British Journal of Educational Psychology (2005), 75, 223–237

The British Psychological Society
www.bpsjournals.co.uk

Optimizing the number of steps in learning tasks for complex skills

Rob J. Nadolski*, Paul A. Kirschner and Jeroen J. G. van Merriënboer
Open University of The Netherlands, The Netherlands

Background. Carrying out whole tasks is often too difficult for novice learners attempting to acquire complex skills. The common solution is to split up the tasks into a number of smaller steps. The number of steps must be optimized for efficient and effective learning.

Aim. The aim of the study is to investigate the relation between the number of steps provided to learners and the quality of their learning of complex skills. It is hypothesized that students receiving an optimized number of steps will learn better than those receiving either the whole task in only *one* step or those receiving a large number of steps.

Sample. Participants were 35 sophomore law students studying at Dutch universities, mean age $=$ 22.8 years ($SD = 3.5$), 63% were female.

Method. Participants were randomly assigned to 1 of 3 computer-delivered versions of a multimedia programme on how to prepare and carry out a law plea. The versions differed only in the number of learning steps provided. Videotaped plea-performance results were determined, various related learning measures were acquired and all computer actions were logged and analyzed.

Results. Participants exposed to an intermediate (i.e. optimized) number of steps outperformed all others on the compulsory learning task. No differences in performance on a transfer task were found. A high number of steps proved to be less efficient for carrying out the learning task.

Conclusions. An intermediate number of steps is the most effective, proving that the number of steps can be optimized for improving learning.

Learning in competency-based multimedia practicals takes place in a self-contained electronic learning environment. Such practicals provide context-relevant practice to students for attaining complex skills such as diagnosing diseases, literature searching, modelling stress-factors that cause burnout, or preparing a plea in court (Brown, Collins,

**Correspondence should be addressed to Rob J. Nadolski, Educational Technology Expertise Centre, Open University of The Netherlands, Valkenburgerweg 177, 6419 AT Heerlen, The Netherlands (e-mail: rob.nadolski@ou.nl).*

DOI:10.1348/000709904X22403

& Duguid, 1989; Nadolski, Kirschner, van Merriënboer, & Hummel, 2001; Westera & Sloep, 1998). These practicals are assumed to be instrumental in allowing learners to develop the cognitive schemata necessary for the performance and transfer of complex skills. Many researchers (e.g. Hannafin, Land, & Oliver, 1999; Jonassen, 1999; Mayer, 1999; Van Merriënboer, 1997) agree that transfer-oriented learning can best be achieved through the use of realistic learning tasks consisting of a task description, an authentic environment to carry out the task, and cognitive feedback on the quality of the task performance.

This, however, is often easier said than done. Realistic whole tasks are often too difficult for novice learners without some form of simplification. Learning tasks, therefore, need first to be designed to pedagogically model reality (Achtenhagen, 2001), and then further simplified through segmentation of the whole task into smaller task assignments or steps. For example, in Fig. 1, the whole task of 'preparing and pleading a case in court' has been segmented into nine meaningful task assignments. This systematic approach to problem solving (SAP) for the whole learning task is provided via a process worksheet (Van Merriënboer, 1997) in combination with references to the task assignments that guide the learners through the distinct steps (see Table 1). Trainers (beginning practitioners) play a key role in identifying the SAP since they themselves are not far removed from the target population. Their SAPs, acquired through the use of thinking-aloud protocols while solving the problem through a means-ends analysis, can (with relatively small changes) be used for ID purposes (Nadolski *et al.*, 2001). The SAP for the whole learning task consists of several phases and involves a domain-specific problem-solving strategy together with its associated heuristics. In other words, a process worksheet guides the learners through the problem-solving process of the whole learning task. The 2-part process is similar to what Achtenhagen (2001, p. 364) calls 'modelling a model of reality under a didactic perspective'. A step consists of a task assignment and carrying out its activities attributes to solve a phase. Following such steps does not automatically lead to a correct solution for the whole task, as is the case in following the steps of a procedure, but leads to a strategy that narrows the search space, and to the selection of the operators that most likely lead to a correct solution. Most tasks can be segmented into smaller steps, which then can be broken down again and again. This method of cognitive task analysis through means-ends analysis reveals steps and associated SAPs (see e.g. Nadolski *et al.*, 2001; Potter, Roth, Woods, & Elm, 2000; Zachary, Ryder, & Hicinbothom, in press), though not capable of giving a completely unequivocal answer to the level of detail needed, provides highly useful general answers. The question at this point is how to determine what the optimal number of steps is: when do you stop? The central question of this article is thus, how the number of steps in a learning task can be optimized and whether this will lead to better performance.

Optimizing the number of steps and providing an accompanying process worksheet brings learning tasks within the reach of the learners' capabilities. Tasks with too few steps are too difficult and mentally demanding for learners, which hampers learning and subsequent transfer. Learners may not accurately process the necessary information because they are overwhelmed by the difficulty of the task (i.e. cognitive overload), or they may revert to surface processing (i.e. superficial, non-meaningful learning) in order to keep their cognitive load within the threshold limit (Craik & Lockhart, 1972; Sternberg & Frensch, 1991). Tasks with too many steps may also hamper learning because of their incoherence caused by redundant information between steps and/or an excess of details. This, too, will make them too mentally

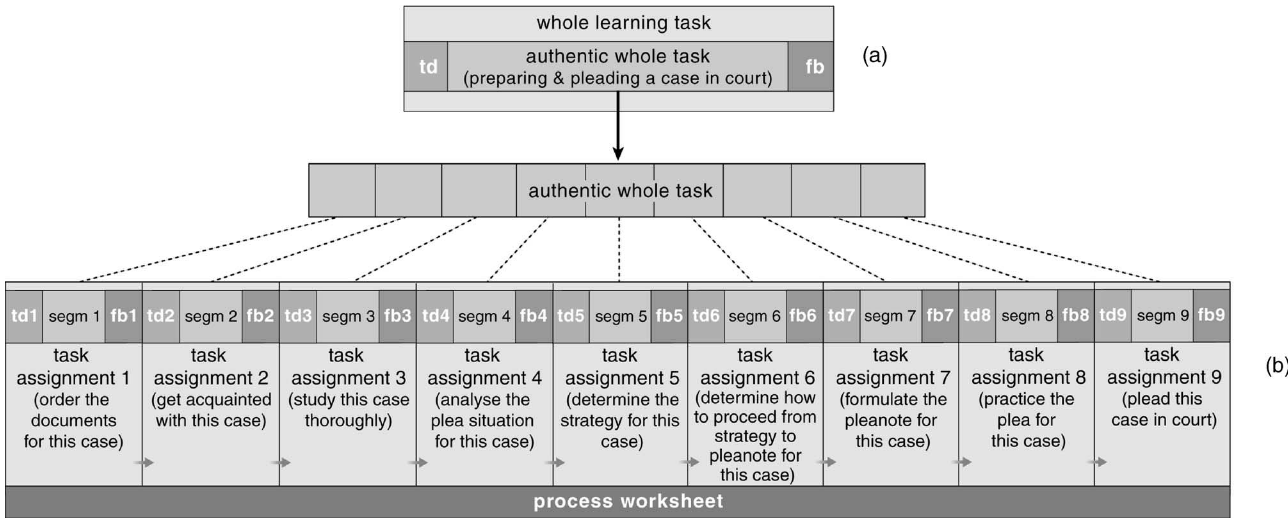

Figure 1. Authentic whole task, segmented in task assignments. (a) A non-segmented whole learning task consists of an authentic whole task preceded by task description (td) and followed up with cognitive feedback (fb). (b) Each task assignment (numbered 1–9) also consists of a task description, an authentic subtask (i.e. a segment from the authentic whole task) and cognitive feedback. Each task assignment (i.e. step) represents a phase in the problem solving process of the whole task. The process worksheet presents the phases in the problem solving process of the whole task and refers to the task assignments.

Table 1. A systematic approach to problem solving (SAP) for preparing and pleading case X in court

Subgoals (Phase)	Heuristics
1. Order documents for case X.	You might try to order the documents chronologically, categorically (e.g. legal documents, letters, notes), or by relevance.
2. Get acquainted with case X.	You might answer questions as, 'Which law subdomain is relevant here?' or, 'How do I estimate my client's chances?'
3. Study case X thoroughly.	You might answer questions such as, 'What is the specific legal question in this case?', 'What law sections are relevant in this case?' or, 'What legal consequence is best for my client?'
4. Analyze the situation for a plea for case X.	You might answer questions as, 'Which judge will try the case?', 'Where?', 'What time of the day?'
5. Determine the strategy for a plea for case X.	You might weigh the importance of the results of (3) and (4) with respect to your own capabilities (plea style) when deciding about aspects to include in your plea.
6. Determine how to proceed from plea-strategy to plea-note in case X.	You might write a concept plea-note in a certain format using results of (3) and (5) in spoken language, always keeping your goal in mind and using a certain style to express yourself.
7. Determine how to proceed from plea-note to plea in case X.	You might transform the plea-note into index cards and then practice this for yourself, paying attention to various presentation aspects (verbal and non-verbal behaviour).
8. Practice plea for case X.	You might ask friends to give feedback or video record your own attempts for self-evaluation.
9. Plead case X in court.	You might pay attention to the reactions of the various listeners.

demanding (Mayer & Moreno, 2002). In addition, learners may regard the steps as too specific for the learning task in question, preventing them from constructing the generalizations or abstract schemata necessary for learning transfer. Thus, like many other instructional design problems (see Clark, 1999), determining the number of steps is an instructional design problem that requires a solution through optimization.

The number of steps may directly influence cognitive processing and schema construction, but may also affect the amount of effort learners bring into the training situation (see Bonner, 1994). In our study, effort is operationalized by a combination of measures of motivation, time-on-task, and mental effort. The more effort invested, the better the result. However, this is not absolute. First, there is a limit above which a further increase of effort does not yield a further increase in performance. Increasing task motivation is not helpful for a learning task that is so difficult that it is impossible to solve. Investing more mental effort only leads to a better result if the task performer is not already cognitively overloaded by too much information. Second, above a certain threshold level, longer time-on-task also does not lead to a better result. Third, task motivation, mental effort and time-on-task may be irrelevant if a learning task is extremely easy. Finally, interaction effects might also occur where, for instance, higher motivation leads to a shorter time-on-task with a better result (Bonner, 1994). This means that we not only need to study learning results, but also the amount of effort (i.e. costs) invested in reaching those results when considering optimizing the number of steps. In our study, effort is determined by measuring task motivation (Bonner, 1994; Maynard & Hakel, 1997), subjectively experienced cognitive load (Paas & van Merriënboer, 1994), and time-on-task (Karweit, 1984).

Finally, several studies have shown the usefulness of not only determining learning effectiveness, but also the learning efficiency of instructional design measures (see Admiraal, Wubbels, & Pilot, 1999; Kalyuga, Chandler, & Sweller, 1998; Van Merriënboer, Schuurman, de Croock, & Paas, 2002). Although efficiency can be operationalized in many different ways, higher efficiency always indicates either equivalent results at lower costs, higher results at the same costs, or, ideally, higher results at lower costs. Therefore, when studying the effects of the number of steps, it is important to study both the task results (i.e. performance on the task) and the costs associated with reaching those results (based on motivation, cognitive load, and time-on-task) to determine the task efficiency.

Our study is designed to examine the effects of the number of steps on the performance and efficiency of learning tasks and transfer tasks. Three conditions (no steps, an intermediate number of steps, and a high number of steps) are compared for law students learning to prepare a plea in court. The first hypothesis is that an intermediate number of steps will lead to the highest performance for the learning tasks, while no significance differences between the other two conditions will be present. The second hypothesis is that a high number of steps will show the lowest transfer performance. The final hypothesis is that task efficiency will be lowest for a high number of steps and highest for an intermediate number of steps.

Method

Participants

Thirty-five students enrolled in the experiment and were randomly assigned to the three conditions (no steps; $N = 11$, intermediate number of steps; $N = 12$, high number of steps; $N = 12$). All participants (22 female, 13 male; mean age $= 22.8$ years, $SD = 3.5$) were sophomore law students studying at Dutch universities. None of the participants had prior plea experience. Comparability of students with respect to domain knowledge was assured since first-year law curricula of Dutch universities are practically identical both with respect to courses and textbooks. Comparability of plea experience was assured by a background questionnaire which will be discussed later.

Materials

Learning materials

The competency-based multimedia practical, *Preparing a plea* (Wöretshofer *et al.*, 2000) was adapted for this research. The goal of the practical is prepare students to carry out a plea in court. Students receive multiple-step whole task training using a high variability sequence of learning tasks (one compulsory learning task and two additional, non-compulsory learning tasks). Support is faded as the learners gain expertise, beginning with concrete modelling examples through working with tasks with process worksheets (see Nadolski *et al.*, 2001). The non-compulsory learning tasks enable variability of practice. Three versions of the practical were produced: one with no steps (a whole task = min condition), one with four steps (intermediate number of steps with feedback between the steps = int condition), and one with nine steps (high number of steps with feedback between the steps = max condition). For all versions, objective task complexity of the assignments (1 = *very simple*, 2 = *simple*, 3 = *complex*, 4 = *very complex*) was determined by 32 different respondents using the task-complexity instrument developed by Nadolski, Kirschner, van Merriënboer, and Wöretshofer (in press). The mean objective complexity for the tasks were for the min condition 3.6 ($SD = 0.7$) for the int condition 2.9 ($SD = 0.4$) and for the max condition 2.0 ($SD = 0.4$).

As indicated in Fig. 2, all versions first presented identical information and support tools. The tools included modelling examples of persons conducting a plea, a 'plea checker' to analyze pleas, discussions of ethical issues in pleading a case, numerous tips for communicability aspects in pleading a case, and judicial-procedural aspects of plea preparation. Participants had all necessary documents available while working on the learning tasks. The material for the learning tasks also contained process worksheets that differed for each condition. A specific task assignment was available for each step included in the process worksheet. The steps in the several versions are related. The task assignment in Step 1 of the version used for the intermediate condition, for example, aims at an outcome that is comparable to those for Steps 1–5 in the maximum condition. If two steps followed each other, the latter step always included cognitive feedback on the previous step (between-step support) and a new task assignment. All step information provided remained available for all following steps. After working on the learning task(s), participants received one transfer task. The material for the transfer task was the same in all conditions.

Tests

Two instruments were developed to measure the coherence, content, and communicability of participants' plea performance results on both the compulsory learning task and the transfer task. Two judges blindly and independently scored participants' videotaped pleas. All videotaped pleas were scored by both judges. Elements related to these three measures were rated on 4-point scales (0 = *poor*, 1 = *weak*, 2 = *sufficient*, 3 = *good*). Examples of items for coherence were: 'starts with a conclusion to be drawn from the plea', 'provides a content overview of the plea before going into details', 'indicates how successive arguments relate to relevant facts' and – for each major point – 'provides conclusion with respect to major point [x]'. Since the coherence of the plea is directly related to the number of points needed to be made by the 'pleader', the number of items differs for the two instruments. Coherence for learning task and transfer task was scored on seven and nine items, respectively (Cohen's κs = .7 and .6, respectively). Content for learning task was scored on 39 items, content related to transfer task was scored on 46 items (Cohen's κs .6 and .7, respectively), and communicability was scored on 14 items for both instruments (Cohen's κs .6 and .7, respectively). Existing plea-measurement instruments (e.g. Edens, Rink, & Smilde, 2000) were regarded as too general to be used here.

Also, a series of condition-specific questionnaires was used to gather data on:

(1) Prior plea experience (e.g. prior written and oral presentation skills, membership of debating club), computer literacy, attitude towards learning with computers, age, and gender.

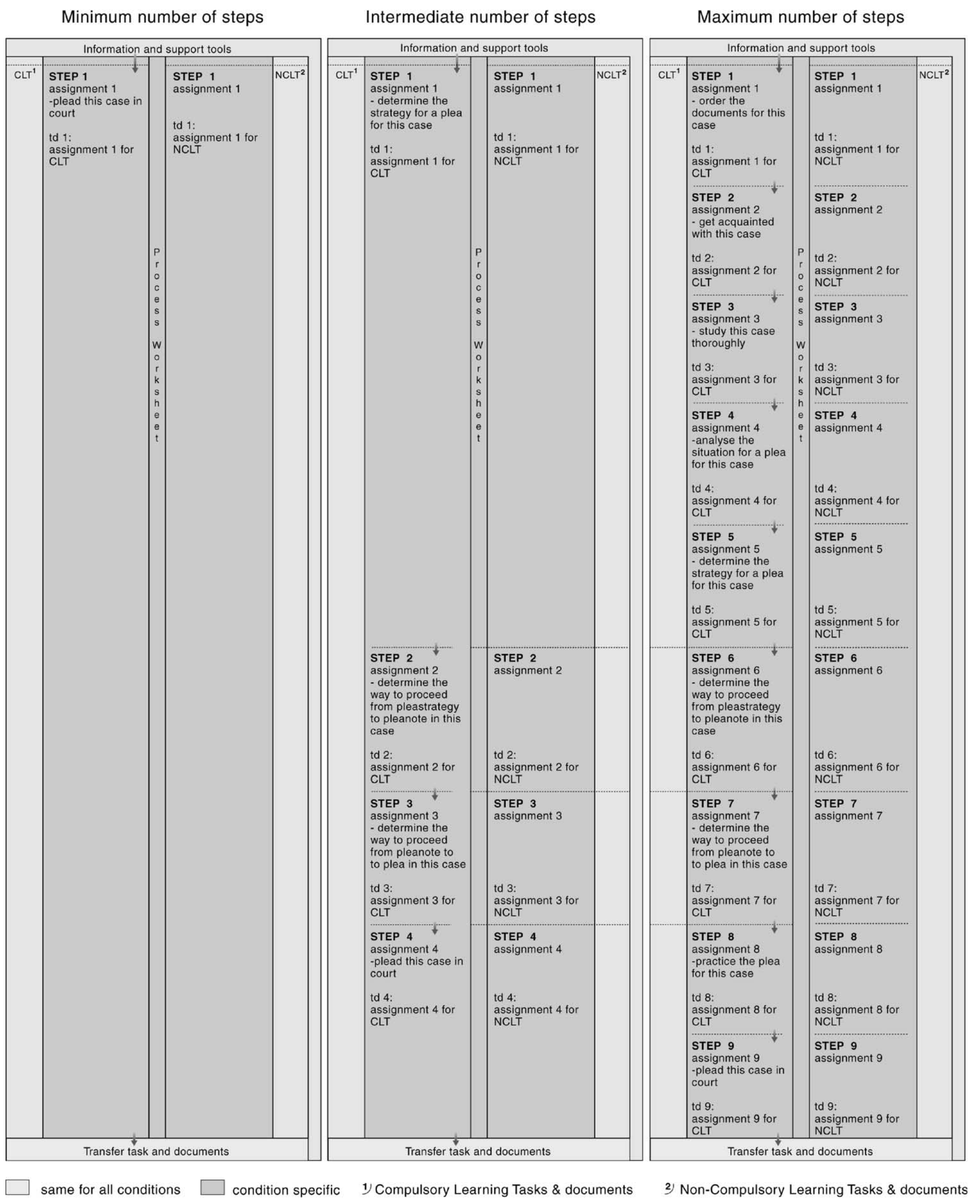

Figure 2. Overview of practical material for all conditions.

(2) Time-on-task per step.

(3) Mental effort. A 9-point categorical scale developed by Paas, van Merriënboer, and Adam (1994) was used to measure the experienced cognitive load of each of the constituting steps for the learning task (before and after each step, respectively), the learning task itself and the transfer task. Experienced cognitive load was used as an indication for mental effort; the less mentally demanding the task, the lower the cognitive load.

(4) Motivation per step. A 3-item 7-point Likert scale taken from Maynard and Hakel (1997) was used. The items were, ‘I was motivated to perform well on this task

assignment', 'This task assignment was interesting to me', 'I put a lot of effort into coming up with the best possible solution'.

(5) Satisfaction/perceived efficacy. Participants indicated their satisfaction with the quality of the instructional materials by giving their opinion on the adequacy of the number of steps, whether the goal of the practical ('I've learned to conduct a plea') was reached, the relevance of the supportive information, and the relevance of between-step support.

Finally, as all conditions were computer-delivered, all study behaviour was logged and analyzed.

Procedure

Before participating, (potential) participants were informed of the experiment, the practical (60 study hours in 1 month), and of the necessary prior knowledge and skills. Participants were randomly assigned to one of the three experimental conditions and were required to work individually. All learning materials (including instructions) and questionnaires were sent to the participants' home address. They were strongly advised to work step-by-step since the programme offered the possibility of skipping consecutive steps. Logging results showed that the participants did not skip any steps. Within a step, however, maximum learner control existed. In other words, participants were free to decide if and when to consult step specific information and how long they would like to spend working on the task assignment in this step. After 2 weeks (approximately 30 study hours), participants were required to hold their plea for the compulsory learning task, which was videotaped for later evaluation. About 2 weeks later, they were required to hold their plea for the transfer task, which was also videotaped. As it is legally required for lawyers in The Netherlands to submit a plea note when pleading a case in court, participants were required to include a plea note for both tasks. Participants were required to return the questionnaires in a stamped self-addressed envelope and to electronically send their logging results 1 week after completion of the course. Upon completion of the experiment, participants were thanked for their participation and received the promised remuneration for their participation (about US $120). Two judges blindly and independently scored participants' videotaped pleas.

Results

The data collected for determining computer literacy and attitude towards learning with computers showed no differences between conditions. The data collected for prior plea experience confirmed that none of the participants had prior plea experience and, thus, that the conditions were equivalent. Analysis of the logging results indicated that participants did not make use of the non-compulsory learning tasks.

Performance

Data means and standard deviations for performance on the pleas for the learning task and transfer task are presented in Table 2.

Analyses of variance showed a significant difference in the coherence of the learning task plea between the three conditions, $F(2, 32) = 3.48$, $MSE = .37$, $p < .05$, $\eta^2 = .18$.

Table 2. Performance on the learning and transfer tasks

	Min ($N = 11$)		Int ($N = 12$)		Max ($N = 12$)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Learning task						
Coherence	2.09	0.8	2.75*	0.5	2.33	0.5
Content	1.82	0.4	2.75**	0.5	2.17	0.6
Communicability	2.82	0.4	2.75	0.5	2.5	0.5
Transfer task						
Coherence	1.55	0.5	1.33	0.5	1.33	0.5
Content	0.91	0.5	1.08	0.5	0.83	0.6
Communicability	2.82	0.4	2.75	0.5	2.5	0.5

Notes. All performance variables were measured on a 4-point scale. (0 = poor, 1 = weak, 2 = sufficient, 3 = good).
*$p < .05$, **$p < .01$.

Contrast analyses, using Bonferroni's correction, revealed a significantly better coherence for the intermediate condition than for the minimal and maximum conditions, $t(32) = 2.5$, $p < .05$ (one-tailed). No significant difference was found between the minimal and maximum conditions, $t(32) = 1.0$, $p = .3$.

Analyses of variance also revealed a significant difference between the conditions with respect to the content of the learning task plea, $F(2, 32) = 10.87$, $MSE = .24$, $p < .01$, $\eta^2 = .41$. Contrast analyses, using Bonferroni's correction, revealed a significantly better result for the intermediate condition compared with the minimal and maximum conditions, $t(32) = 4.4$, $p < .01$ (one-tailed). No significant difference was found between the minimal and maximum conditions, $t(32) = 1.7$, $p = .1$. There were no differences in the communicability of the learning task plea.

Results on the coherence, content, and communicability of the plea for the transfer task did not show any significant differences between the conditions.

Time-on-task

Data means and standard deviations for the time spent on the learning and transfer tasks are presented in Table 3. A contrast test using Bonferroni's correction showed significantly more time-on-task for the learning task for the max condition than the combined int and min condition, $t(32) = 2.42$, $p < .05$ (one-tailed). There was no significant difference between the min and int condition, $t(21) = .18$, $p = .86$.

There were no significant differences on the time-on-task values for the transfer task between the conditions, although the expected pattern of an increasing time-on-task with an increasing number of steps, was observed.

Mental effort and motivation

The mental effort values for both the learning task and the transfer task did not show any differences between the conditions (see Table 4). In all conditions, both tasks demanded an average mental effort.

Table 3. Time-on-task (in minutes)

	Min		Int		Max	
	M	*SD*	*M*	*SD*	*M*	*SD*
Time-on-learning task	635	378	660	233	933*	368
Time-on-transfer task	325	104	382	228	473	557

Notes. Based upon self-report. All spent time was reported in multiples of 5 minutes. The transfer task was the same for all conditions; no support was given.
$*p < .05$, one-tailed.

Table 4. Mental effort and motivation on the learning and transfer tasks

	Min		Int		Max	
	M	*SD*	*M*	*SD*	*M*	*SD*
Learning task						
Mental effort	4.64	1.4	5.5	2.2	5.08	2.4
Motivation	5.79	0.9	5.75	0.9	6.14	0.8
Transfer task						
Mental effort	4.73	1.7	5.92	1.8	5.5	2.5
Motivation	5.33	1.6	5.47	0.9	6	0.9

Notes. Mental effort is measured on a 9-point categorical scale (Paas *et al*., 1994; 1 = very, very low mental effort, 9 = very, very high mental effort). Motivation is measured on a 3-item, 7-point categorical scale (Maynard & Hakel, 1997; 1 = very, very low, 7 = very, very high).

The motivation values for both the learning task and the transfer task did not differ between the conditions. In all conditions, participants were highly motivated.

Task efficiency

There were no significant correlations between mental effort, motivation, and time-on-task. Data means and standard deviations for learning task efficiency and transfer task efficiency are presented in Table 5.[1] A contrast test using Bonferroni's correction showed that the max condition was significantly less efficient on all three measures of quality of the learning task than the combined int and min condition (learning task coherence: $t(32) = 2.41$, $p < .05$ (one-tailed); learning task content: $t(32) = 2.44$, $p < .05$ (one-tailed); learning task communicability: $t(32) = 2.47, p < .05$(one-tailed)). There were no differences for learning task efficiency between the int and the min condition (learning task coherence, $t(21) = .15$, $p = .89$; learning task content: $t(21) = .10$, $p = .92$; learning task communicability: $t(21) = .27$, $p = .79$). Both

[1] *All efficiency measures were calculated using the procedure described by Tuovinen and Paas (2004) for determining instructional condition efficiency. They describe a 3-factor instructional condition efficiency which we extended to a 4-factor instructional condition efficiency here. In formula: 4-Factor Efficiency = (P-E-T-M)/SQRT (4), P = performance, E = mental effort, T = time-on-task, M = motivation. Students' scores on all factors are standardized (the total mean was subtracted from each score and the result was divided by the standard deviation), giving z scores for each factor.*

Table 5. Efficiency for the learning and transfer tasks

	Min		Int		Max	
	M	*SD*	*M*	*SD*	*M*	*SD*
Learning task						
Coherence	3.0	10.6	2.4	6.1	− 5.2*	10.3
Content	2.9	10.5	2.5	5.9	− 5.2*	10.3
Communicability	3.3	10.6	2.3	5.9	− 5.3*	10.3
Transfer task						
Coherence	2.8	3.5	0.7	6.6	− 1.9	15.7
Content	2.2	3.4	0.3	6.2	− 2.4	15.4
Communicability	2.3	3.4	0.3	6.3	− 2.4	15.6

Notes. $*p < .05$, one-tailed.

Table 6. Satisfaction and perceived efficacy of the training

	Min		Int		Max	
	M	*SD*	*M*	*SD*	*M*	*SD*
Number of steps adequacy[a]	3.82*	0.9	4.83	0.9	5	1.1
'I've learned to conduct a plea'[b]	5.45	1.4	5.67	1.1	5.42	1.1
Relevance of supportive info[c]	2.45	0.7	2.75	0.5	2.66	0.8
Between-step support[d]			2.05	0.4	2.30	0.6

Notes. [a]The number of steps-adequacy was indicated on a 7-point categorical scale with respect to the assertion 'the number of steps is' (1 = far too few, 4 = perfectly all right, 7 = far too many).
[b]The assertion 'I've learned to conduct a plea' was scored on a 7-point categorical scale (1 = very strongly disagree, 7 = very strongly agree).
[c]The relevance of the supportive info was indicated on a 4-point categorical scale (0 = strongly irrelevant, 3 = strongly relevant).
[d]Satisfaction for between-step support was indicated on a 8-item, 4-point categorical scale (0 = very dissatisfied, 3 = very satisfied) and could only be collected for the intermediate and max condition.
$*p < .05$.

conditions were equally efficient. There were no differences for the conditions with respect to transfer task efficiency.

Satisfaction/perceived efficacy

Participants in the min condition significantly differed with those in the other conditions with respect to their opinion of the adequacy of the number of steps (see Table 6). The min condition was most satisfied, $F(2, 32) = 4.74$, $MSE = .98$, $p < .05$, $\eta^2 = .23$. *Post hoc* tests using Bonferroni's correction indicated that both the min and the max condition differed significantly ($p < .05$). There was no significant difference between the int and the max condition nor between the int and the min condition ($p = .06$). In other words, participants in the min condition thought their one-step approach to be slightly more adequate as those in the max condition. Participants in both other conditions were not unsatisfied, but indicated that the number of steps could be slightly decreased. Satisfaction did not significantly correlate with mental effort, motivation, nor time-on-task.

Feelings of efficacy ('I've learned to conduct a plea') did not differ between conditions, which was also the case with respect to their opinions on 'the relevance of the provided supportive information' and their satisfaction as to the between-step support.

Discussion

Limiting the number of steps in learning to solve complex tasks leads to optimal learning-effectiveness, as was predicted in Hypothesis 1. Too many steps made the learning task less coherent; though time on task was increased, no concomitant increase in learning was observed. Although as efficient as the optimal condition with respect to amount learnt per unit time, too few steps led to a lower performance on the learning task. Since there were no differences with respect to prior knowledge and skills, the results can be attributed solely to the variation in the number of steps. None of the other variables that might affect learning (motivation, increased cognitive load, perceived adequacy, time on task, satisfaction, and so forth) proved to be significant and, thus, cannot explain this result.

Contrary to the expectations in Hypothesis 2, performance on the transfer task did not differ between conditions. As the conditions also did not differ with respect to the effort learners expended in the transfer situation, explanations for the equal transfer-performance require more detailed analyses of the logging data. Our data shows that learners did not make use of the non-compulsory learning tasks for practicing what they had been taught. This was contrary to our expectation that they would make use of these high variability practice tasks to practice what the different steps required. As such, the learning situation was limited to a single set of learning task assignments as researchers agree that transfer cannot be expected in such circumstances (e.g. Gagné, Yekovich, & Yekovich, 1993; Van Merriënboer, 1997).

As predicted in Hypothesis 3, learning efficiency was lowest for learners confronted with a high number of steps. Learners in this condition showed a significantly higher time-on-task, but this did not lead to better learning. Contrary to our expectations, no differences in time-on-task were found between learners in the no-steps and intermediate number-of-steps conditions, and this explains why the intermediate number of steps was not the most efficient condition. This lack of superior efficiency for the intermediate condition cannot be attributed to differences in cognitive load/mental effort. There were also no differences in transfer task efficiency. Explanations for this are similar as was the case for the rejection of the transfer hypotheses as was stated in Hypothesis 2.

The practical implications of this study are quite straightforward. There is clear empirical evidence for the value of optimization the number of steps in learning tasks. Too many steps leads to lower performance and, thus, does not justify the extra, but apparently unnecessary, costs of developing such instructional materials. In other words, development costs can be reduced since less instructional material is needed. Too few steps leads to lower performance. Although the costs here are also reduced, the reduction in learning precludes this option. The impact on learning is, in our opinion, not exclusive to law, but is applicable to all domains/situations where whole learning tasks are not too simple, and where it is possible to identify functionally relevant steps in order to apply step optimization. A consideration specific to our situation is that the instructional material presupposed that students had roughly the same prior knowledge and skills. A current trend in education is to develop personalized, student-centred instruction that takes differences in prior knowledge and skills into account. This clearly results in more instructional material, as optimization on the number of steps needs to be done for student-groups differing in prior knowledge and skills. Data on mental effort and time on task seems a workable solution for tailoring the instructional material 'on the fly'.

The results point out several directions for future research. First, since optimization of the number of steps leads to better performance, other instructional measures such as the availability of different types of within-step support, could also be of importance. We are currently conducting experiments on the relation between the number of steps and kinds of within-step support to gain more insight into this. A second research direction is the effects of the number of steps on transfer task performance and efficiency. In our study, the high variability practice learning tasks were not compulsory and the participants could not be expected to invest more than 60 study hours for the practical (the course containing the practical was a 100 hour module). High variability practice has shown its value in various studies (Cormier & Hagman, 1987; Paas & van Merriënboer, 1994; Quilicy & Mayer, 1996; Shapiro & Schmidt, 1982; Singley & Anderson, 1989). A shorter time-on-task for whole learning tasks could bypass this practical obstacle, though the risk of choosing whole tasks that are too simple would exist. Finally, further research is needed to investigate whether the sort of learner control exercised influences the support given.

Acknowledgements

The authors would like to thank the law students from various Dutch universities for their participation in this study and Nick Broers from the University of Maastricht for his statistical advice.

References

Achtenhagen, F. (2001). Criteria for the development of complex teaching-learning environments. *Instructional Science*, *29*, 361–380.

Admiraal, W., Wubbels, T., & Pilot, A. (1999). College teaching in legal education: Teaching method, students' time-on-task, and achievement. *Research in Higher Education*, *40*, 687–704.

Bonner, S. (1994). A model of the effects of audit task complexity. *Accounting, Organizations and Society*, *19*(3), 213–234.

Brown, J. S., Collins, A., & Duguid, S. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*, 32–42.

Clark, R. E. (1999). Yin and Yang cognitive motivational processes in operating multimedia learning environments. In J. J. G. van Merriënboer (Ed.), *Cognition and multimedia design*. Heerlen, The Netherlands: Open University of The Netherlands.

Cormier, S. M. & Hagman, J. D. (Eds.). (1987). Transfer of learning: Contemporary research and applications. San Diego, CA: Academic Press.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*, 671–684.

Edens, F. M., Rink, F., & Smilde, M. J. (2000). De studentenrechtbank: een evaluatieonderzoek naar beoordelingslijsten voor presentatievaardigheden [The student law court: an evaluation study on assessmentlists for presentation skills]. *Tijdschrift voor Onderwijsresearch*, *24*, 265–274.

Gagné, E., Yekovich, C., & Yekovich, F. (1993). *The cognitive psychology of school learning*. New York: Harper Collins.

Hannafin, M., Land, S., & Oliver, K. (1999). Open learning environments: Foundations, methods, and models. In C. M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory* (Vol. 2, pp. 115–140). Mahwah, NJ: Erlbaum.

Jonassen, D. (1999). Designing constructivist learning environments. In C. M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory*, (Vol. 2, pp. 215–239). Mahwah, NJ: Erlbaum.

Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. *Human Factors*, *40*, 1–17.

Karweit, N. (1984). Time-on-task reconsidered: Synthesis of research on time and learning. *Educational leadership: Journal of the Association for Supervision and Curriculum Development*, *41*, 32–35.

Mayer, R. E. (1999). Designing instruction for constructivist learning. In C. M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory* (Vol. 2, pp. 141–159). Mahwah, NJ: Erlbaum.

Mayer, R. E., & Moreno, R. (2002). Aids to computer-based multimedia learning. *Learning and Instruction*, *12*, 107–119.

Maynard, D. C., & Hakel, M. D. (1997). Effects of objective and subjective task complexity on performance. *Human Performance*, *10*(4), 303–330.

Nadolski, R. J., Kirschner, P. A., van Merriënboer, J. J. G., & Hummel, H. G. K. (2001). A model for optimizing step size of learning tasks in Competency-based Multimedia Practicals. *Educational Technology Research and Development*, *49*, 87–103.

Nadolski, R. J., Kirschner, P. A., van Merriënboer, J. J. G., & Wöretshofer, J. (in press). Development of an instrument for measuring the complexity of learning tasks. *Educational Research and Evaluation*.

Paas, F. G. W. C., & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive load approach. *Journal of Educational Psychology*, *86*, 122–133.

Paas, F. G. W. C., van Merriënboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills*, *79*, 419–430.

Potter, S. S., Roth, E. M., Woods, D. D., & Elm, W. C. (2000). Bootstrapping multiple converging cognitive task analysis techniques for system design. In J. M. C. Schraagen, S. F. Chipman & V. L. Shalin (Eds.). *Cognitive task analysis* (pp. 317–340). Mahwah, NJ: Erlbaum.

Quilicy, J. L., & Mayer, R. E. (1996). Role of examples in how students learn to categorize statistics word problems. *Journal of Educational Psychology*, *28*, 175–224.

Shapiro, D. C., & Schmidt, R. A. (1982). The schema theory: Recent evidence and developmental implications. In J. A. S. Kelso & J. E. Clark (Eds.), *The development of movement control and coordination* (pp. 113–150). New York: Wiley.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Sternberg, R. J., & Frensch, P. A. (1991). *Complex problem solving: Principles and mechanisms*. Hillsdale, NJ: Erlbaum.

Tuovinen, J. E., & Paas, F. G. W. C. (2004). Exploring multidimensional approaches to the efficiency of instructional conditions. *Instructional Science*, *32*, 133–152.

Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills*. Englewood Cliffs, NJ: Educational Technology Publications.

Van Merriënboer, J. J. G., Schuurman, J. G., de Croock, M. B. M., & Paas, F. G. W. C. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test and training efficiency. *Learning and Instruction*, *12*, 11–37.

Westera, W., & Sloep, P. B. (1998). The Virtual Company: Toward a self-directed, competence-based learning environment in distance education. *Educational Technology*, *38*, 32–37.

Wöretshofer, J., Nadolski, R. J., Starren-Weijenberg, A. M. A. G., Quanjel-Schreurs, R. A. M., Aretz, C. C. W. M., van der Meer, N. H. W., Martyn, G., van den Brink, H. J., Slootmaker, A., & Berkhout, J. (2000). *Preparing a plea* [Pleit voorbereid] (version 1.0) [multimedia CD-ROM]. Heerlen, The Netherlands: CIHO.

Zachary, W. W., Ryder, J. M., & Hicinbothom, J. H. (in press). Cognitive task analysis and modeling of decision making in complex environments. In J. Cannon-Bowers & E. Salas (Eds.), *Decision making under stress: Implications for training and simulation*. Washington, DC: APA.

Received 27 November 2002; revised version received 8 December 2003