complexity tasks, learners were expected to have sufficient cognitive capacity to process the information individually. Therefore, both groups and individuals are expected to perform equally well, but the advantage of being able to divide the information processing among group members is expected to be smaller than the disadvantage of having to invest additional effort associated with information communication and action coordination. Therefore, the learning process was expected to be more efficient for individuals than for group members (i.e., equal learning performance with lower mental effort). As a result, individual learners were expected to be able to develop higher quality schemas than group members, which would manifest itself in a more efficient learning outcome for those who learned individually (i.e., higher transfer performance with lower mental effort).

**Method**

*Participants*

Participants were 83 Dutch high school sophomore students (46 girls, 37 boys) with an average age of 15.52 years (SD = 0.67). They participated in the study as part of their biology curriculum. No differences in prior knowledge were expected because all participants had followed the same biology courses in the previous three years and the topic of this study (i.e., heredity) was new to them. In addition, participants were randomly assigned to the different
experimental conditions to exclude possible differences in prior knowledge. Participants did not receive any academic or financial compensation for their participation.

**Materials**

Materials were in a domain of biology concerned with heredity, specifically the passing of both genotypic and phenotypic biological traits from parents to offspring through genes which carry biological information (e.g., eye color in humans, fur length in dogs, leaf shape in plants). In this domain, a general introduction containing necessary concepts and instruction on how to solve a cross-breeding problem in the form of a worked out example, simple and complex problem-solving tasks, and transfer tasks were designed. Two biology teachers approved the materials as suitable for the participants. All materials were paper based.

*General Introduction.* Relevant terminology, rules and theory underlying heredity, as well as a worked out example on solving heredity problems were presented in the introduction. It provided participants with the definition of genes, information on an organism’s genotype and phenotype, homozygosity or heterozygosity of an organism’s dominant or recessive genes, the pedigree chart, and the rules concerning the Punnett square (i.e., a diagram for predicting the outcome of a particular cross-breeding experiment). The worked out example demonstrated how to solve a heredity problem by combining relevant terminology, rules, and theory.

*Learning tasks.* The learning tasks for participants learning individually (i.e., individual learning condition) or in collaboration (i.e., collaborative learning condition) were the same. Low-complexity and high-complexity learning tasks were used, both consisting of a certain number of information elements concerning a biological trait in a family (i.e., ear shape of a guinea pig family, thumb length of a human family, fur color of a rabbit family) and two questions about the proportion of possible genotypes of the offspring and/or the specific genotype and phenotype of a family member. Participants had to combine the information elements that
they were given and solve the problem (i.e., correctly answer the questions). Each piece of information provided in a learning task was relevant but insufficient by itself for solving the problem. Successful problem solution entailed combining it with the other information elements. In the domain of heredity this, for instance, could mean that information element 1 is the mother’s eye color: blue; element 2 the father’s eye color: brown; and element 3 the dominance of brown eyes over blue eyes. To answer the question as to what the eye color of the offspring will be, the participant(s) need to combine all three pieces of information. In both the individual and group learning conditions all information elements needed to successfully solve the problem, were presented in a booklet (see Figure 1 for an example).

Task complexity or intrinsic cognitive load was determined by using Sweller and Chandler’s (1994) method based on the number of interactive task elements. Low-complexity tasks contained three information elements and high-complexity tasks contained nine elements that had to be combined to solve a heredity problem (see Figure 1a for an example of a low-complexity task, and 1b for an example of a high-complexity task). The low-complexity and high-complexity learning tasks consisted of three series of tasks connected to three main family-trait combinations. To keep the amount of practice of terminology, rules and theory concerning heredity equal in both conditions, the number of learning tasks in a series differed. In the low-complexity condition, each series consisted of four learning tasks which were identical except for the family members’ names. In the high-complexity condition, each series consisted of two learning tasks, which were identical except for the family members’ names. Thus, there was a total of 12 low-complexity learning tasks consisting of 3 information elements per task and 6 high-complexity learning tasks consisting of 9 information elements per learning task.

<Insert Figure 1 about here>
Transfer-test tasks. To determine how well participants were able to apply the knowledge and skills acquired in the learning phase (i.e. the quality of the acquired schemas) to new problem-solving situations, six transfer-test tasks were designed. Although the same basic terminology, rules, and underlying heredity-theory had to be used to solve these problems, the transfer tasks differed from the learning tasks with respect to the families and traits used, the kind of information elements given, the structure and the questions asked (e.g., genealogical tree, X-chromosome linked inheritance, dihybrid crossings, etcetera; see Figure 1c for an example).

Cognitive-load measurement. To measure the participants’ cognitive load after each task in the learning and test phase, the subjective 9-point cognitive-load rating scale developed by Paas (1992) was used. Participants were asked to rate how much effort it took them to solve a problem on a scale ranging from very, very low effort (1) to very, very high effort (9). This cognitive load measure has been used in numerous studies dealing with cognitive load and has proven to be non intrusive, valid and reliable (Paas, Van Merriënboer, & Adam, 1994).

Performance measurement. Successfully completing the learning and test tasks meant correctly answering two questions per task on the heredity characteristics of a certain trait in a family. Every question could be scored on multiple elements, with 1 point for a correct element and 0 points for an incorrect or not included element. This resulted in a maximum score of 40 points that could be earned for the 12 low-complexity learning tasks, 32 points for the 6 high-complexity learning tasks, and 24 points for the 6 transfer-test tasks. The minimum score for all tasks was zero. For the statistical analysis, the performance scores on learning and test tasks were transformed into proportions by dividing the participant’s score on the 12 low-complexity learning tasks, the 6 high-complexity learning tasks, and the 6 test tasks by the maximum score.

Efficiency measurement. The combination of performance and cognitive load measures provides a reliable estimate of the efficiency of instructional methods, both in terms of learning
process and learning outcomes. Paas and van Merriënboer’s (1993; see Van Gog & Paas, 2008) computational approach was used to calculate efficiency as a function of task complexity. The basic idea underlying this approach is that instructional conditions requiring less learning effort with equal or higher performance are more efficient than instructional conditions requiring more learning effort with equal or lower performance.

Regarding the learning process, learning efficiency (i.e., the mental efficiency while learning) was calculated by standardizing each participant’s scores for the low and high learning task performance and mental effort invested. For this purpose, the grand mean was subtracted from each score and the result was divided by the overall standard deviation, yielding z-scores for effort \( R \) and performance \( P \). Finally, a performance efficiency score, \( E \), was computed for each participant using the formula: \( E = [(P - R)/2^{1/2}] \). High learning efficiency was indicated by relatively high learning task performance in combination with relatively low mental-effort rating while low learning efficiency was indicated by relatively low learning task performance combined with relatively high mental-effort rating. Regarding efficiency learning outcome efficiency, the same computational approach was used, although this time by standardizing each the participant’s scores for performance and mental effort on the transfer-test tasks. High learning efficiency was indicated by relatively high test task performance in combination with relatively low mental-effort rating and low test efficiency by relatively low test task performance in combination with relatively high mental-effort rating.

\textit{Time-on-task.} During the learning phase, time on the task series was fixed to 10 min and managed by a proctor. The test phase had to be finalized within a maximum of 90 min. The specific time spent on each test task was recorded by the participants themselves with the aid of a digital clock.
Design and Procedure

Participants were randomly assigned to the four conditions with 21 learning individually from low-complexity learning tasks, 20 learning individually from high-complexity learning tasks, 21 learning in 3-person groups from low-complexity learning tasks, and 21 learning in 3-person groups from high-complexity learning tasks. All participants were guided through the learning and test phase by a proctor, who read the instructions on the procedure and rules out loud. At the start of the learning phase, all participants had to individually study the general introduction for 15 min. Participants were then assigned to one of the four conditions to carry out the first series of problem-solving tasks in 10 min. After each task in the series, independent of condition, all participants had to rate the amount of invested mental effort on the 9-point cognitive load rating scale. The second and third series of learning tasks followed the same procedure. The instructions given to all participants preceding the series of learning tasks consisted of reading all information elements thoroughly, reading the questions, and finally trying to answer the questions as correctly as possible using all information elements. For participants in the group condition, it was also stressed that working together was necessary for solving the problem. Face-to-face communication was very important, but transaction costs were minimized so as not to cause an overload from extraneous load. Therefore, participants were only permitted to communicate about task related topics. To prevent participants from offloading information in their WM (Scaife & Rogers, 1996), both individuals and group members were not allowed to write anything down. After this learning phase, the test phase required all participants to work individually for a maximum of 90 min on six transfer tasks. The amount of invested mental effort was measured after each transfer task using the same 9-point cognitive load rating scale as in the learning phase. Use of pen and paper was permitted in this phase.
Results

The data were analyzed with 2 (learning condition: individual vs. group) by 2 (task complexity: low vs. high) between-subjects multivariate analyses of variance (MANOVAs). The results for the learning and test phases are described separately. Dependent variables for the learning phase were performance, mental effort and learning efficiency. For the test phase, performance, mental effort, test efficiency and time-on-test-task were the dependent variables. A significance level of .05 was used for all analyses. Data from two participants was incomplete due to registration problems in the test phase. Case-wise deletion of those participants was used to analyze the data. Table 1 shows the means and standard deviations for performance, mental effort, learning efficiency, test efficiency and time-on-test-task in the learning and test phases as a function of learning condition and task complexity. Cohen’s $f$ is used as a measure of effect size, with $f$-values of .10, .25, and .40, corresponding to small, medium, and large effects, respectively (Cohen, 1988).

Learning phase

The omnibus MANOVA comparing learning condition (individual vs. group) and task complexity (low vs. high) indicated a main effect for learning condition, multivariate Wilks’ Lambda = .71, $F(2, 78) = 16.02, p < .001$, and task complexity, multivariate Wilks’ Lambda = .23, $F(2, 78) = 129.42, p < .001$. In addition there was a significant interaction between learning condition and complexity, multivariate Wilks’ Lambda = .91, $F(2, 78) = 3.99, p < .05$. Univariate ANOVAs were conducted to further examine these significant effects.

The ANOVA for learning performance revealed a significant main effect for learning condition with participants learning in groups outperforming participants learning individually, $F(1, 79) = 17.69, MSE = 0.03, p < .001, f = 2.15$. The main effect for task complexity was significant with participants having a higher performance score on low-complexity tasks than on
high-complexity tasks, $F(1, 79) = 111.06, \text{MSE} = 0.03, p < .001, f = 1.03$. In addition, it revealed a significant interaction between learning condition and task complexity, $F(1, 79) = 6.93, \text{MSE} = 0.03, p < .02, f = 0.18$. To determine the nature of this ordinal interaction, a post hoc analysis of simple effects was conducted using Tukey-Kramers method of multiple comparisons with unequal sample sizes (honestly significant difference [HSD] = 0.15, $\alpha = .05$). This analysis indicated that on low-complexity learning tasks participants learning in groups and participants learning individually performed equally well, but that on high-complexity tasks participants learning in groups significantly outperformed participants learning individually.

The ANOVA on the perceived amount of mental effort invested in solving the learning tasks revealed a main effect for learning condition, $F(1, 79) = 27.28, \text{MSE} = 1.03, p < .001, f = 0.29$, indicating that participants learning in groups invested lower amounts of mental effort than participants learning individually. The main effect of task complexity was significant, $F(1, 79) = 240.96, \text{MSE} = 1.03, p < .001, f = 1.48$, indicating that participants invested less mental effort on low-complexity tasks than on high-complexity tasks. It also revealed a significant ordinal interaction between learning condition and task complexity, $F(1, 79) = 4.23, \text{MSE} = 1.03, p < .05, f = 0.11$. To determine the nature of this ordinal interaction, a post hoc analysis of simple effects was conducted using Tukey-Kramers method of multiple comparisons with unequal sample sizes (HSD = 0.84, $\alpha = .05$). Whereas there was no difference in perceived amount of invested mental effort in learning from low-complexity tasks between participants that learned in groups and those who learned individually, group members reported investing significantly less mental effort than individuals in learning from high-complexity problems.

For learning efficiency, the ANOVA revealed a main effect for learning condition, $F(1, 79) = 30.54, \text{MSE} = 0.43, p < .001, f = 0.31$, indicating that participants learning in a group
had higher efficiency scores than participants learning individually. The main effect for task-complexity was significant, $F(1, 79) = 228.53$, $MSE = 0.43$, $p < .001$, $f = 1.39$, indicating that participants performing low-complexity tasks learned more efficiently than those performing high-complexity tasks. As expected, the learning condition × task complexity interaction was significant, $F(1, 79) = 7.99$, $MSE = 0.43$, $p < .02$, $f = 0.15$. To determine the nature of this ordinal interaction, a post hoc analysis of simple effects was conducted using Tukey-Kramers method of multiple comparisons with unequal sample sizes (HSD = 0.56, $\alpha = .05$). While learning efficiency for low-complexity tasks did not differ between participants learning in groups and learning individually, a significantly more favorable relationship between invested mental effort and performance was found for participants learning in groups than for participants learning individually for high-complexity tasks (see Figure 2).

Test phase

The omnibus MANOVA comparing learning condition (individual vs. group) and task complexity (low vs. high) did not reveal a significant main effect for learning condition, multivariate Wilks’ Lambda = .97, $F(2, 78) = 16.02$, ns. It did, however, reveal a significant main effect for task complexity, multivariate Wilks’ Lambda = .78, $F(2, 78) = 129.42$, $p < .001$, and a significant interaction between learning condition and task complexity, multivariate Wilks’ Lambda = .90, $F(2, 78) = 3.99$, $p < .05$. Again, univariate ANOVAs were conducted on the dependent variables to further examine these significant effects.

The ANOVA for test performance did not yield a significant main effect for learning condition, $F(1, 77) = 1.24$, $MSE = 0.04$, ns. The main effect for task complexity was significant, $F(1, 77) = 14.84$, $MSE = 0.04$, $p < .001$, $f = 0.14$, indicating that participants who had learned from the low-complexity tasks performed better than participants who had learned from high-
complexity tasks. The analysis did not reveal an interaction between learning condition and task complexity, $F(1, 77) = 1.43, MSE = 0.04, ns$.

For the perceived amount of mental effort invested in solving test problems, the ANOVA did not reveal a main effect for learning condition, $F(1, 77) < 1, MSE = 2.21, ns$, but did reveal a main effect for task complexity, $F(1, 77) = 7.107, MSE = 2.21, p < .01, f = 0.29$. The main effect was qualified by the crossover interaction effect of learning condition and task complexity, $F(1, 77) = 6.27, MSE = 2.21, p < .02, f = 0.27$, reflecting that the participants who had learned individually from low-complexity tasks perceived investing less mental effort for the test tasks than those who had learned from those tasks in a group. The reverse was found for participants who had learned from high-complexity problems; those who had learned in a group perceived investing less mental effort than participants who had learned individually.

The ANOVA performed on test efficiency scores did not reveal a main effect for learning condition, $F(1, 77) < 1, MSE = 1.25, ns$, but did reveal one for task complexity, $F(1, 77) = 14.40, MSE = 1.25, p < .001, f = 0.42$. Here too, the main effect was qualified by a crossover interaction effect between learning condition and task complexity, $F(1, 77) = 4.93, MSE = 1.25, p < .05, f = 0.23$, indicating that when participants had learned from low-complexity tasks individually, they carried out the test tasks more efficiently – as indicated by a more favorable relationship between test effort and test performance – than those who learned within groups. The reverse effect emerged for high-complexity tasks; those participants who had learned within groups performed the test tasks more efficiently than those who learned individually (see Figure 3).

The ANOVA performed on time-on-test-task did not yield a significant main effect for learning condition, $F(1, 77) = 1.15, MSE = 45.08, ns$, but did yield one for task complexity, $F(1, 77) = 4.60, MSE = 45.08, p < .05, f = 0.24$. The main effect was qualified by a crossover interaction between learning condition and task complexity, $F(1, 77) = 4.46, MSE = 45.08$,.
$p < .05, f^2 = 0.23$, indicating that participants who had learned individually from low-complexity tasks spent less time working on the test tasks than those who had learned from those tasks within groups. The reverse was found for high-complexity tasks where participants who had learned within a group spent less time working on the test tasks than those who had learned individually.

Discussion

Educators and researchers have expressed high expectations for the value and power of collaborative learning as educational approach, but results of research on its use have often been disappointing, not unequivocal with respect to student learning and performance, or both. A primary cause for this could be that the tasks presented to students in collaborative learning settings are not complex enough to necessitate working in teams and that this impedes both learning and performance. In trying to better understand the impact of task complexity on the effectiveness of collaborative learning environments, CLT was used as an indicator of the efficiency with which both groups and individuals learn. This efficiency is affected by the trade-off between the possibility to divide information processing amongst the WMs of the group members (i.e., collective working memory effect) and the associated costs of information communication and action coordination.

This study was designed to investigate the effects of task complexity on the efficiency of collaborative versus individual learning, both in terms of learning process and learning outcome. It was expected that when learning from high-complexity tasks, the advantage of being able to use each others’ WM through sharing cognitive load would be relatively larger than the disadvantage of having to invest mental effort in the communication of information and the coordination of actions (i.e., transaction costs). As a result, group members would process information elements more deeply and, thus, construct higher quality schemas in their LTM than
learners working individually. In contrast, when learning from low-complexity tasks, it was expected that the limited WM of an individual would be sufficient to process all information elements and therefore sharing the cognitive load within a group would not be advantageous for the quality of schema constructed. Working in collaboration on low-complexity tasks was even expected to have a detrimental effect on group member learning, for the ‘advantage’ of being able to make use of each others’ WM, would be outweighed by the relatively high mental effort associated with the transaction costs needed to do this. In conclusion, it was hypothesized that for high-complexity tasks, group members would learn in a more efficient way - both in terms of the learning process and outcomes - than individual learners, while for low-complexity tasks, individual learning would be more efficient.

The results largely confirmed this interaction hypothesis with respect to learning process efficiency. Although groups achieved higher performance, invested less mental effort and were, thus, more efficient than individuals when learning from the high complexity learning tasks, individuals performed equally well, invested the same amount of mental effort, and were equally efficient on the low complexity learning tasks. In other words, as expected, learners profited from being able to expand their limited WM capacity in the group when working on high complexity tasks. Furthermore, the limited WM capacity of individuals was apparently sufficient for working on the low-complexity tasks since both group members and individuals performed equally well on those learning tasks. However, unexpectedly, learners within a group did not experience a disadvantageous effect of having to invest the mental effort associated with the transaction costs when working on low-complexity tasks. This effect can be explained in two ways. The first explanation relates to the way that the learning tasks were designed. All problems had a single correct answer, all elements required for their solution were known, and the solution required using logical processes. Solving these well-structured problems (Kitchner, 1983) may have
resulted in learners paying minimal transaction costs to arrive at and agreeing on a common solution by combining and integrating individual ideas into a shared and collective one was not that difficult. The second explanation is related to the experimental design. The research was set up to restrain and, thus, minimize off-task communication. In a more natural setting, there might be significantly more discussion and, thus, significantly higher transaction costs. Future research should investigate the contribution of ill-structured tasks on the effects of task complexity on the efficiency of learning in a group as well as the effects of unrestrained communication between group members.

With respect to the efficiency of learning outcomes, the data confirmed the expectations. Participants who had learned individually from low complexity tasks performed equally well but invested less mental effort, and were therefore more efficient when carrying out the transfer-test tasks than participants who had learned from those tasks in a group. Although participants who had learned individually from high complexity tasks showed equal performance, they invested more mental effort and were, thus, less efficient when performing the transfer-test tasks, than those who had learned from those tasks in a group. In other words, with respect to the efficiency with which cognitive schemas were constructed during the learning process, learners profited from having learned from high-complexity tasks in collaboration, while learners profited from having learned from low-complexity tasks individually. This interaction effect increases when time-on-test is taken into account. Time-on-test has been identified as an important factor that, together with transfer-performance and mental effort, should be taken into account when determining the quality of the cognitive schemas constructed in the learning phase (Tuovinen & Paas, 2004; Van Gog & Paas, 2008). With respect to time-on-test, participants who had learned individually from low complexity tasks carried out the test tasks more quickly than those who
had learned in a group. Participants who had learned individually from high complexity tasks carried out the test tasks less quickly than participants who had learned in a group.

The results presented in this article are both a replication and an extension of two previous studies focusing on the efficiency of individual versus group learning as a function of task complexity. With regard to replicating previous findings, these results confirmed the hypothesis formulated and confirmed in F. Kirschner et al. (2009b) that when learning from high-complexity tasks, participants who learned in groups would demonstrate more efficient transfer performance (i.e., efficiency of the learning outcomes) than participants who learned individually. The result that the efficiency of learning (i.e., efficiency of the learning process) in a group was higher than the efficiency of learning individually when learning from high-complexity tasks, but not when learning from low-complexity tasks was also found in F. Kirschner et al. (2009c), although they used a different domain (i.e., mathematics) and the participants were younger. With regard to extending the results of previous studies, this study included both low-complexity and high-complexity tasks, focusing on the efficiency of both the learning process and the learning outcomes. This resulted in a more complete picture of the effects of task complexity on the efficiency of collaborative learning.

Another important extension of our previous studies is related to the conditions that were created for collaboration. In contrast to our previous studies, in which each group member was provided with unique information elements that had to be exchanged to solve the problems, in the present study each group member was provided with the same (i.e., all) information elements. Although, the latter condition could be argued to be less advantageous to the collaboration process, because group members are not forced to exchange information elements to solve the problems, the results regarding the relationship between high task complexity and individual/group learning were similar to those found in the previous studies, indicating that
learning in collaboration is more efficient than learning individually. In future research, it would be interesting to further investigate the effects of the need to collaborate on the relationship between task complexity and individual/group learning.

Within the context of CLT, the results presented here have implications for the design of certain types of collaborative learning environments, such as those making use of problem-based learning (Hmelo-Silver, Duncan, & Chinn, 2007; Schmidt, Loyens, Van Gog, & Paas, 2008) or project-based learning (Van Bruggen, Kirschner, & Jochems, 2002). One implication is that the tasks provided to these collaborative learning environments should be of such complexity that individual learners are not capable of successfully executing the task by themselves. From this perspective it is extremely important for teachers and instructional designers, that future research identifies the level of task complexity at which it becomes more efficient to assign tasks to groups than to individuals. All too often, students are given collaborative learning tasks that they could better and more efficiently carry out individually, often leading to minimal collaboration and/or stop-gap measures such as requiring students to post a minimal number of comments on a discussion board or send each other a minimal number of emails (P. A. Kirschner et al., 2008).

Having said this, we note that such a recommendation cannot be made in absolute terms since task complexity, which is defined as the number of interacting elements in a task and the degree of interaction between them, is relative to and dependent on the quality and quantity of schemas in the learner’s long term memory (i.e., their expertise; Paas et al., 2003). What is complex for the novice may be moderately complex for the more experienced learner and simple or even trivial for the expert. Therefore, comparable to the expertise-reversal effect (Kalyuga, Ayres, Chandler, & Sweller, 2003), one can expect that the task-complexity threshold at which group learning becomes more efficient than individual learning will vary as a function of learner
expertise. An interesting hypothesis for future research is that for group learning to be efficient, task complexity needs to be higher for more advanced learners and experts than for novices.

It should be noted that, although specific hypotheses in this study were formulated regarding the possible beneficial and/or deleterious effects of inter-individual communication and coordination of information (i.e., transaction costs), those activities were not explicitly monitored. Monitoring and analyzing communication and coordination activities would be a valuable addition in future research as it could be informative of the possibly different collaborative dynamics under different collaborative circumstances such as providing group members with all or just part of the necessary information, or providing them with worked examples or conventional problems. In addition, relating information on the amount and type of communication and coordination between group members to their subsequent performance could also be informative of the type of load imposed by these inter-individual activities (Janssen, Kirschner, Erkens, Kirschner, & Paas, 2010).

Finally, it should be noted that the learning conditions in this study can be considered artificial, for example, learners were not allowed to use pen and paper while solving the learning problems. In this sense, it is not clear to what extent the results can be generalized to real classroom settings. It can be assumed that the complex pattern of interactions between cognitive, motivational, and social factors that characterize a real life context would add ‘noise’ to the data and cause the effects to be less pronounced than in this study. We acknowledge that, ultimately, research on group-based learning requires an interrelated perspective integrating cognitive, motivational, and social aspects. However, to disentangle the contributions of each of these factors to learning processes and outcomes in group-based learning, they must be studied within tightly constrained experimental environments, one at a time, keeping all other aspects constant.
The present study showed that task complexity is a key factor in determining learning efficiency in individual and collaborative learning environments with strong implications for the design of such environments. However, the results need further experimental confirmation with tasks that are ill-structured, using learners with varying levels of expertise and age, and in more ecologically valid collaborative learning settings to provide more specific guidelines for designing group learning environments.
Task Complexity and Collaborative Learning Efficiency

References


Table 1

*Means and Standard Deviations of the Dependent Variables in the Learning and Test Phase as a Function of Learning Condition and Task Complexity*

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Learning condition</th>
<th>Individual</th>
<th></th>
<th></th>
<th>Group</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td></td>
<td>M</td>
<td>SD</td>
<td></td>
</tr>
<tr>
<td><strong>Learning performance</strong> a (0-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low complexity</td>
<td>0.87</td>
<td>0.21</td>
<td>0.93</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High complexity</td>
<td>0.34</td>
<td>0.18</td>
<td>0.61</td>
<td>0.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Learning mental effort</strong> (1-9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low complexity</td>
<td>2.35</td>
<td>0.85</td>
<td>1.64</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High complexity</td>
<td>6.27</td>
<td>1.14</td>
<td>4.64</td>
<td>1.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Learning efficiency</strong> c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low complexity</td>
<td>0.88</td>
<td>0.65</td>
<td>1.27</td>
<td>0.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High complexity</td>
<td>-1.72</td>
<td>0.71</td>
<td>-0.51</td>
<td>0.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test performance</strong> a (0-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low complexity</td>
<td>0.64</td>
<td>0.21</td>
<td>0.63</td>
<td>0.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High complexity</td>
<td>0.41</td>
<td>0.20</td>
<td>0.52</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test mental effort</strong> (1-9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low complexity</td>
<td>4.06</td>
<td>1.39</td>
<td>4.73</td>
<td>1.69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High complexity</td>
<td>5.77</td>
<td>1.28</td>
<td>4.78</td>
<td>1.54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Transfer efficiency</strong> c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low complexity</td>
<td>0.66</td>
<td>1.08</td>
<td>0.33</td>
<td>1.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Performance</td>
<td>Mental Effort</td>
<td>Performance</td>
<td>Mental Effort</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>---------------</td>
<td>-------------</td>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High complexity</td>
<td>-0.84</td>
<td>1.01</td>
<td>-0.06</td>
<td>1.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time on test (min.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low complexity</td>
<td>24.25</td>
<td>8.12</td>
<td>29.00</td>
<td>5.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High complexity</td>
<td>24.20</td>
<td>7.16</td>
<td>22.65</td>
<td>5.42</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Performance is the proportion of correct answers on the learning or test tasks.

b Based on the z-scores of mental effort and performance in the learning phase.

c Based on z-scores of mental effort and performance in the test phase.
Figure 1. A low (a) and high (b) complexity learning task, as well as a transfer-test task (c).
Figure 2. Learning phase: Learning Condition x Task Complexity interaction.
Figure 3. Test phase: Learning Condition x Task Complexity interaction.