2014

Figuring out how a device works: What you ask is what you get?

UNIVERSITY OF TWENTE

Department of Behavioral Science, Learning Sciences

Examination Committee

Dr. Ard W. Lazonder

MSc. Judith ter Vrugte

Bachelor thesis

Brian Hendrik Mueller

Universiteit Twente

1/23/2014

Table of Contents

Samenvatting

In deze Bachelor thesis word er verder onderzoek gedaan over de consequenties van de Goal specificity effect (GSE). De GSE houd in dat al de manier waarop een doel geïntroduceerd ook de manier beïnvloed waarop de respondent met het doel omgaat. Er wordt ook na gekeken of er met de doelstelling de strategie beïnvloed kan worden die het kind gebruikt om zijn doel te bereiken.

De proefpersonen waren verdeeld in 2 condities. In het experiment worden kinderen met een slot machine bekend gemaakt, bestaand uit 4 knoppen en 4 wielen. De ene conditie kreeg een non-specific doel (What is the function of each button), de tweede conditie een specifiek doel (How to activate each reel). Aan de hand van hun toetsen en hun antwoorden word een analyse gedaan welke dan de een indicatie geeft over of de hypothesen goed waren of niet.

Hypothese 1: Kinderen met een non-specific doel behalen een diepere begrip en meer causaal kennis van het in het onderzoek gebruikte system vergleken met kinderen in de andere conditie.

Hypothese 2: De specificiteit van het doel heeft invloed op welke strategie de kinderen gaan gebruiken.

Het experiment was uitgevoerd aan de OGS Paderborn (Offene Gesamtschule Paderborn) met 25 kinderen tussen 6,5 en 10,5 jaar. Nadat de kinderen aan de researcher zijden dat ze nu bekend zijn met de machine, kregen ze hun doel waar ze door experimenteren achter moesten komen. Tijdens het experiment worden hun verzoeken gedocumenteerd om aan de hand van hum knoppen te zien of er een verschil in strategie is.

Er kon in het experiment gedocumenteerd worden dat de kinderen met een specifiek doel meer vragen goed hebben beantwoord dan de kinderen met een niet-specifiek doal. Er kon helaas niet achter gekomen worden of er een verschil was in strategie keuze was.

Summary

This Bachelor thesis seeks to shed further light on the consequences and implications of the Goal specificity effect (GSE). The GSE describes how already the way a goal is presented to the respondent influences how it will be dealt with. It shall also be investigated if and how the goal influences strategy choice with which the respondents try to achieve their goal.

The respondents were randomly assigned into two conditions. In the experiment the children were introduced to a slot machine, made up out of 4 buttons and 4 wheels. The one condition got a non-specific-goal (What is the function of each button) whereas the other group got a specific goal (How to activate each reel). Regarding the buttons the children used to figure out how the machine works, an analysis was performed to investigate if the hypotheses were correct or not.

Hypothesis 1: Respondents introduced to a non-specific goal develop a more profound causal understanding of the experimental environment than children in the other condition.

Hypothesis 2: The specificity of the goal influences strategy choice.

The experiment was performed at the OGS (Offene Gesamtschule Paderborn) with 25 children between 6,5 and 10,5. After a brief experimentation phase, the children reported when they felt confident using the regular slot machine. Once this was achieved the wicked slot machine was introduced, where buttons were not assigned to the wheels as they were before. Now the respondents were either set up with a specific or non-specific goal, depending on the condition. During the experiment the researcher took notes about which buttons the children activated, which later were used to analyze if there was a difference in strategy between both conditions.

This experiment could document that children with a specific goal answered more questions correctly than children with a non-specific goal. The experiment failed though to identify if there was a difference in strategy choice.

Introduction

Scientific thinking is becoming increasingly important and has gotten more recognition throughout the last couple of years. Scientific thinking is defined as "the application of the methods or principles of scientific inquiry to reasoning or problem-solving situations, and involves the skills implicated in generating, testing and revising theories, and in case of fully developed skills, to reflect on the process of knowledge acquisition and change" (Zimmerman, 2007, p. 173). There is a growing body of research conducted on what is called "psychology of science" (Zimmerman, 2000). Broadly defined, psychology of science is the study of scientific thought or behavior and this topic plays a major role in the fields of education and developmental psychology.

One of the fundamental principles in science is the linkage between cause and effect. Finding such causal relationships is not easy though and involves several steps and subgoals. Much like Sir Isaac Newton's famous apple incident which supposedly led him to discover the laws of gravity, in first instance, through observation a theorie is formed. Next, hypotheses are generated which are then tested by conducting experiments or making observations. There are strategies and tactics on how to achieve a desired outcome or produce a certain insight. Some subgoals might be more valuable than others at a certain time and others might render different subgoals obsolete. Experimentation requires a complex coordination of many subgoals (Dunbar & Klahr, 1989). Schauble and Glaser (1990) noted that elementary children for instance only have a fragment of the understanding regarding experimentation compared to individuals with more experience.

Dewey (1913) was one of the first to differentiate two ways of scientific activity. He suggested that there are two basic scientific purposes of any activity: producing a *desired outcome* and producing fundamental *knowledge.* Since producing a desired outcome has wider applicability to everyday purposes, it is hypothesized that producing a desired outcome has developed earlier than the more analytical form of thinking.

Consistent with Dewey's thoughts on scientific activity, the same differentiation is made between two modes of inquiry when a learner is faced with a certain problem. The first involves learning through manipulation and experimentation until the desired goal is achieved, the second mode

includes learning by formulating theories, testing hypotheses and evaluating the evidence. Schauble referred to these two methods of investigation as *engineering model* and *science model* of experimentation (Schauble, Klopfer, & Raghavan, 1991). The engineering model is more goal oriented, designed to produce a certain goal state whereas the scientific model is more appropriate when trying to gain fundamental understanding of a certain matter.

For example, a student might be faced with a certain environment s/he needs to gain knowledge about (for instance s/he needs to know how to operate a machine). There are now two basic approaches on how to solve the problem. S/he might do so by experimenting and simply pressing buttons, integrating only the positive results in his approach and try to get closer to the desired goal state. All information that does not serve the purpose of getting closer to that goal can be denied in this approach. Eventually, the goal state will be achieved. The other option is to start off with a similar explorative approach but then incorporate the different aspects found into a holistic theory that incorporates all aspects and functions of the machine. It appears that although the first method might deliver faster results simply by its random nature (hitting the right button by accident), the second approach generates more knowledge and deeper understanding about the machine and would allow for the following incident to be solved quickly (understanding why exactly this button was the right one). Here, the motivation and intention of the student is essential, whether the goal is to develop insight or simply the desired goal state. Of course this depends on the context and wording of the specific problem.

Various research has indicated that the goal introduced to learners already has an impact on how they will perform. For instance, if accepted, difficult goals lead to higher performance then easier goals (Locke & Latham, 1990; Locke, Shaw, Saari, & Latham, 1981). Furthermore, Locke and Latham (2006) found that focusing on a specific performance goal in a new, complex task, can create a *tunnel vision,* in which the learner focuses more on reaching the goal rather than acquiring the necessary knowledge and skills to reach it. These two examples, even though only providing a small glimpse, illustrate how a presented goal can have an influence on how the respondent will perform.

To illustrate this further, imagine a learner who tries to learn to play the guitar. Setting yourself the goal to learn to play the guitar is a rather non-specific goal. It gives hardly any instruction on where to start, nor when the goal is achieved. On the other hand there is the specific goal, such as

learning to play a certain song. This goal gives the learner clear instructions such as where to start (at the beginning of the song), when to end (at the end of the song) and when the goal is acquired (when I can play the song). Comparing these two goals highlights the advantages of a specific goal, but it also shows the disadvantages. It might be that a certain type of learner only learns to play the song step by step, reaching the desired goal, but failing in achieving a deeper understanding of how to play the guitar and ignoring aspects such as what a scale is or how a chord is built up.

Introducing a learner with a specific goal, prompts the use of means-ends analysis (Newell & Simon, 1972). This is a problem solving strategy in which the actual state is brought through manipulating the variables into the desired state; a strategy comparable to what was earlier described as engineering model of experimentation. Sweller (1988, 1994) argued that this strategy, although efficient for problem solving, is not suited for learning about the problem. In this problem solving strategy, the goal is achieved once the desired state is reached, which not necessarily requires the student to understand the environment s/he is in. In addition, Sweller argues, it requires considerable amount of cognitive effort to perform. The student at least needs to know about the current state, the desired state and the operators which can be manipulated to reach the desired state.

In addition, previous research has also shown that the learner's intentions of respondents don't always match the goal they were presented with. For instance, Schauble (1990) has shown that children who were asked to identify which car features affect the car's speed, were preoccupied by constructing as fast cars as possible (Schauble, 1990). This finding illustrates how the more scientific nature of the identification of car parts as rather value-free undertaking, becomes a task in which the goal is to build an as fast car as possible. The goal has turned into producing something rather than understanding the underlying causal mechanisms.

However, various aspects of how exactly the presented goal influences strategy choice and which differences might appear in acquired knowledge, still remain unclear. Respondents with a non-specific goal, using the science model, are expected to develop a more profound understanding of the causal relationships between variables and outcomes compared to the group introduced to a specific goal (Schauble, Klopfer, & Raghavan, 1991).

As mentioned earlier, Locke and Latham (2006) found that certain goals can influence task performance, by for instance creating a tunnel vision in which the learner focuses more on reaching the goal than acquiring knowledge. This illustrates that the goal can have an impact on what aspects of a given task the learner focuses on and therefore gains understanding about. The main focus of the present study will therefore be to further investigate if the specificity of the goal results in a difference of gained knowledge between the two conditions.

Children introduced to a specific goal are expected to give more correct responses to the first set of questions in contrast to the children who are introduced to a non-specific goal, who are expected to give more correct answers on the second question. The presented goal is expected to influence the knowledge the children achieve, therefore they should be able to answer the sets of questions the best that resemble their initial goal. Although there has been previous research regarding this topic, the comparing nature of this study and the possibility to observe the influence of the goal's specificity directly, is new and therefore will generate further insight.

Additionally, the question remains if and how the presented goal influences strategy choice. If a certain type of goal provokes the use of a specific strategy to meet this inquiry, presenting different groups with several different types of goals should shed light on the question whether this is the case or not. A difference in strategy choice would give additional information about how a possible difference in gained knowledge might have occurred.

Method

Participants

Participants were 24 children (10 boys and 14 girls) from a primary school near Paderborn, a medium sized city in Germany. This specific primary school offers daycare in vacations, and this research was performed with pupils being in this daycare during the fall break of 2013. Participants were selected by the researcher according to the criteria that the sample would give a good representation of the school's population in ethnical/cultural diversity and age distribution. Specific criteria was a balanced number of males and females in both conditions and an even distribution of age between roughly 7 and 10 years old representing the age distribution naturally present at this specific primary school. Mean age for the sample was 7 years and 7 months, ranging from 6 years and 9 months to 10 years and 2 months. Participants were selected in alternating order into one of the two conditions. One condition mirrored the engineering model ($n = 13$), the other the science model ($n = 11$).

Materials

Children in both conditions worked with the slot machine displayed in Figure 1. This machine appears to be a regular slot machine with four buttons, four reels, and a lever. Usually, to make a machine easily understandable for its users, designers make use of the natural mapping of mental models concept. This refers to the notion that users automatically make inferences between the manner in which a system is designed (i.e.., the location of buttons seems to be "logical" and makes sense to the user). However, in comparison to a regular slot machine, the current study employed a 'wicked' slot machine. In the wicked slot machine however, the function of the buttons was arranged in a counterintuitive way. The blue button was the 'main switch': it needed to be turned on for any of the four reels to spin. With the blue button switched on, the status of the red and green buttons (either on or off, so four unique possibilities) determined which of the four reels would spin. The yellow button had no effect.

Figure 1: The Wicked Slot Machine

Children in both conditions were assigned to figure out how the wicked slot machine worked. Their interactions with the device were guided by a predefined goal that differed among conditions. Children in the engineering goal condition were asked to find out how to activate each reel; children in the science goal condition had to discover the function of each button.

Two sets of questions were used to assess children's final understanding of the wicked slot machine. The first set contained four items about the spinning of the reels (e.g., "How do I activate the red reel?") that were asked in sequential order. The second set comprised four items that went after the function of the buttons (e.g., "What is the function of the red button?"). All questions were read aloud by the researcher who wrote down children's answers.

During the experiment two laptops were employed: one to run the wicked slot machine with an external mouse attached to make it more easily to use for the children, and one for the researcher to take notes. The trials the children performed were noted on a spreadsheet. The answers to the researcher´s questions were noted on paper.

Procedure

The entire research was conducted in a room made available by the school, giving the opportunity to conduct this research with minimal disturbance for the children. They were individually invited into the room where they were informed that there was no time limit and that if there were any circumstances that made them feel uncomfortable, they could stop.

Since the knowledge about how to operate a computer cannot be taken for granted at such an early age, the first step was to assure that the children were able to operate the wicked slot machine. If not already familiar with computers, the researcher would explain how to operate the mouse and which button was the action button.

After having the setup thoroughly explained the child was introduced to a regular version of the slot machine. It was explained that each reel could be activated by pressing the corresponding button under the reel and then pressing the lever. Children appeared to be widely unfamiliar with slot machines, which was not surprising given the rather mature audience regular slot machines target. To ensure that missing knowledge about the slot machine would not impact the way children would later interact with the wicked slot machine, an extensive practice-phase of around 7 to 15 minutes preceded every child's wicked slot machine experiment. Children were asked to report to the researcher once they felt comfortable with using the slot machine.

Following the practice-phase the actual research-phase was initiated. Children were told that now the slot machine was under a curse and had been bewitched. The researcher needed their help to figure out how it worked now. Depending on their condition, children then were asked either to find out how to activate each reel, or to figure out the function of each button. The monitors were set up in a way that the researcher could see what buttons the child was pushing. Every combination the child tried was noted. To be able to draw conclusions about their strategy choice, the children were asked to indicate what their motivation was and why they were trying a specific combination of buttons. This way, the motivation and strategy would become apparent to the researcher.

No time limit had been set in this stage of the experiment in order to exclude time pressure as an influencing factor. Once feeling confident, the child could determine him/herself when s/he felt that s/he had sufficient knowledge to answer the question s/he was introduced to. After finishing

the research phase, children answered the two knowledge based questions presented earlier: first the one that fitted the condition s/he was in, then the question that corresponded with the goal of the other condition.

Results

As shown in Table 1, the average time per task was approximately 6 to 8 minutes. Especially regarding the relatively high standard deviations, it becomes clear that the relative close averages on time spent do not hold that much argumentative power. In both conditions there were participants who finished the experiment very quickly, on the other side there were also a few who took much longer than average. An independent-samples t-test was conducted to compare the amount of time spent on the task in both conditions. There was no significant difference in duration between condition 1 and condition 2, $t(22) = -0.63$, $p = .268$.

There was a considerable difference in the average number of trials, which was almost two-thirds higher in condition 2. The performed t-test nevertheless revealed that this difference was not statistically significant, $t(22) = -1.47$, $p = .079$.

Table 1

Mean Time and Number of Trials and Correct Answers By Condition.

	Condition 1 (Engineering goal)		Condition 2 (Science goal)	
	M	*SD*	*M*	*SD*
Time on task (sec)	412.2	218.2	484.3	338.4
Number of trials	24.8	16.0	37.3	25.2
Correct answers question 1 (reel)	2.2	1.3	0.9	0.8
Correct answers question 2 (button)	0.9	1.4	1.0	0.9

Table 1 also shows the mean number of correct answers per condition. It appears that there was a difference in the number of correct answers given on question 1 between the two conditions. This difference was significant, $t(22) = 2.42$, $p = .012$, which means that children in condition 1 gave significantly more correct answers on the question how to activate each real than children in condition 2. The slight differences between the correct answers on question 2, identifying the

correct function of each button, were not significant, t (22) $=$ 0.25, p $=$ .400. Both groups performed equally well on the second question.

However, the average scores do not reveal what exactly the children in both conditions did and did not know. It might be for instance that the average number of correct answers was the same in both conditions but one of the groups scored better on a different aspect, resulting in a different main emphasis of correct answers.
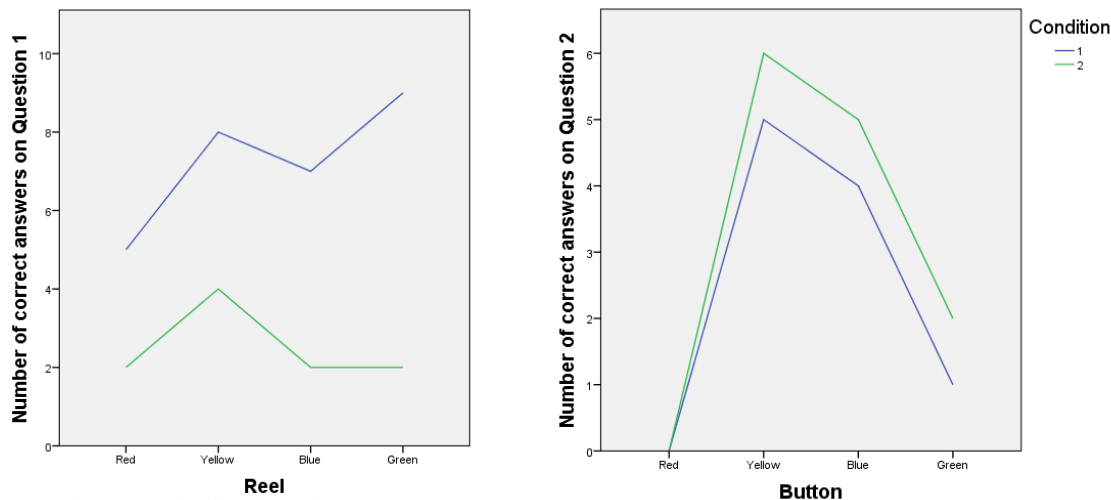

Figure 2. Distribution of correct answers to both questions

Figure 2 shows that there was a similar distribution of the correct answers in both conditions. Even though in condition 1 there were more correct answers on the question how to activate the green reel, the general shape of the graph is very similar to the one of condition 2. The same occurs when comparing the two graphs showing the frequency count of correct answers on question 2.

The children in both conditions had to answer two sets of four questions. Although these questions were different in nature, the knowledge required to answer them correctly was not entirely different. For example, if a child is able to answer correctly how the green reel is activated (by pressing the blue button), s/he might also have some idea how to respond to the question what the function of the blue button is. To analyze if there was such a relationship, Pearson correlations among the two answers were conducted. First, both conditions were analyzed separately to make sure that possible differences between the conditions are addressed.

In condition 1 there was no significant relationship between the scores on the two questions, $r =$ .28, $n = 13$, $p = .35$. In condition 2 the correlation was even weaker than in condition 1, with $r =$ .08, $n = 11$, $p = .82$. Regarding there being no correlation between the scores on the two questions in the two conditions, it appears obvious that also the entire sample correlation would not be significant. This proved to be the case, $r = .19$, $n = 24$, $p = .36$. Together these findings show that there is no meaningful correlation either positive or negative between the answers on the two questions.

Regarding the strategy choice of the two groups, we looked at how many of the trials in each condition were made up of none, single or multiple buttons. The children in both samples were given no instructions about how many buttons to put together, so the choice was entirely up to them and shows how they try to figure out how the machine works. In both conditions the respondents showed a comparable distribution of combinations in their trials (see Table 2).

Table 2

Composition of the Trials per Condition.

	Proportion of activated buttons (%)				
	None	Single	Double	Triple	Quadruple
Condition 1 (Engineering goal)	.01	.25	.37	.19	.17
Including Main switch (blue)	n.a.	.09	.22	.16	.17
Condition 2 (Science goal)	.007	.23	.40	.20	.15
Including Main switch (blue)	n.a.	.08	.22	.16	.15

The two samples show a comparable make-up with no significant difference in how many buttons were used together by the children in each condition. The same applies for the combinations including blue, where also the two samples show no significant difference.

Discussion

The main purpose of this paper was to shed further light on the topic of how the specificity of a presented goal influences knowledge acquisition and strategy choice made by child learners. Based on the literature of Schauble et al. (1991) two experimental conditions were designed that provoked either the engineering model or the science model of experimentation. Children in the first condition, confronted with a specific goal, were expected to give significantly more correct answers on the question how to activate the individual reels than children from the second condition, introduced to a non-specific goal. Children from this condition were expected to develop more causal knowledge and a deeper understanding of the machine, indicated by giving more correct answers on the second question, what the function of each button is. Further it was expected to observe a difference in the strategy children would chose, explaining the difference in generated knowledge.

Expectancies went towards a difference in gained knowledge between the two conditions, with both conditions having an advantage while answering the question that resembles their initial goal. This hypothesis was partly supported by the results. Children in condition 1, who had been introduced to a specific goal, answered significantly more questions correct, showing that they had picked up significantly more knowledge being introduced to a specific goal than the second condition working with a non-specific goal. However, this difference revealed itself only in the answers to the first question (How to activate each reel); the second question was answered with no significant difference in number of correct answers. The distribution of correct answers regarding the different reels in the experiment was relatively alike, suggesting similar "knowledge clusters".

The difference in knowledge gained could only partly be supported in this experiment. In line with our expectations the specific goal led to better performance on the first question, how to activate each reel, possibly by having provoked means-end analysis as suggested by Newell and Simon (1972). In contrast to our expectations, the distribution of correct answers was alike in both conditions. Additionally there was no difference in the distribution of combinations between the two conditions, meaning that no significant difference could be found. Schauble, Klopfer and Raghavan (1991) discovered that the science model was associated with broader

exploration and greater attention to causal relationships among variables, in this experimental setup these findings were not reflected in the scores since both conditions scored comparable on the second question, what the function of each button is.

A difference in knowledge implies a difference in strategy while dealing with the wicked slot machine. Since children were randomly selected into either condition, this would be the only way to explain this difference, providing that circumstances were the same for both groups. One of the goals was to make observations about how the presented goals influenced strategy choices the children made, such as for instance more means-ends analysis, by interviewing them during the experiment. During the interview it became apparent that the questions were distracting the respondents and did not yield any useful information. It is therefore not possible to determine differences or similarities in strategy use. Disregarding the consciousness aspect, the analysis revealed many similarities between the two conditions. No significant difference in time was found, which shows that in both conditions the children had more or less the same sense of when they had tried enough. The same applies to the number of trials and the distribution on combinations which was comparable throughout the entire sample. Both groups gave the main blue switch the same amount of attention measured by the number of hits and the integration into the combinations. This suggests that both groups gave the blue button, acting as main switch, a similar emphasis.

Longitudinal neuroimaging studies have shown that the brain continues to develop even into the 20s (Johnson, Blum, & Giedd, 2009). Therefore the age of the children might have had an influence on the results gathered by the experiment. In order to successfully use a plan in any given situation, the individual has to be able to overcome several smaller tasks that are involved in the planning process. First of all, the child has to have an overview over the task ahead. Has to be able to choose a suitable start, the plan has to be consciously present, progress made has to be noted and results have to be remembered before being able to determine when one is ready. The possibly not yet sufficiently developed verbal skills played a role in this experiment since children in both conditions had trouble reporting their conscious decision regarding their strategy choice. The development of the brain might also have had an influence on the number of correct answers. On the one hand whilst answering the questions, some respondents were not able to

report a combination they had successfully identified during the experiment, apparently simply due to forgetting it. On the other hand, a significant correlation has been found documenting that the children gave more correct answers in both conditions if they were older. Failing to report the right answer might therefore also be partly due to inadequate verbal or memory skills. The data gathered during this experiment, does show a significant influence of age. In both conditions older children did report significantly more correct answers than younger ones which supports this argument.

It would be interesting to see if samples of older children would show similar results. Since the children in our sample had difficulties planning and reporting their strategy choice it would be interesting to see if there is a difference in strategy choice once the respondent group is older and possible better equipped verbally, as well as cognitively, to deal with such a rather complex task. Observing a possibly unconscious decision making process has proven difficult in this experimental setup.

One of the purposes of this paper was to shed further light on the issue if a science model of experimentation generates more causal knowledge than an engineering approach and if more knowledge about bringing a machine in a desired state would be developed in the engineering approach. This experiment documents that children in this age group do generate more knowledge about how to bring a machine into a desired state. This knowledge might be applicable in child education and implemented into schools, helping with the decision at which age which type of goal presentation might be the most suited. If there was a more profound understanding of this topic throughout the entire maturation of a child/young adult, it could be used to streamline the current way of education and make it more efficient. Presenting a specific goal to this age group yielded benefits in numbers of correct answers, presenting a nonspecific one did not. A possible starting-point of further research therefore could be if these results stay constant throughout maturation. Also, how this knowledge can be utilized best in modern child education would require further examination.

Bibliography

Dewey, J. (1913). *Interest and effort in education.* Hillsdale, NJ: Erlbaum.

Dunbar, K., & Klahr, D. (1989). Developmental differences in scientific discovery strategies. In D. Klahr, & K. Kotovsky, *Complex information processing: The impact of Herbert A. Siman* (pp. 109-144). Hillsdale, NJ: Erlbaum.

Johnson, S. B., Blum, R. W., & Giedd, J. N. (2009). Adolescent Maturity and the Brain: The Promise and Pitfalls of Neuroscience Research in Adelescent Health Policy. *J Adoles Health* , 216-221.

Kuensting, J., Wirth, J., & Paas, F. (2010). The goal specificity effect on strategy use and instructional efficiency during computer-based scientific discovery learning. *Computers & Education* , 668-679.

Locke, A., & Latham, P. (1990). *A theory of goal setting and task performance.* Englewood Cliffs, NJ: Prentice-Hall.

Locke, A., Shaw, N., Saari, M., & Latham, P. (1981). Goal setting and task performance: 1969-1980. *Psychological Bulletin, 90* , 125-152.

Locke, E., & Latham, G. (2006). New directions in Goal-setting Theory. *Current Directions in Psychological science* , 265-269.

Newell, A., & Simon, A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice Hall.

Schauble, L. (1990). Belief revivion in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology* , 31-57.

Schauble, L., & Glaser, R. (1990). Scientific thinking in children and adults. In D. Kuhn (Ed.), Developmental perspectives on teaching and learning thinking skills. *Contributions to human Development, 21* , 9-26.

Schauble, L., Klopfer, L., & Raghavan, K. (1991). Students' transition from an engineering model to a science model of experimentation. *Journal of Research in Science Teaching* , 859-882.

Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science, 12* , 257-285.

Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and instruction, 4* , 295-312.

Zimmerman, C. (2000). *The Development of Scientific Reasoning Skills.* Edmonton: University of Alberta.

Zimmerman, C. (2007). The development of scientific skills in elementary and middle school. *Developmental Review* , 172-223.