Interactive Learning Environments
2004, Vol. 12, No. 3, pp. 251–264

Taylor & Francis
Taylor & Francis Group

Asking Questions During Self-Directed Inductive Learning: Effects on Learning Outcome and Learning Processes*

Pascal Wilhelm[1] and Jos J. Beishuizen[2]
[1]University of Twente, The Netherlands, and [2]Free University Amsterdam, The Netherlands

ABSTRACT

Asking learners standardized questions during performance of a self-directed inductive learning task might be a useful way to complement think aloud protocol data. However, asking questions might also scaffold the learning process and thus influence the exact processes one wants to study. In the study described in this paper two groups of learners performed a computerized self-directed inductive learning task in which they conducted experiments to discover the relations between five independent variables and one dependent variable. In one condition, the learners thought aloud, in the other the learners were asked additional standardized questions pertaining to specific reasoning steps during learning. Measures of learning outcome and learning processes were collected. It appeared that the questions did not influence learning outcome. With respect to learning processes no differences were found, except that learners in the no questioning condition more often repeated experiments. It was concluded that the questions do not seem to threaten the validity of research findings.

INTRODUCTION

To combine the use of the think aloud method and asking standardized questions might be a useful and complimentary way to gather data on learning outcome and learning processes in a learning task. However, asking specific questions might also scaffold learning and thus influence the variables under study. In this paper, we present a study into the effects on learning outcome

*The research described in this paper was supported by a grant from the Dutch National Science Foundation (NWO 575-22.002).
Address correspondence to: Pascal Wilhelm, Faculty of Behavioral Sciences, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands. Tel.: +31-53-4894207. Fax: +31-53-4892849. E-mail: p.wilhelm@utwente.nl

10.1080/10494820512331383449$16.00

and learning processes of using thinking aloud and standardized questioning at the same time. To compare, we included a condition that only used thinking aloud. The effects were studied in the context of self-directed inductive learning.

Imagine a learner who has been given a balance scale similar to the one used in research by Inhelder and Piaget (1958). The balance scale has two arms with several pegs located at equal intervals along each arm and the learner can place weights on any of these pegs. The learner is given the task of discovering what rules determine whether the arm tips to the right, to the left or remains level after a lever that holds the scale motionless has been released. Presented in this way, this task can be regarded a typical self-directed inductive learning task (Wilhelm & Beishuizen, 2003); the learner needs to conduct experiments to discover the rules and is free to conduct as many experiments in whatever order he or she likes.

To establish learning outcome, we can simply interview the learner on what rules were discovered after he or she has finished. To study learning processes, we could ask the learner to think aloud while he or she is performing the task. However, does the extra task of verbalizing thought processes influence the way the task is normally addressed? Ericsson and Simon (1993) posed that the question whether cognitive processes involved in problem solving are affected by thinking aloud depends on the type of instruction that is given. On the hand, if the instruction requires verbalization of information that would not otherwise be attended to, then task performance is affected. For example, Van Someren and Elshout (1985) showed that chess players who were asked to reflect on their moves did better in a subsequent "quiet" game, compared to chess players who were not asked to reflect on their moves. Berry (1983) found a similar transfer effect with Wason's selection task. Learners first performed a concrete version of the task, followed by an abstract version. All learners successfully performed the concrete version of the task. However, only the learners who were asked to verbalize their reasoning during or following performance of the concrete version of the task did better on the subsequent abstract version of the task.

On the other hand, if the thinking-aloud instruction does not call upon reporting information that is otherwise not attended to, then thinking aloud does not affect task performance. Ericsson and Simon (1993) stated that "...information that is heeded during performance of a task, is the information that is reportable; and the information that is reported is information that is heeded (p. 167)." This information is present in short-term memory and when learners think aloud, it is this information that is reported. Several studies reported by Ericsson and Simon (1993) show that thinking aloud does not affect learning processes, but it does affect the speed at which they occur.

Thus, given an appropriate amount of time, thinking aloud while working with the balance beam seems to be a valid research method to gain insight into cognitive processes during task performance. However, despite its apparent usefulness thinking aloud still has its drawbacks. For example, one cannot be certain in advance whether the protocols contain all information one is interested in. For various reasons, protocols may be incomplete. Therefore, some researchers ask their participants questions during task performance so that they can be sure that the information they are interested in is gathered. An additional benefit is that the protocols can be analyzed in a more efficient way because only the responses to the questions need to be attended to. Of course, this approach may also lead to a loss of information, but if the content of the responses to the questions are the focus of the research, this might be taken for granted.

A good example of the questioning method is described in Kuhn, Garcia-Mila, Zohar, and Andersen (1995). They asked their learners (fourth-graders and adults) for intentions, predictions, inferences and justifications for these inferences during the performance of self-directed inductive learning tasks in different domains. As with the use of the think aloud method, the question whether learning is affected by the research method itself can also be asked here. Klahr and Carver (1995) criticized Kuhn et al.'s approach by stating that the set of questions they used (e.g., "What are you going to find out? What do you think the outcome of this experiment will be? What have you found out?"), presents an underlying goal structure for systematic experimentation to the learners and thus scaffolds learning. They stated that an improvement in learning outcome over time Kuhn et al. observed, would not have occurred when the questions were not asked. Because the study included tasks in two different domains, Klahr and Carver (1995) stated that Kuhn et al.'s study should be viewed not only as a study of transfer but also as a study of transfer of training. Thus, asking specific questions may pose a threat to the validity of the research findings.

Kuhn et al. (1995) opposed to the critique of Klahr and Craver by stating that the questions were necessary for collecting the data they were interested in, namely reasoning processes underlying experimentation strategies and inferences. They stated that it is not easy to get people to think and since they were interested in how the exercise of thinking influences learning outcome and learning processes over time, the questions served a critical methodological purpose. Still, Kuhn et al. admitted that studying the influence of the presence versus absence of the questions is on their "to do" list.

For our research (see De Jong et al., in press) we adopted the question methodology of Kuhn et al. (1995). In one study (Veenman, Wilhelm, & Beishuizen, 2004) we presented four inductive learning tasks in two different domains to different age groups and used similar questions as in Kuhn et al.'s study. The research question in this study pertained to the relative influence of intellectual and metacognitive skills on inductive learning from a developmental perspective. If the questions would have no effect on learning outcome and learning processes, then the use of these questions would be methodologically sound, but if learning outcome and learning processes would be affected by the questions, then the research findings could be confounded. Therefore, we decided to test the hypothesis of Klahr and Carver (1995), which states that the questions may scaffold learning and thus lead to better learning outcomes compared to a situation in which the questions would be absent.

Two groups of learners performed a computerized inductive learning task. In both conditions learners had to think aloud, in one condition the learners received additional standardized questions. The questions we asked learners pertained to intentions ("What are you going to find out?"), inferences ("What have you found out?") and arguments for inferences ("How do you think this outcome came about?" "How do you know that?" "Do the experiments you conducted show you that what you are saying is true? Can you show me?"). In line with Ericsson and Simon (1993), our hypothesis is that questions that stimulate learners to attend to information that is otherwise not attended to have the potential to affect learning. Especially questions that foster learners to reflect or elaborate on their thinking might have this potential. In our view, the questions "What are you going to find out?" and "What have you found out?" have the quality to impose a goal structure for systematic experimentation on the learner, because the learner may come to anticipate the questions and will try to answer them. These questions also have the potential to stimulate reflective thinking, since they stimulate the learner to become explicit about intentions and inferences. The questions "How do you think this outcome came about?" "How do you know that?" "Do the experiments you conducted show you that what you are saying is true? Can you show me?" have the potential to evoke reflective thinking, because a learner is asked to elaborate on the outcomes and has to justify them.

Taken together, we expected to see effects on both learning outcomes and learning processes. In the first place, we expected the learners in the questioning condition to show better learning outcomes. For this purpose we used a comprehension score, which was inferred from a structured interview conducted after a learning session. In addition, we calculated the prediction error rate (Kuhn et al., 1995). Prediction error rate is the mean difference between the predicted outcome and the actual outcome of each experiment conducted by a learner. A low prediction error reflects a better understanding of what can be learned in the task than a high prediction error.

Secondly, we expected specific effects on learning process measures. To this end, learning process measures were extracted from computer log files collected during task performance. The measures pertained to time (time on task and time per experiment), the experiments conducted (total number of experiments and total number of unique experiments), and the number of variables changed per experiment (indicative of usage of the Control-of-Variables Strategy, Chen & Klahr, 1999). When the hypothesis of Klahr and Carver (1995) that the questions scaffold learning is true, we expected the students in the questioning condition to experiment more systematically. This means they would vary less variables per experiment and conduct more unique experiments than learners in the no questioning condition. Because questioning inevitably takes time, we expected time spent per experiment and time on task to be longer in the questioning condition.

METHOD

Participants

Thirty-five students from the Faculty of Social Sciences of Leiden University voluntarily participated in the study. Mean age was 21.6 years (range 18–26 years) and the number of males and females was approximately equal. A cognitive ability measure was used to match the learners before they were randomly assigned to the questioning or no questioning condition. Unfortunately, data of four learners was lost. Sixteen learners were assigned to the questioning group, 15 learners to the no questioning group. They received 16 EUR for their participation.

Materials

Learning Tasks

All students performed the “Peter task”, which was implemented in the FILE system (Flexible Inquiry Learning Environment; Hulshof, Wilhelm, Beishuizen, & van Rijn, in press). The interface is shown in Figure 1.

The task was presented to learners in the context of a story. The story is about the problem of arriving at school in time, and features a boy who has to make several choices about how he rides his bicycle to school. These choices will determine the number of minutes he will arrive late for school. Learners are given the task of finding out how different variables (e.g., how breakfast is arranged, the type of shoes that are worn) influence the number of minutes it takes to ride to school.

For the purpose of explaining Figure 1, it is divided into six different regions (regions A to F). The learners select a level for each of the input variables to conduct an experiment and are presented with the outcome on one output variable. The input variables are shown on the left side of the screen (region A). Each variable is shown in its own row, as an array consisting of two or more

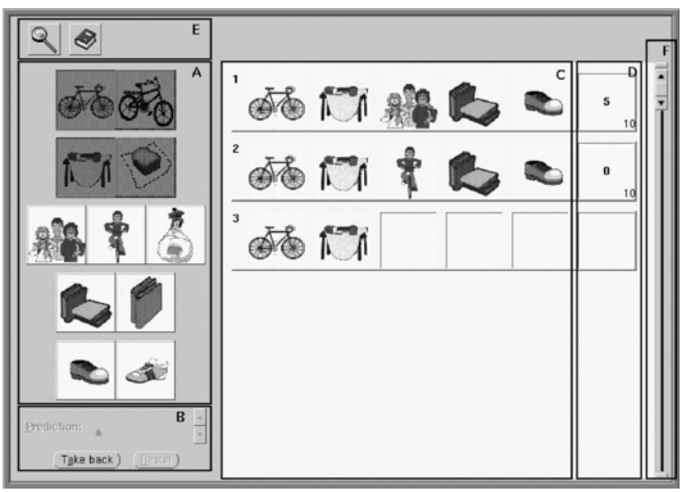

Fig. 1. Interface of Peter task. Section A: independent variables and their levels, Section B: prediction menu, Section C: experiment window, Section D: outcomes (boldface) and predictions, Section E: button used to show learner-selected sets of experiments (left) and instruction button, Section F: scroll buttons.

small pictures, one for each level of the variable. For example, in the first row, a racing bike and a normal bike are presented, which represent the two levels of the first variable "type of bicycle". After selecting a level of a particular variable, a picture referring to the chosen level is added to the experiment window in the middle part of the screen (region C). FILE disables the row, making it impossible to select another picture from that row, and gives visual feedback by graying out the row. When students have selected a level for all five variables, the "Outcome" button is enabled (region B). After pressing this button the outcome of the newly constructed experiment is shown on the right side of the screen (region D). Also, after the "Outcome" button is pressed the rows of variables are enabled again, and students can construct the next experiment. Next to the "Outcome" button, there are two additional interface elements in region B. Firstly, before the outcome is shown, learners have to predict it. The prediction can be edited using the keyboard or the mouse (by pressing the up and down buttons shown at the top right part of region B). Secondly, if a learner selects a level for a variable and later on decides that the selection of that level is incorrect, it can be taken back by pressing the "Take back" button. This button removes the last selection made from the experiment, and re-enables the variable to which it belongs. Once completed, an experiment stays fixed and cannot be changed by the learner. If learners conduct more experiments than the number that can fit on the screen (region C), earlier experiments will scroll off the screen. By using the scroll bar (region E), learners can scroll back and forward through experiments they have carried out. A different way to examine previous experiments, or to compare different sets of experiments, is to select a set of experiments and to display these in a separate window. Selecting an experiment is done by simply clicking on one of the experiments shown on the screen. The background color of this experiment changes to indicate that the experiment has been selected. When a learner clicks on the magnifying glass (region F), a window is shown with the selected experiments. Again, when more experiments have been selected than can be shown at once, a scroll bar can be used to get the other experiments into view.

Cognitive Ability

The following tasks were selected for measuring cognitive ability: (1) Number Series (Elshout, 1976); (2) Abstract Syllogisms (Conclusions, Elshout, 1976); (3) Hidden Figures (Flanagan, 1951) and (4) Spatial Insight (DAT, Evers & Lucassen, 1983). This test battery was constructed to contain subtests that call upon inductive reasoning ability (Elshout, 1976; Wilhelm & Beishuizen, 2003).

Procedure

The cognitive ability test battery was administered in a group session, the learners performed the Peter task in individual sessions. All learners were told that they had to perform a task during which they had to think aloud. A 2-min tape recording of someone thinking aloud was played as instruction. Learners in the questioning condition were also told that they would be asked questions during task performance which they had to answer.

After the instruction the learners read a task instruction from the screen and practiced with the task to get used to the interface. It was made sure that the learners used all functions in the task at least once during practice. No outcomes of experiments were visible during practice. After practice, they could commence experimenting. They were told that they could conduct as many experiments as they wished and that the test session stopped if the learner thought he or she knew about all the effects, or a time period of 45 min was exceeded. At the end of each learning session, the learners were interviewed about the effects they had found.

The questions asked during the learning session in the questioning condition were standardized and concerned the learner's research intentions and inferences. Before each experiment (except for the first one) they were asked: ''What are you going to find out?'' and after each experiment (except for the first one conducted) they were asked: ''What have you found out?'' If the inference question yielded an answer containing only the outcome of the experiment, the experimenter asked: ''How do you think this outcome came about?'' If the inference question yielded an answer containing a statement about the effect of an independent variable (e.g., ''If Peter takes less books he arrives earlier at school''), the experimenter asked: ''How do you know that?'' The purpose of this question was to motivate the learner to relate their inferences to the evidence they generated. If this question yielded theory motivated answers (e.g., ''If you take less books you carry less weight which makes you faster''), the experimenter asked the following question to focus the learner on the evidence he/she had generated: ''Do the experiments you conducted show you that what you are saying is true? Can you show me?''

Data Collection

Two sets of data were analyzed. First, a learning outcome measure (comprehension score) was calculated from the interviews conducted at the end of each learning session. In these interviews, the experimenter asked the learner to describe the effects of each independent variable in detail (e.g., "What difference does type of bicycle make?"). On the basis of the learner's answers a comprehension score was calculated, which served as the learning outcome measure. Interrater reliability of this measure was established in a former study (Wilhelm & Beishuizen, 2003) and yielded a percentage agreement of 81%. In this study, the experimenter scored all interviews.

In Table 1, correct statements about the effects in the learning task are depicted in abstract form. Input variable B (type of bicycle) and A (breakfast) interacted. This interaction was disordinal in nature, meaning that the effect of one input variable was reversed under the influence of another. In the Peter task, riding to school on a race bike goes faster than on an ordinary bike when breakfast is eaten at home. This effect reverses when breakfast is eaten on the road. This effect can be explained by the fact that it is more uncomfortable to eat breakfast on the road on a race bike compared to an ordinary bike. Input variable C (speed) had a main effect, one of its three levels resulted in a different effect than the other two levels, which had no effect. When Peter rides with his friends, he comes 5 min later, riding at his own pace or as fast as he can makes no difference. Input variables D and E were irrelevant. Type of shoes worn and the number of books carried to school had no effect on arrival time.

Table 1. Scoring Scheme Comprehension Score.

	Correct	Without restricting condition	Incorrect
$A1 = A2$ (if $B = 1$)	2	1	0
$A1 > A2$ (if $B = 2$)	2	1	0
$B1 > B2$ (if $A = 1$)	2	1	0
$B1 < B2$ (if $A = 2$)	2		0
$C1 < C2$	2		0
$C1 < C3$	2		0
$C2 = C3$	2		0
$D1 = D2$	2		0
$E1 = E2$	2		0

Note. Letters A–E refer to independent variables, numbers 1, 2 and 3 refer to the levels of the independent variables. Variables A and B interact, variable C has a curvilinear effect, variables D and E are irrelevant. $A1 > A2$ (if $B = 2$) means that level 1 of variable A produces a better outcome than level 2 of variable A if level 2 of variable B is chosen. Students received 1 point if they omitted the "if..." – statement in describing the interaction effects. Maximum score: 18.

A total of nine statements (see Table 1) cover the effects present in the task. Points were awarded to the students when their answers to the questions matched a correct statement. In Table 1, letters A–E refer to the independent variables (e.g., "type of bicycle", "type of shoes"), 1, 2 and 3 refer to their levels (e.g., "race bike", or "sneakers"). The expression $C1 < C2$ means that level 2 of Variable C produces a better higher outcome than level 1 when all other variables are held constant. Two points were awarded for each correct statement about an effect, zero points when the statement was incorrect or absent. If a student did not mention the restrictive condition in describing the effects of the interacting variables (e.g.: "... riding on a race bike goes faster than on an ordinary bike *if breakfast is eaten at home*"), one point was awarded.

Secondly, learning process measures were extracted from computer log files (time on task, time per experiment, total number of experiments, number of unique experiments and mean number of variables changed per experiment). The prediction error was also extracted from the log files, but pertains to learning outcome.

RESULTS

Cognitive Ability

Scores for each subtest were transformed into z-scores because the tests contained different numbers of items. Mean z-score for cognitive ability in the questioning group was $-.24$ (*SD*: 2.95) and in the no questioning group it was $-.72$ (*SD*: 2.98). This difference was not significant, $F(1, 28) = 0.196$, $p = .661$.

Comprehension Score and Learning Process Measures

In Table 2, means and standard deviations for the comprehension score and the learning process measures are depicted. An ANOVA showed no significant differences between the questioning and the no questioning condition with respect to comprehension score, prediction error, time on task, number of unique experiments, and mean number of variables changed per experiment. Time per experiment was significantly longer in the questioning condition, $F(1, 28) = 8.74$, $p < .05$, and learners in the questioning condition conducted less experiments, $F(1, 28) = 6.32$, $p < .05$. All learners finished the task within the time period of 45 min. Although learners in both conditions did not differ

Table 2. Comprehension Score and Learning Process Measures.

	No questioning ($n = 15$)		Questioning ($n = 16$)	
	M	(*SD*)	*M*	(*SD*)
Comprehension score	13.2	(4.1)	11.5	(5.5)
Time on task	1129	(455)	1540	(991)
Time per experiment	54	(25)	93**	(43)
Total experiments	22.9*	(8.9)	16	(6.1)
Unique experiments	17.6	(5.6)	14.3	(5.3)
Usage of CVS	1.9	(0.40)	1.9	(0.33)
Prediction error	1.1	(0.36)	1.2	(0.41)
Experiment ratio	0.80	(0.13)	0.91*	(0.11)

Note. Maximum comprehension score is 18 points. Time on task and Time per experiment in seconds. CVS: Control-of-Variables Strategy. Usage of CVS: mean number of variables changed per experiment. Prediction error: mean difference between prediction and actual outcome of each experiment a learner conducted. Experiment ratio: number of unique experiments divided by total number of experiments.
$^{*}p < .05$, $^{**}p < .01$.

in the number of unique experiments conducted, the ratio between the number of unique experiments and the total number of experiments was smaller in the questioning condition. For the questioning group this ratio was 0.91 (14.3/16, *SD*: 11), for the no questioning condition it was 0.80 (17.6/22.9, *SD*: 13). This difference was significant, $F(1, 28) = 6.27$, $p < .05$.

CONCLUSIONS AND DISCUSSION

In this study, we could not confirm the hypothesis of Klahr and Carver (1995) that the questioning methodology of Kuhn et al. (1995) we adopted scaffolds the learning process. No effect on learning outcome was detected; both the comprehension score and the prediction error rate did not differ between the questioning and the no questioning group. With respect to learning process measures, usage of the Control-of-Variables Strategy (Chen & Klahr, 1999) and number of unique experiments conducted were unaffected by the questions.

Differences between the questioning group and the no questioning group emerged with respect to the time spent per experiment and the total number of experiments conducted. Time on task did not differ significantly between the groups. This disconfirms our hypothesis with respect to time on task. It seems that after having conducted a particular number of (unique) experiments, both groups of learners finish the task. During this period, learners in the questioning condition spent more time per experiment than learners in the no questioning condition. The latter group conducts the same number of unique experiments, but more frequently repeats experiments, which compensates for the time questioning takes in the questioning group. This explains why the ratio between the number of unique experiments and the total number of experiments conducted differed significantly between the conditions. Thus, of the total number of experiments conducted the learners in the questioning condition conducted more unique ones. An explanation for this might be that because the questions make the learners spent more time on each experiment, the experiments are conducted in a more thoughtful way, which prevents learners from conducting the same experiments twice or more as learners in the no questioning condition did more frequently. This may be interpreted as a more systematic approach to conducting experiments evoked by the use of questions, as Klahr and Carver (1995) predicted. On the other hand, usage of CVS (an important indicator of systematic experimentation) and the total number of unique experiments did not differ between the conditions. Therefore, we are reluctant to conclude that questioning clearly scaffolds learning through imposing a systematic approach on conducting experiments.

Why did the questions we used not have the effect predicted by Klahr and Carver (1995)? It might be that the nature of the task played a role here. Self-directed inductive learning is an open-ended learning activity, which calls upon self-regulation on the learner's part. Learners are known to have difficulties with this type of learning (see De Jong & Van Joolingen, 1998). Therefore, it might be that the task was too difficult for the questions to have an effect on learning outcome and learning processes. No floor effects in comprehension score were found, but from experience with the task we know that learners especially have difficulty discovering the disordinal interaction effect in the tasks used (see Wilhelm & Beishuizen, 2003). Another possibility is that the questions we used did not address information that would otherwise not be attended to. Analysis of think-aloud protocols in the no questioning condition might shed light on this issue. The question is then whether the mere instruction to think aloud also produces verbal utterances pertaining to research intentions and inferences. However, in her reply to Klahr and Carver's criticism, Kuhn et al. (1995) stated that the questions served as a means to get people to think. Based on this observation one might doubt whether mere thinking aloud would have produced verbal utterances pertaining to intentions and inferences. Therefore, it is more likely that the questions did have the potential to stimulate learners to focus on information that would otherwise not been attended to, but that they did not have the expected scaffolding effect.

To conclude, we found no clear evidence of an effect of our questioning methodology on learning outcome and learning processes. Therefore, the position of Klahr and Carver (1995) that using the questioning methodology in a study of learning changes that study from a study of learning into a study of training can be questioned. This might suggest that the questioning methodology, at least in self-directed learning tasks, can provide for an valid additional source of information in conjunction with think aloud protocols.

A few limitations of the study should be kept in mind. In the first place, the number of learners in this study was limited. Further research should include more learners to ensure sufficient statistical power. In the second place, one should also consider that other factors, such as age, might moderate the effect of the questions. Moreover, one should be careful to generalize these findings to other types of learning tasks since they may be task-specific.

ACKNOWLEDGEMENTS

We are grateful to Philip Niewold who let us use part of the empirical data he collected for his master thesis (Niewold, 1998) for the purpose of writing this paper.

REFERENCES

Berry, D.C. (1983). Metacognitive experience and transfer of logical reasoning. *Quarterly Journal of Psychology*, *35A*, 39–49.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, *70*(5), 1098–1120.

De Jong, T., Beishuizen, J., van Someren, M., Veenman, M., Hulshof, C., Wilhelm, P., Prins, F., & van Rijn, H. (in press). Determinants of discovery learning. In P. Gärdenfors, U. Riis, & P. Johansson (Eds.), *Cognition, education and communication technology.* Mahwah, NJ: Lawrence Erlbaum Ass.

De Jong, T., & Van Joolingen, W.R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, *68*(2), 179–201.

Elshout, J.J. (1976). *Karakteristieke moeilijkheden in het denken.* Unpublished doctoral dissertation, University of Amsterdam.

Ericsson, K.A., & Simon, H.A. (1993). *Protocol analysis: Verbal reports as data.* Cambridge, MA, MIT Press.

Evers, A., & Lucassen, W. (1983). *Differentiële Aanleg Testserie.* The Netherlands: Swets & Zeitlinger.

Flanagan, J.C. (1951–1956). *Flanagan Aptitude Classification Tests.* Chicago Science Research Associates.

Hulshof, C.D., Wilhelm, P., Beishuizen, J.J., & Van Rijn, H. (in press). FILE: A tool for the study of inquiry learning. *Computers in Human Behavior.*

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking.* London: Routledge & Kegan Paul.

Klahr, D., & Carver, S.M. (1995). Scientific thinking about scientific thinking. In: D. Kuhn, M. Garcia-Mila, A. Zohar, & C. Anderson (Eds.). *Strategies of knowledge acquisition.* Monographs of the Society for Research in Child Development, 60 (4, Serial No. 245).

Kuhn, D., Garcia-Mila, M., Zohar, A., & Anderson, C. (1995). *Strategies of knowledge acquisition.* Monographs of the Society for Research in Child Development, 60 (4, Serial No. 245).

Niewold. (1998). *Inductive reasoning enigmas: The structure of the inductive reasoning process in a computerized environment.* Unpublished master's thesis, University of Leiden, Leiden, The Netherlands.

Van Someren, M.W., & Elshout, J.J. (1985). Het effekt van zelreflectie op leren probleemoplossen. In: J.G.L.C. Lodewijks & P.R.J. Simons (Eds.), *Zelfstandig leren.* Lisse, The Netherlands: Swets & Zeitlinger.

Veenman, M.V.J., Wilhelm, P., & Beishuizen, J.J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction*, *14*, 89–109.

Wilhelm, P., & Beishuizen, J.J. (2003). Content effects in self-directed inductive learning. *Learning and Instruction*, *13*, 381–402.