

Factors Influencing Strategy Use In Inquiry Learning Tasks

Joke Noppers, 0104442

Enschede, 24 October 2011

University of Twente

Author Note

The author wishes to thank Ard Lazonder and Pascal Wilhelm for their help and guidance in writing this paper. Ivo Nouwens is acknowledged for his insightful comments and general help. I also like to thank all the research participants, for making this study possible. Lastly, I would like to thank my parents for their support.

Abstract

This study investigated three factors that may be of influence on the type of investigative strategy learners use in an enquiry learning task. These factors are existing domain knowledge, goal orientation and systematic thinking. Learners with high domain knowledge were expected to employ a theory driven strategy. Learners with less domain knowledge were expected to use a data-driven strategy. Learners with a high ability in systematic thinking were expected to be more consistent in their strategy use than low-ability learners. The adoption of a Scientist or Engineering goal orientation is hypothesized to influence the adoption of a theory driven or data-driven strategy. Participants were 23 volunteers. Their domain knowledge, systematic thinking and goal orientation were measured by pen-and paper tests. Then they engaged in an enquiry learning task. Participants' strategy use in this task was compared with their scores on the other factors. The results support the hypothesis about domain knowledge and strategy use. The two other hypotheses are not supported by the results. The results of this study and their implications are discussed.

Keywords: inquiry learning, Scientific Discovery as Dual Search model

Factors Influencing Strategy Use In Inquiry Learning Tasks

Learning is not something that happens to us. Learning is something we do. Sometimes learning is the rote memorization of facts, sometimes it is seeing connections, and sometimes it is both. Learning can also take the form of finding out or discovering – a mode of learning known as inquiry learning. In inquiry learning, students learn by actively investigating a phenomenon and drawing conclusions about the nature of that phenomenon (De Jong & van Joolingen, 1998, Kuhn, Black, Keselman, & Kaplan, 2000). But how do human beings do this? What kind of strategies do people use to reach their goal of understanding something? To understand something, we often use the methods scientists use in investigating phenomena. That is, we generate hypotheses, design and conduct experiments and draw conclusions based on outcomes.

Klahr and Dunbar (1988) presented a model of scientific reasoning, the Scientific Discovery as Dual Search model (SDDS). This model describes the process of scientific discovery as a search in two problem spaces: an experiment space and a hypothesis space. The hypothesis space consists of all possible hypotheses that can be formed about the nature of a given problem. The experiment space consists of all possible experiments that can be performed to discover the nature of this problem. When faced with a problem, learners can do two things. They can conduct experiments, and let the experimental data guide them towards an understanding of the problem. Then, they are searching the experiment space. They are using a strategy that is driven by experimental data. The other option is a search through the hypothesis space. Before conducting experiments, learners think about the subject and form hypotheses about it. Then, these hypotheses are put to the test. This is a theory-driven strategy. Klahr and Dunbar saw this distinction in their seminal 1988 study. They asked research participants to discover the function of a mystery button on a robot tank. This

robot could be programmed to move around on the floor. Participants could perform experiments by programming the robot and watching how it responded to different commands. The mystery button was named RPT N. All participants assumed that the N on the button stood for the number of repetitions of the program. However, this assumption was incorrect. The letter N stood for the number of commands in the program that was repeated. To discover the function of the RPT N button, participants had to realize that their assumption about the meaning of N was wrong, and they had to discover the actual meaning of N. Klahr and Dunbar made their distinction based on how participants discovered the actual meaning of N. Did they find it out by reasoning about the robot? Or did they need a experiment to realize this? Klahr and Dunbar termed the former group ‘Theorists’ and the latter ‘Experimenters’ since finding the solution through a critical experiment would point to the use of a data-driven strategy. In this study, nearly all (19 out of 20) participants succeeded in finding the solution in the allotted time. The Theorists however, needed less time and less experiments than the Experimenters to solve this problem.

Existing knowledge influences the search through the hypothesis space by providing ideas for hypotheses and leads learners to believe that some hypotheses are more likely to be true than others. Lazonder, Wilhelm, and Hagemans (2008) further investigated the role of existing knowledge in inquiry learning. They predicted that when people are faced with a problem in a familiar domain, they are more likely to adopt a theory-driven approach. When faced with a task they are unfamiliar with, learners are unable to generate hypotheses and start searching the experiment space. They gave their research participants an abstract and a concrete problem to solve. These problems were isomorphic, but had different ‘cover stories’. In the abstract task participants had to deduce the influence of a number of variables (represented by geometrical shapes) on a dependent variable (a numerical score). In the concrete task, these variables were real-world variables like the influence of a healthy diet or

smoking on a dependent variable: an athlete's time on a 10 kilometer race. Because of the abstract nature of the geometrical shape task, the research participants had no prior knowledge about the task. Therefore, the abstract task was an unfamiliar task. In the concrete task, the participants could use their real-world knowledge about athletics and exercise, to form hypotheses. Their results supported the predictions. In the familiar task, subjects used more theory-driven strategies than in the abstract task. This study also yielded some other relevant results. 17 of the 21 participants who used a theory driven- strategy in the familiar (concrete) task, also used this approach in the unfamiliar (abstract) task. These participants seemed to show a preference for a theory-driven approach, regardless of whether they had prior knowledge about the task or not. This indicates that, apart from prior knowledge, there is some other factor that influences strategy use. Another interesting effect this study found was, that some of their research participants used a mix of theory-driven and data-driven strategies. Sometimes these participants searched the hypothesis space and other times they let experiments guide them through the inquiry task. If domain knowledge has an influence on the 'choice' to use a theory-driven strategy or a data-driven strategy, what is it, that accounts for the use of a strategy that switches between theory-driven and data driven approaches? And besides prior knowledge, what else may be of influence on the adoption of different strategies? This study seeks to answer these questions. To be more specific, this study tries to take a closer look at two personal traits, finding out if these traits have an influence, and if so, how this influence works. These traits are systematic thinking and goal orientation. In the next paragraph, these traits, will be further explained. It is also explained how these traits might be related to strategy use.

The first factor worth mentioning is goal orientation. When people engage in a task, they may have particular goals in mind. Their main goal may be to just finish the task, but more subtle mindsets can exist too. Do these have an influence on strategy choice?

Schauble, Klopfer and Raghavan (1991) found that people can engage in an inquiry task with different goals. One goal is to produce desirable outcomes, to be able to manipulate variables in such a way that a certain outcome can be produced. Another goal can be to reach an understanding of the problem. Schauble et al. (1991) compared these two goals with the goals of the fields of engineering and science. Engineers perform experiments and calculations in order to reach a certain goal, like building a bridge. They may investigate properties of different metals, not to know more about metal, but to know which, and how much metal is needed for their construction to work. Scientists on the other hand, investigate metal for its own sake, basically because they are motivated to know more about metal. So in general, engineers strive to produce a certain desirable outcome, scientists strive towards understanding of a system.

In their study they presented research participants with different problems, one more consistent with the engineering model (finding out under which conditions boats travel the fastest) and the other more consistent with the science model (lowering an object into fluid to observe the effects of buoyancy on objects with different mass and volume). They also varied the context of each task. One version had a task description which was worded consistent with an 'Engineering' frame. The other version was worded consistent with a 'Scientists' frame. They found that the science task and the science framing were associated with broader exploration.

The question is, whether these goals influence the choice for a data-driven or theory-driven strategy. And if they do, what kind of influence might this be? It is possible to speculate about the influence of goal orientation on strategy use. It may be that, with an engineering goal in mind, people will systematically search the experiment space until the desired outcome is reached. So engineering goals might correspond with a data-driven strategy choice. But, it can be argued that an engineering goal orientation might also be

associated with a preference for theory-driven strategies. The work of Schauble et al. (1991) showed that people with an Engineering goal orientation tended to focus on the variables they believed to cause the best outcome. Dunbar (1993) found similar results. He re-created a real scientific discovery, the discovery of the concept of inhibition in molecular biology, by Jacob and Monod (1961). Inhibition means that on the molecular level of a single cell, the production of certain chemicals is triggered by turning an inhibitory mechanism off instead of turning an activating mechanism on. Dunbar used a simplified computer simulation to model the experiments performed by Jacob and Monod on *E. Coli* bacteria. He found that very few research participants were able to find this concept. They had a great difficulty finding ‘the right answer’ (inhibition) because it was hard to discard ‘the wrong answer’ (activation) which is a much more intuitively appealing concept. Dunbar also found that the few participants who did manage to find ‘the right answer’ tended to state different goals than the other participants. Six out of seven successful participants stated that their goal was to find out what accounted for the surprising experimental results. Ten out of thirteen unsuccessful participants stated that the goal was to manipulate conditions to find the situation in which no chemicals were produced. Working towards a ‘optimal’ (no chemicals) situation can be viewed as an Engineering goal orientation. Trying to find out what accounted for the unexpected results is a much more explorative view and can be seen as a Scientists’ goal orientation. These two studies show that participants with an Engineering goal orientation tended to conduct their experiments in a certain direction, the direction were they believed the optimal situation was to be found. So, this points to a Theorist strategy, where experiments are performed to check an existing theoretical notion.

The same questions can be asked about a Scientists’ goal orientation. To have a Scientist goal orientation means trying to seek understanding. Perhaps, trying to find this understanding, one might explore all possibilities. This would point to an Experimenters

approach. But what does ‘exploring all possibilities’ mean? Would someone with a Scientists’ goal orientation explore this possibilities through the experimental way? Or might he conduct this search in the hypothesis space? Someone who is trying to gain an understanding about a certain subject, is likely to reason about this subject. He will create a mental model of the thing he is trying to understand. Hypotheses are part of such a model, so this would point to a preference for a theory-driven approach. So what can be said about goal orientation and strategy use? Are the Engineers, who like to take the shortest way to the best results, the ones most likely to adopt a theory-driven approach? Or will they start performing experiments, without hypothesizing about the underlying system? Are the Scientists, in their search for understanding, more likely to perform all kinds of experiments? Or do they tend to search the hypothesis space? Each of the statements above has some validity. Saying which one is true is therefore, difficult. The hypothesis of this study is that goal orientation has influence, but that the direction of this influence is unknown.

The second factor that may have an influence on strategy use in a discovery learning task is systematic thinking. In the following paragraph we will discuss how systematic thinking might influence strategy use. Systematic thinking may be a precondition for successfully tackling problems. When thinking in a systematical way, you structure your thinking. This may enable you to work towards a clear mental representation of the problem. Having such a concept in mind and systematically weighing the different options available to solve this problem, may divide a big problem into a series of less difficult little steps and therefore, bring a solution within reach. It is hypothesized here, that systematical thinking has an influence on strategy use. The systematical handling of a problem may involve a systematical and consistent use of available strategies. A less systematical approach may bring about a more inconsistent use of strategies, using different strategies on a trial-and-error

basis, without much conscious consideration. This study hypothesizes that it is a lack of systematical thinking that accounts for the use of mixed strategies.

For systematic thinking, ‘metacognition’, the ability to reflect on one’s own cognitive processes (Flavell,1976), is a necessary precondition. Tackling a problem in the way described above means that you think about you own mental representation of the problem, and about different steps that you may take to solve this problem. This is thinking about your own thinking and therefore, it can be shared under the concept of metacognition. So, it can be argued here that metacognition is a precondition for systematic thinking, which is, in turn a precondition for consistent strategy use. Research shows that there is indeed a connection between metacognition and strategy use in general. Borkowski and Carr (1987) stated that “metacognition is largely responsible for the decision to be strategic” (p. 63) and that “metacognition subsequently directs conscious use of strategies, including the monitoring of strategies”(p. 63). People who have a high ability for systematic thinking tend to be consistent in their strategy use. They are less likely to exhibit ‘messy’ problem solving strategies, since they will notice inconsistencies in their performance. Individuals with a lower ability in systematic thinking seem to lack the skill to monitor their strategy use. A study by Belmont and Butterfield (1971) found that retarded individuals use poor and inconsistent learning strategies. In a review of strategy use in reading, Garner (1990) found evidence of poor cognitive monitoring predicting failure to use reading strategies. These findings imply that people with a low ability in systematic thinking tend to be inconsistent in their strategy use. Therefore, the influence of systematic thinking may exert itself on the consistent use of strategies. Low ability in systematic thinking may be a factor that accounts for the use of mixed strategies in inquiry learning tasks.

In summary: This study investigates data-driven and theory-driven strategies in enquiry learning tasks. It tries to look at which factors may influence the use of such

strategies. Prior results have pointed out that existing domain knowledge has an influence on the adoption of a data-driven or theory-driven strategy. But this is not the only factor that exerts an influence on strategy use. Therefore, this study seeks to identify other factors that may influence strategy use. More specific, it aims to identify factors that influence the adoption of different strategies, and factors that influence how consistent these strategies are applied. It is hypothesized that goal orientation may have an influence on the adoption of a theory-driven strategy or a data-driven strategy. This is a two-sided hypothesis. Scientists as well as Engineers may be more inclined to adopt theory-driven strategies. This study tries to establish whether there is an effect and which direction this possible effect has. It is also hypothesized that systematic thinking has an influence on the consistency of strategy use. People with a high ability for systematic thinking will be more consistent in their strategy use than people with a low ability for systematic thinking. This study tries to measure strategy use in an enquiry learning task and looks whether connections with measures of the other factors arise.

Method

Participants

A total of 23 research participants volunteered for this study. 8 of the 22 participants were female. All participants were Dutch. Participants' ages ranged from 17 to 63 years old. The mean age of the participants was 34,9, SD 15,1. Participants' educational levels ranged from lower professional education to university.

Instruments

Participants' demographic characteristics were determined by a background questionnaire, inquiring about their age, sex, nationality and educational level.

The choice for an inquiry task that fitted the goals of the study required some effort. To uncover differences in strategy use, it is important to have variability in existing domain knowledge. It is hard to find any correlation between strategy use and other factors, if there is no variability in existing domain knowledge. Asking people about a domain most people have a lot of knowledge about, results in high scores. Asking about a domain very few people have knowledge about will result in equally low scores. With this in mind, a suitable 'cover story' with an intermediate difficulty for the inquiry learning task was chosen. The cover story was about Peter, an athlete training for a 10 kilometer run. Most people will know at least something about running and exercise in general. For instance: A lot of people will know that smoking has a detrimental effect on an athlete's performance. However, there are also some elements of running knowledge mostly known to running experts alone. For instance, there is the influence of Body Mass Index (BMI). Most people associate a high body mass index with obesity and will think that a high body mass index will hamper running performance. However, since muscles are heavier than fat, most well trained athletes also have a high body mass index, and muscular heavy athletes will, in general, perform better than less muscular, lighter competitors. So, since the domain of athletics contains some 'basic facts' that are known to the general public and also some more specialist knowledge, this domain was hypothesized to yield varying existing domain knowledge scores.

The Peter running task is a Flexible Inquiry Learning Environment, (FILE) developed by Hulshof, Wilhelm, Beishuizen and Van Rijn, (2005). It allows the user to generate and test hypotheses by conducting experiments. These experiments are performed by changing the levels of different independent variables and observing their influence on one dependent

variable. There are different ‘cover stories’ for FILE tasks available, the Peter Running story being one of those cover stories. For the Peter Running cover story, the goal was to find out the influence of food, training type, training frequency and Body Mass Index on an athlete’s performance. Participants had to find out the magnitude of these effects and whether they interacted. The FILE task is ran on a computer, displaying the visual interface on the computer screen. In a FILE task, independent variables typically have 2 or 3 different values, which are represented by small pictures. The user can select these different values by clicking on those pictures; When all independent variables are assigned a value, the user can make his own prediction of the outcome and ‘run’ an experiment. After running the experiment, the resulting value of the dependent variable is presented. The user can check to see if the outcome matches his prediction. All the previous experiments in a session are stored. The user can scroll down to look at previously conducted experiments. Also, the user can use a special magnifying glass window to compare a special selection of previous experiments. Different ‘cover stories’ are available. As said before, this study used the cover story of Peter, the athlete and his 10 kilometer run. The inquiry task was framed in a question to the research participant: ‘Can you provide Peter with two advices about preparing for the running match?’ The 5 variables that could be manipulated were: training frequency (one time a week, or 3 times a week), type of training (interval training or endurance training), food intake (carbohydrate-rich food, regular food or junk food), muscle type (type 1 or type 2), and body mass index (a body mass index of 20, 25, or 30). Apart from the main effect for each factor, there was an interaction effect between training frequency and body mass index. Higher Body Mass Index amplified the positive effect of higher training frequency. Peters running time ranged from a best possible performance of 37 minutes to a worst possible performance of 63 minutes.

There were several options available for measuring systematic thinking. One such option was the use of a self-report questionnaire. However, there is some evidence that self-report measures of study behavior do not predict actual behavior (Prins, Busato, Elshout, & Hamaker, 1998; Verhey, Veenman, & Prins, 1999). Lompscher (1995) found that this is also the case for self-reported metacognitive skillfulness. Veenman and Elshout (1999) proposed that metacognition should be assessed by evaluating students' actual study behavior by observing and evaluating problem solving behavior and strategy use on a learning task. So measuring problem solving behavior while participants are working on the Peter Running task could be another option. Still, some caution is in order here because metacognition and strategy use cannot be measured on the same task. This is because this study wants to measure a causal influence of the extent of systematical thinking and metacognition on strategy use. It would be unwise to measure systematic thinking by evaluating strategy use on the same task. In that case one is looking for connections between the very same thing, which is not going to be very informative. So to give participants a chance to show their proficiency in systematic thinking, a different task is needed. A candidate for a test of metacognitive skills is the Raven matrices or the Raven Advanced Progressive Matrices (APM). Planning, adopting a strategy, and monitoring the solution process leads to more effective performance in solving test items. Sternberg (1985) argued that "the extent that existing IQ tests do in fact measure intelligence and predict consequential real-world performance, it is in large part because they implicitly measure metacomponential functioning" (p. 299). Thus the APM as an intelligence test can be considered to measure the control processes mentioned by Embretson (1995). Gray, Chabris and Braver (2003) found that performance on the Raven test is positively correlated with activity in the left lateral prefrontal cortex. The prefrontal cortex is associated with fluid cognition (Braver et al., 1997; MacDonald, Cohen, Stegner, & Carter, 2000; Rypma, Prabhakaran, Desmond,

Glover, & Gabrieli, 1999) According to Blair (2006), fluid cognition involves inhibition of irrelevant information or other information that is likely to interfere with the maintenance of information, and the planning and execution of sequential steps or actions. These actions are metacognitive control processes. As measures of general intelligence such as the APM have strong relations with measures of fluid cognition (Embretson 1995; Engle, Tuholski, Laughlin, & Conway, 1999; Gustafsson 1984, 1988; Kyllonen & Christal, 1990), the APM might be a good candidate for measuring systematic thinking and predicting consistency in problem solving. Embretson found that general metacognitive control processes accounted for 71% of the performance on the Abstract Reasoning Test, a test that is based on Carpenter's theory of solving the Raven's progressive Matrix test. (Carpenter, Just & Shell, 1990). For practical reasons, a short form of the Raven Advanced Progressive Matrices, consisting of only 12 items (Arthur & Day, 1994) was chosen to assess Systematic thinking. This short form has psychometric properties similar to those of the long form of the APM. Like the original APM, a single-factor model fits the data of the short form. Reliability is lower than the full APM, with a Cronbach's alpha of .67 (Arthur, Tubre, Paul & Sanchez-Ku, 1999). Arthur and Day (1994) found an alpha of .66. These alpha's may be lower than in the original APM, because the item difficulty in the 12-item APM is increasing faster than in the 36-item APM. This makes the test items in the short form more heterogeneous. Test-retest reliability is fairly good: an alpha of .76 was found (Arthur et al., 1999). Convergent validity was established by comparing the short-form APM with the Wesman Personnel Classification Test (Wesman, 1965). A correlation of .57 was found between the PCT and APM (Arthur et al., 1999).

A pretest was used to assess participants' existing domain knowledge. This pretest consisted of 10 items. These items addressed the influence of the factors training frequency, training type and body mass index on athletic performance. The questions were worded in a

what-if format. They described a hypothetical initial situation, (“Suppose you have trained three times a week for a 10 kilometre race”) and a alternative situation, (“if you had trained two times a week”). The items also described three post-change answering options about the direction of the effects (“you probably would complete the race... A, faster, B, the same, C, slower”). Participants then chose the most likely post-change situation as an answer. The content validity of the what-if items was ensured by a representative coverage of the content of the running task, and an accurate representation of variables and relations in the task. Seven items were a comparison between two values of the same factor. Three items addressed the effect of a common factor in general. ‘If you train more, you probably will run’.. A, faster, B, the same, C, slower”. Every combination of values was addressed once. The items dealt with the direction of the effect, magnitude of the effects were to be discovered in the inquiry learning task. All items used the same factors, relations, contexts, and values as in the Peter Running Task. The internal consistency of the items was more or less satisfactory (14 items; $\alpha = .58$).

In his 2005 study about the discovery of the inhibition mechanism, Dunbar found strong relations between stated learning goals and actual behavior in his discovery learning task. Since the goals stated in this study can be seen as Engineering or Scientist goals, it may be safe to assume that self-report measures are an accurate predictor for Scientist/Engineering goal-oriented behavior. Therefore, a tree-item questionnaire was chosen to measure Goal orientation. The items asked which goals participants worked toward while solving the Peter Running task. Translated in English, items were worded like: What is it that you were the most interested in? The items had two answering options, one pointing toward an Engineering goal orientation “I was interested in identifying the factors that give the best result” and one pointing towards a Scientist goal orientation “I was interesting in how all factors influenced each other”.

Design and data analysis

Existing domain knowledge was assessed by the answers on the items of the pre-test. These items were multiple-choice items with three answering options. The answers participants gave were matched against the underlying model of the Peter running Task. One point was given for each correct response. To ensure if these items were understood correctly a pilot test was performed with one individual. This pilot test revealed no problems or ambiguities in the pre-test.

Performance on the short form of the APM was scored by counting the number of right answers. The maximum score was 12. Goal orientation was measured by comparing the number of 'Scientist' answers to the number of 'Engineering' answers.

As mentioned earlier in this study, there are two different ways to look at strategy use. One can look at which strategy is adopted, or how consistently this strategy is used. This study tries to measure both concepts. Following the example of Lazonder et al. (2008), the adoption of different strategies was assessed by analyzing the scoring sheets looking at the specificity of the hypotheses the subjects stated. While participants had at least some hypotheses that guided them through the Peter Running Task, some participants used more elaborate and detailed hypotheses than others. An utterance like 'I am going to test whether junk food will increase my time by two minutes' points to a fully developed hypothesis about the influence of junk food. The only thing the participant has to do is to put it to the test. Therefore, it can be argued that highly specific utterances indicate a tendency towards a theory driven approach. An utterance like "I am going to see what eating junk food does" indicates that there is no such hypothesis. The participant is letting the data speak ("seeing what junk food does"). So, non-specific utterances indicate a tendency toward a data-driven approach. To judge these utterances in a more or less systematic way, points were awarded to each utterance. One point was awarded for stating the presence of an effect ("I think

training frequency has an influence”). Two points were awarded for stating the direction of an effect (“I think the more training, the better”). Three points were awarded for stating the magnitude of an effect. To check the quality of the ratings, two raters coded the hypotheses of eight randomly selected scoring sheets from the Peter running task. Inter-rater agreement was satisfactory (Cohen’s *kappa* 0,67). Because the differences were only gradual, somewhere the distinction had to be drawn. The average score was chosen as the cut-off point. Above-average scores were coded as theory driven strategies, below-average scores were coded as data-driven strategies. The average score was 0.83.

Hypothesis specificity was also used for measuring consistency of strategy use. The utterances on the scoring sheets were sorted on their temporal order of appearance, and then divided by four. This creates four different ‘quartiles’ for each research participant. This makes it possible to compare different quartiles from the same participant. A high mean specificity score on one quartile and a low mean specificity score on another quartile is seen as an indication for inconsistent strategy use. This is because high mean specificity scores point to a theory-driven strategy, while low specificity scores point to a data-driven strategy. A participant receiving high and low specificity scores seems to be using a mix of both strategies. Quartile mean scores in the last quartile were not used in assessing consistent strategy use. Many participants may have checked their conclusions in the last experiments they performed. Sudden changes in hypothesis specificity would indicate that these participants changed their strategy, while they were only double-checking their results. Since inconsistent strategy use is associated with both high and low specificity scores, great variance in these scores can be seen as a indicator for inconsistent strategy use. So, a participant’s total variance of mean specificity scores is taken as the measure for inconsistent strategy use.

Procedure

Participants who volunteered for this study completed their session individually and in their own homes. Participants received the same instructions and followed the same procedure.

Participants first had to complete a short background questionnaire to capture data on demographic characteristics and then they had to complete the pre-test to measure existing domain knowledge about exercise and sports. Then they completed a paper-and pencil version of the short form of the Raven Advanced Progressive Matrices.

After that they were seated behind a laptop computer, on which they had to complete the inquiring learning task. The experimenter first introduced the participants to the FILE inquiry learning environment, by showing another, simpler FILE task, (which had a cover story about the price of a skiing holiday). The experimenter explained the experimental procedures. Then she performed an 'experiment' by randomly clicking on some variable values, to show how the interface works. After this, she started the Peter Running Task. Participants were introduced to the story of Peter, the athlete, who wants to run the 10 kilometer as fast as possible. She asked participants to provide Peter with two advices. Then, participants started experimenting. They could see on an index card the visual representation of each variable. The participants received an answering sheet for their final answer and a draft sheet to take notes. The participants had 40 minutes to finish the task. Participants were asked by the experimenter about the hypotheses they were testing. Every time the participants ran an experiment, or scrolled down to look at previously conducted experiments, or selected a group of previous experiments in the magnifying glass screen, the experimenters asked them two non-directive questions: What are you going to investigate? And: what do you think will be the outcome? The answers were written down on a scoring sheet. Lazonder et al. (2008) found this scoring method to be reliable. In their study they

scored five respondents' answers by two raters, finding an inter-rater agreement of 90%. After that participants were presented with the goal orientation question. They could mark their answer in one of the two answer boxes on the sheet of paper the question was printed on. Then they were told that the experiment was over and were thanked for their participation.

Results

First, APM scores were used to classify research participants as being either high or low on systematic thinking. Twelve participants (52 %) scored 9 points or lower on the APM. Therefore, it was decided to make 9 the cut-off point. The high-scoring group consisted of 11 participants, who scored at least 10 points or higher. The next step was to examine whether systematic thinking was related to consistent strategy use. Participants who scored high on the APM were hypothesized to show more consistency in their strategy use than participants with a low APM score. A one sided *t*-test, found no significant results, $t(21) = 0.43$, $p .34$, indicating that hypothesis specificity of participants in the high-scoring and low-scoring group varied to a comparable extent. Descriptive statistics of the high-scoring group and low-scoring group are shown in Table 1.

The next step was to analyze whether goal orientation influenced the use of a particular strategy. A preliminary analysis was performed to ensure that goal orientation was independent of systematic thinking. As this proved to be the case,

$\chi^2 (1, N = 23) = 2.25$, $p = .19$, goal orientation was analyzed without taking participants' proficiency in systematic thinking into account. Participants were then classified as having either a Scientists ($n = 16$) or Engineers ($n = 7$) goal orientation. Mean hypothesis specificity scores were used to classify the strategies participants used. These strategies were either data-driven (mean hypothesis specificity lower than 0.83, $n = 12$), or theory-driven (mean

hypothesis specificity higher than 0.83, $n = 11$). The cross tabulation in Table 2 shows the distribution of Scientists and Engineers to the two types of strategies. A chi-square test indicated no relation between goal orientation and strategy choice, $\chi^2(1, N = 23) = .10, p = .75$. Descriptive statistics of the data-driven group and the theory-driven group are shown in Table 1.

The final analysis investigated the influence of prior knowledge on strategy use. A one sided t -test with strategy choice as independent variable and participants' pretest scores as dependent variable yielded a significant effect for strategy choice, $t(21) = 2.67, p = .014$, with Theorists receiving higher scores on prior knowledge than Experimenters. Descriptive statistics of the high-knowledge group and the low-knowledge group are also presented in Table 1. Two additional ANOVA's were performed to check for possible differences in prior knowledge between other subsamples. The first compared prior knowledge in the high and low APM groups. The differences between these groups did not reach significance, $F(1, 21) = 1.24, p = .28$. The second analysis compared participants with a Scientist or Engineering goal orientation, and also did not reach significance, $F(1, 21) = .01, p = .92$.

Table 1
Summary of scores for systematic thinking and hypothesis variation, and prior knowledge

	N	M	SD
Systematic thinking¹			
High-scoring-group	11	11.00	.89
Low-scoring-group	12	7.00	2.04
Total	23	8.91	2.57
Hypothesis variation			
High-scoring group	11	.86	.47
Low-scoring group	12	.96	.70
Total	23	.91	.59
Prior Knowledge			
Theory-driven group	11	8,55	1,44
Data-driven group	12	6,25	2,49
Total	23	7,35	2,33

¹ As indicated by participants' APM scores.

Table 2
Cross tabulation comparing the Theorist/Experimenter/group with the Scientist/engineering groups

Strategy choice	Goal orientation		
	Scientist	Engineer	Total
Theory-driven-group	8	3	11
Data-driven-group	8	4	12
Total	16	7	23

Discussion

The aim of his study was to shed some light on different factors that influence strategy use. What influences the choice for one strategy over another? What influences the consistent use of strategies? This study proposed some possible candidates for these factors and hypothesized about the influence these factors may have on strategy use. According to these hypotheses, high existing domain knowledge is associated with a choice for theory-driven strategies. Systematic thinking was hypothesized to influence consistent strategy use. A high ability for systematic thinking was associated with a consistent use of strategies. A Scientist or Engineering goal orientation, was hypothesized to have an influence on the adoption of a data-driven or theory-driven strategy. The hypotheses put forward in this study however, do not specify in which direction goal orientation affected strategy choice.

The study finds an effect for prior domain knowledge and strategy use. Subjects who have more existing domain knowledge tend to use more theory-driven strategies than participants with less existing domain knowledge. However, there are other possible explanations for the effect found in this study. The Theorists turn out to have more domain knowledge. It could be the case that more domain knowledge turns people into Theorists (as hypothesized by Lazonder et al. (2008) or it could be the case that some underlying factor accounts for both the tendency to use theory-driven approaches and the tendency to acquire more domain knowledge.

No significant effects were found for the other hypotheses. Participants who scored high on systematic thinking did not show more consistency in their strategy use. Goal orientation turned out to be fairly independent from strategy choice. There are reasons to think that some real effects might have remained hidden, due to specific features and shortcomings of this study.

First, the sample used is relatively small. In order to reach statistical significance, any effect needs to be relatively large. This may conceal some existing effects. Then, there are the questions concerning validity. Are the used measures valid indicators for the underlying constructs they are supposed to represent? This question holds for all measures, but some measures are of special concern here.

One of these is the construct of consistent strategy use. Lazonder et al. (2008) used hypothesis specificity to identify the space participants were searching. Highly specific hypotheses indicate elaborate moves through the hypothesis space (theory-driven). The measure for (in)consistent strategy use is not based on hypothesis specificity alone, but also on the variance of this specificity. This means that an extra assumption is added. Not only is high specificity indicative of theory-driven searches, but high and low together indicates switching between theory-driven and data-driven searches. It was not checked whether his assumption is valid. Hypothesis specificity could fluctuate for a variety of other reasons. For instance, participants who have searched the hypothesis space extensively, may show a sudden drop in hypothesis specificity. Not because they have changed their strategy, but because they are double-checking their earlier made theoretical moves. Some participants may sometimes feel the need to ‘explain’ their trial and error searches because they think that is the socially acceptable answer. Finding results when searching the experiment space, some participants may form their own theory on how it all fits together, and utter elaborate hypotheses. It is unknown whether hypotheses variation is indicative of actual strategy changes.

There are also questions whether these findings generalize to other settings, and other tasks. The Peter Running Task focuses on inquiry learning. Inquiry learning is only one of the reasoning processes we use in daily life. We probably seldom use it in isolation in the

way the research participants did when working on the Peter Running Task. To what extent the findings of this study generalize to real-life learning and reasoning is unknown.

A last potential imperfection concerns the omission of learning outcomes in the study. Do factors that were hypothesized to influence strategy use, also have an influence on learning outcomes? Perhaps in an indirect way, through strategy use? Klahr and Dunbar (1988) found that the Theorists, needed less time and less experiments than the Experimenters to solve the robot tank problem. It would have been informative to look at such possible influences. So, why are learning outcomes missing in this study?

When the study was designed, it was the initial plan to include learning outcomes in the model. Learning outcomes were to be measured as the number of experiments it took participants to arrive at the solution of the Peter Running Task. The less experiments participants needed to solve the Peter Running Task, the better. But later on, some methodological questions arose. The number of experiments may not be the best way to measure how effective an inquiry task is solved. First, this way of measuring puts Experimenters at a disadvantage. They search the experiment space and by doing so, conduct a lot of experiments. A Theorist's search may take as long as an Experimenter's search, but since it leads through the hypothesis space, less experiments are recorded. This means that the Theorists will score better. Hypotheses about the influence of strategy choice on learning outcomes may become self-fulfilling. Secondly, having done a certain number of experiments does not say anything about arriving at the right conclusion or not. This measure of learning outcomes was not very useful, yielding no reliable information about learning outcomes. Therefore learning outcomes were not used for data analysis.

How are these results relevant for research in the field of inquiry learning? The study does confirm Klahr and Dunbar's hypotheses (1988) and the findings of Lazonder et al. (2008). For the rest of the factors, it finds no statistical significance. However, it has only a

limited sample of participants and limited validity. This means that these results do not rule out systematic thinking and goal orientation as interesting research avenues. It may be useful to look at strategy use at a more detailed level than this study did, with the use of more appropriate models than the hypothesis variance-based model this study used. Lazonder et al. also took this approach in their 2008 study. Maybe a possible effect for systematic thinking on strategy use could then be determined. It may also be useful to investigate the other hypotheses formulated in this study, with a larger sample to uncover any true effects that may have remained hidden. It is also useful to continue the search for other factors that may influence strategy choice, for instance personality traits, social contexts, learning contexts, primes and task aspects. The hypotheses about learning outcomes that were never put to the test, may also be incorporated in these studies.

What are the practical implications of the results of this study? This study hopes to offer some insight in the choice and use of different research strategies. It tries to offer educators a hint in how they might steer learner's strategy use in a certain direction. It also tries to offer educators some understanding why different strategies are used, and when they are used. This study confirms that more domain knowledge tends to be associated with more theory-driven strategies and offers a educators a hint on how to influence strategy choice. Educators could supply learners with the knowledge they need for a search through the hypothesis space. Sometimes the aim of a learning task is a systematic search trough the experiment space. Searches trough an experiment space could be useful to help students learn to deduce information from empirical data. With the results of this study in mind, educators could design learning tasks in such a way that students have little or no prior knowledge and therefore turn to searching the experiment space. Educators also may use the result of this study to know what kind of searching behavior to expect when offering learners a certain task.

References

- Angeli, E., Wagner, J., Lawrick, E., Moore, K., Anderson, M., Soderlund, L., & Brizee, A. (2010, May 5). *General format*. Retrieved from <http://owl.english.purdue.edu/owl/resource/560/01/>
- Arthur, W., & Day, D. V. (1994). Development of a short form for the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement*, 54(2), 394-403. Retrieved from <http://epm.sagepub.com/content/54/2/394.short>
- Arthur, W., Jr., Tubre, T.C., Paul, D.S., & Sanchez-Ku, M.L. (1999). College-sample psychometric and normative data on a short form of the. Raven Advanced Progressive Matrices Test. *Journal of Psychoeducational Assessment* 17 (4), 354-361. doi: 10.1177/073428299901700405
- Belmont, J. M., & Butterfield, E. C. (1971). Learning strategies as determinants of memory deficiencies. *Cognitive Psychology*, 2(41), 1-420. Retrieved from <http://www.sciencedirect.com/science/article/pii/0010028571900247>
- Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences*, 29(2), 109-125. doi: 10.1017/S0140525X06009034
- Borkowski, J. G., & Carr, M. (1987). "Spontaneous" strategy use: Perspectives from metacognitive theory. *Intelligence*, 11(1), 61-75. doi: 10.1016/0160-2896(87)90027-4.
- Braver, T.S., Cohen, J.D., Nystrom, L.E., Jonides, J., Smith, E.E., & Noll, D.C. (1997). A Parametric Study of Prefrontal Cortex Involvement in Human Working Memory. *NeuroImage*, 5, 49-62. doi: 10.1006/nimg.1996.0247

- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97(3), 404-431. American Psychological Association. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/238199>
- De Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68, 179 - 202. doi: 10.3102/00346543068002179
- Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science*, 17(3), 397-434. doi:10.1207/s15516709cog1703_3
- Embretson, S. E. (1995). The role of working memory capacity and general control processes in intelligence* 1. *Intelligence*, 20(2), 169–189. Retrieved from <http://www.sciencedirect.com/science/article/pii/0160289695900314>
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128(3), 309-331. American Psychological Association. Retrieved from [http://chomsky.hum.uct.ac.za/Refs/Engle_et_al_\(1999\)_Working_memory,_short-term_memory,_and_general_fluid_intelligence.pdf](http://chomsky.hum.uct.ac.za/Refs/Engle_et_al_(1999)_Working_memory,_short-term_memory,_and_general_fluid_intelligence.pdf)
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. *The nature of intelligence* 12, 231-236.
- Garner, R. (1990). When children and adults do not use learning strategies. *Review of Educational Research*, 60, 517-529. Retrieved from <http://www.jstor.org/stable/1170504>

- Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, 6(3), 316-322. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12592404>
- Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, 8, 179-203. Retrieved from <http://www.sciencedirect.com/science/article/pii/0160289684900084>
- Gustafsson, J.-E. (1988). Hierarchical models of individual differences in cognitive abilities. *Advances in the psychology of human intelligence Vol 4* (Vol. 4, pp. 35-71). Hillsdale, NJ, England: Erlbaum.
- Hulshof, C. D., Wilhelm, P., Beishuizen, J. J., & Rijn, H. V. (2004). FILE: a tool for the study of inquiry learning. *Computers in Human Behavior*, 21(6), 945-956. doi: 10.1016/j.chb.2004.03.014
- Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 3(3), 318-356. Elsevier. doi:10.1016/S0022-2836(61)80072-7
- Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12(1), 1-48. doi:10.1016/0364-0213(88)90007-9
- Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The Development of Cognitive Skills To Support Inquiry Learning. *Cognition and Instruction*, 18(4), 495-523. doi: 10.1207/S1532690XCI1804_3
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence*, 14(4), 389-433. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0160289605800121>

- Lazonder, A., Wilhelm, P., & Hagemans, M. (2008). The influence of domain knowledge on strategy use during simulation-based inquiry learning. *Learning and Instruction*, 18(6), 580-592. doi:10.1016/j.learninstruc.2007.12.001
- Lompscher, J. (1995). Learning strategies in 4th, 6th, and 8th grade students. In C. Aarnoutse, F. De Jong, H. Lodewijks, R.J. Simons, and D. v. d Aalsvoort (Eds), *Abstracts of the 6th European Conference for research on Learning and instruction*, p. 527. Nijmegen: University of Nijmegen.
- MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288(5472), 1835-1838. AAAS. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10846167>
- Prins, F. J., Busato, V.V., Elshout, J. J., & Hamaker, C. (1998). A new contribution to the validation of the (meta) cognitive part of the Inventory Learning Styles (ILS). *Pedagogische Studien*, 75, 73-93.
- Rypma, B., Prabhakaran, V., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Load-dependent roles of frontal brain regions in the maintenance of working memory. *NeuroImage*, 9(2), 216-226. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9927550>
- Schauble, L., Klopfer, L. E., & Raghavan, K. (1991). Students' transition from an engineering model to a science model of experimentation. *Journal of Research in Science Teaching*, 28(9), 859-882. doi:10.1002/tea.3660280910
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence. The behavioral and brain sciences* (Vol. 7, pp. 269-287). Cambridge University Press.

Veenman, M., & Elshout, J. J. (1999). Changes in the relation between cognitive and metacognitive skills during the acquisition of expertise. *European Journal of Psychology of Education*, XIV, 509-523.

Verhey, J., Veenman, M.V.J., & Prins, F. J. (1999). Identifying technical students at risk: self reported learning styles vs. behavioural measures. In R. Poell (Ed.), *Proceedings ORD 1999*, p.74-75. Nijmegen: KUN.

Wesman, A.G. (1965). *Wesman Personnel Classification Test, Revised Manual*. New York: The Psychological Corporation.