INTERNATIONAL DOCTORAL SCHOOL REASON

Biological Observation Competency in Preschool – the Relation to Scientific Reasoning and Opportunities for Intervention

Dissertation zum Erwerb des Doctor of Philosophy (Ph.D.) am Munich Center of the Learning Sciences der Ludwig-Maximilians- Universität München

Janina Klemm

München, 2016

Erstgutachterin: Prof. Dr. Birgit J. Neuhaus

Zweitgutachterin: Prof. Dr. Beate Sodian

Datum der Abgabe: 10.11.2016

Datum der mündlichen Prüfung: 21.02.2017

Teile dieser Arbeit wurden auch in Artikeln in wissenschaftlichen Zeitschriften veröffentlicht:

Klemm, J., & Neuhaus, B. J. (2017). The role of involvement and emotional well-being for preschool children's scientific observation competency in biology. International Journal of Science Education, 39(7), 863-876. DOI: 10.1080/09500693.2017.1310408

Klemm, J., Sodian, B., & Neuhaus, B. J. (2017). Scientific Reasoning in Biology – the Impact of Domain-General and Domain-Specific Concepts on Children's Observation Competency. Manuscript submitted for publication.

Klemm, J., Kohlhauf, L., Boone, B., Sodian, B., & Neuhaus, B. J. (2017). Förderung biologischer Beobachtungskompetenz im Kindergarten. Manuscript submitted for publication.

Für meine Oma Irmgard

Acknowledgements

The research presented in this work was supported by the Elite Network of Bavaria [Project number: K-GS-2012-209]. I would like to extend my sincere gratitude to the opportunities (conference attendance, incubator stay, international knowledge exchange) made possible by the ENB.

One aim of this doctoral program was to enable the students to work interdisciplinary. While this phrase often stays a lip service, my two supervisors made it possible to really combine aspects from both biology education and developmental psychology. Their mutual respect for each other and interest in the research holds exemplary function for me for how to approach interdisciplinary projects. My first supervisor, Birgit Neuhaus, was an ideal supervisor in every respect: she fostered my progress, she challenged me, she respected my views, and always gave me valuable feedback. My second supervisor, Beate Sodian, was always there to help me with all the developmental psychology aspects of my work, but also gave me the freedom to integrate them into an applied research question. Another big thank you goes out to my international supervisory, Bill Boone, whose visits in Munich were both very helpful for my work and a lot of fun. When I started my project, I was fortunate that there was already a preschool expert at the chair. Thank you, Luci, for all the discussions and ideas.

One specialty of the REASON Program is that you have a peer group of other students who work on a similar topic, albeit in different fields. This enriched the scientific work, but also the office life. I was especially grateful for sharing my office with the funniest, nicest people in the world, Ansgar, Maryam, and Sandra. Another special thank you goes to Özgün and Andrea, who were my fellow-developmentalists in the program. But everybody made the program special: Andras, Christian, Christian, Daniel, Dr.Horse, Katharina, Maria, Mauro, all the associate members, all the professors, and our lovely coordinators, Kerstin, Markus, and Irina. Additional to the REASON crew, I had a great group of colleagues in the Biologiedidaktik, who helped me when I lacked biology knowledge, made lunch breaks funnier, and taught me about the virtues of having prosecco on any occasion. Also thanks to the youngest member of the group, Noah, who was very patient with his mother‘s doctoral students. Then, there was also the developmental chair group, who did not only give me lots of valuable feedback in the chair sessions but also turned out to be a very funny crew during pub quiz sessions.

This study would not have been possible if I had not had the help of some great students: thank you especially, Amely, Kathy, Maj, and Ana, for going to all the kindergartens with me traveling, always keeping the good mood, even when the mouse nearly escaped. And of course a big thank you to the kindergartens and parents who agreed to let their children participate in this study.

Then there are all these people who are not directly related to my dissertation project, but who were still really important for its success. My ―Psycho-Stammtisch” on Monday nights, who listened to my complaints and praises about my project. My friends I met during lunchbreaks in order to maintain good work-life-balance. David, who never would believe that I ―could not do this”. Andi, who basically kept me together in the last months.

And thank you to my family. You taught me the love for knowledge, and the value of helping each other out, and you proved both points by proofreading this thing here (several times), and giving me pep talks when I needed them.

Abstract

Observation is a central scientific method. It is especially relevant for inquiry in biology, but also an important prerequisite for other methods such as experimenting or comparing. There is evidence that already children in preschool are able to observe, but not yet at the detailed level that is needed for science. It is known that children‘s domain-general scientific reasoning develops at a similar age. However, the research of the two fields has not been linked yet. Instead, research on observation competency has focused on the relation to prior knowledge of the observed. This study investigates the relations of domain-general scientific thinking and domain-specific biology understanding to children‘s observation competency, as well as the relation of additional cognitive and affective factors to observation competency. The second question in this study is whether preschoolers‘ biological observation competency can be fostered with a training program.

75 preschool children (age range: 4;9 to 6;7) were tested for their biological observation competency, their scientific thinking (domain-general scientific reasoning and domain-specific biology understanding), and several cognitive factors (theory of mind, executive functions, language abilities,and prior knowledge). Additionally, their affective state (emotional well-being and involvement) in the observation situation wasassessed. The experimental group (40 children) participated in atraining program that aims at improving children‘s perception of details, hypothesis-lead investigation, and interpretation of observed contexts. There were 12 weekly sessions, each lasting about one hour. After the training, children‘s biological observation competency was assessed in a posttest.

The results show that both domain-general scientific reasoning and domain-specific biology understanding are related to children‘s observation competency, showing effects beyond the shared influence with language abilities. The expected relation of prior knowledge about the observed objects to the quality of the observation can be confirmed. Children‘s emotional well-being and involvement were linked with their performance in the observation situation as well. In a summarizing model, both the scientific thinking components and the affective factors were significant predictors of children‘s observation competency.

The training materials showed good usability and led to a high participation of the children during the training sessions. While the training showed positive effects on observation competency for children who were already better observers in the beginning,this effect could not be found for children who were worse observers in the beginning. For these children, there was a developmental effect across both experimental and control group. For all children, executive functions at the pretest were a predictor of observation competency at the posttest, indiciating the relevance of executive functions for the development of observation competency.

Zusammenfassung

Beobachten ist eine zentrale wissenschaftliche Methode. Sie ist besonders relevant für Untersuchungen in der Biologie, aber auch eine wichtige Voraussetzung für andere wissenschaftliche Methoden wie Experimentieren oder Vergleichen. Es gibt Hinweise darauf, dass Kinder im Vorschulalter bereits beobachten können, aber noch nicht auf der detaillierten Ebene, die für die Wissenschaft benötigt wird. Es ist bekannt, dass sich domänenübergreifendes wissenschaftliches Denken in einem ähnlichen Alter entwickelt. Bis jetzt wurden die beiden Kompetenzen noch nicht gemeinsam untersucht. Stattdessen konzentrierte sich die Forschung zur Beobachtungskompetenz auf den Zusammenhang mit Vorwissen über den Gegenstand der Beobachtung. Die aktuelle Studie untersucht den Zusammenhang der Beobachtungskompetenz mit domänenübergreifendem wissenschaftlichen Denken und domänenspezifischem Biologieverständnis sowie mit weiteren kognitiven Faktoren und affektiven Faktoren. Die zweite Forschungsfrage der Studie ist, ob die biologische Beobachtungskompetenz der Vorschulkinder mit einem Trainingsprogramm gefördert werden kann.

75 Kinder im Vorschulalter (Alter: 4;9 bis 6;7 Jahre) wurden auf ihre biologische Beobachtungskompetenz, ihr wissenschaftliches Denken (domänenübergreifendes wissenschaftliches Denken und domänenspezifisches Biologieverständnis), und mehrere kognitive Faktoren (Theory of Mind, Exekutive Funktionen, sprachliche Fähigkeiten und Vorwissen) hin untersucht. Darüber hinaus wurde ihr affektiver Zustand (emotionales Wohlbefinden und Involviertheit) in der Beobachtungssituation beurteilt. Die Experimentalgruppe (40 Kinder) nahm an einem Trainingsprogramm teil, das darauf abzielte, die Wahrnehmung von Details, das hypothesengeleitetes Untersuchen und die Interpretation beobachteter Zusammenhänge zu verbessern. Es gab 12

wöchentliche Sitzungen, die jeweils etwa eine Stunde dauerten. Nach dem Training wurde die biologische Beobachtungskompetenz der Kinder in einem Posttest erhoben.

Die Ergebnisse zeigten, dass sowohl das domänenübergreifende wissenschaftliche Denken als auch das domänenspezifische Biologieverständnis in Zusammenhang mit der Beobachtungskompetenz standen, und dies über den gemeinsamen Zusammenhang mit sprachlichen Fähigkeiten hinaus. Der erwartete Zusammenhang des Vorwissens über die beobachteten Objekte mit der Beobachtungskompetenz konnte bestätigt werden. Das emotionale Wohlbefinden und die Involviertheit der Kinder hingen ebenfalls mit ihrer Leistung in der Beobachtungssituation zusammen. In einem Modell, das alle Faktoren zusammenfasst, zeigten sich sowohl domänenübergreifendes wissenschaftliches Denken und Biologieverständnis als auch die affektiven Faktoren als signifikante Prädiktoren der Beobachtungskompetenz.

Die Trainingsmaterialien konnten gut angewendet werden und motivierten die Kinder zu einer aktiven Beteiligung während des Trainings. Im Hinblick auf die Auswirkungen des Trainigs auf die Beobachtungskompetenz fand sich ein geteilter Effekt. Die Kinder, die zu Beginn bereits bessere Beobachter waren, profitierten von dem Trainig. Dies war für die schlechteren nicht der Fall. Für diese Kinder zeigte sich ein Entwicklungseffekt sowohl in der Experimental- als auch in der Kontrollgruppe. Für alle Kinder sind die exekutiven Funktionen im Vortest ein Prädiktor für die Beobachtungskompetenz im Posttest, was darauf hinweist, dass exekutive Funktionen ein relevanter Faktor für die Entwicklung der Beobachtungskompetenz sein können.

Table of Contents

1. Introduction

In recent years, science education in preschool has become an important topic. It is integrated into educational guidelines of preschool, and many new programs, books, and materials are being published. However, there are also critical voices: Why should children learn about something as complicated as science? Do we overburden them with too much teaching and too little fun and free time at this early age? And are we trying to turn all children into scientists now? Applying the concept of scientific literacy, I will present arguments for science education in general and then specifically for young children.

1.1. Scientific Literacy

The term of scientific literacy is at the base of the debate about how much science children should be taught during their school career. In the PISA study, a worldwide study by the OECD for comparing students‘ academic performance in mathematics, science, and reading, scientific literacy is defined as follows:

> Scientific literacy includes an understanding of fundamental scientific concepts [..], familiarity with scientific ways of thinking and working, and the ability to apply this knowledge of scientific concepts and processes, particularly to evaluate aspects of science and technology. It also requires the ability to identify questions that can be answered by scientific inquiry and to draw evidence-based conclusions in order to

> understand and help make decisions about the natural world and changes made to it through human activity. (Stanat et al., 2002, p. 6)

This definition is broad enough to cover different facets of science education, like conceptual science knowledge, inquiry skills, and an understanding of the nature of science. At the same time, it formulates concrete aims, like being able to identify relevant questions or help make decisions about the natural world. The question remains why these abilities are necessary to learn.

Kind and Osborne (2016) identified three main arguments why scientific literacy is important. The economic argument bases on the idea that more scientists and engineers are needed for a country in order to further develop economically and compete internationally. In fact, the launch of the Sputnik satellite in 1957 led to reforms in American science education (Rutherford, 1997). However, Kind and Osborne (2016) argue that in fact, only a small percentage of jobs are in science and that there is no general shortage of scientists. The economic argument is, therefore, not a valid argument for teaching all children basic concepts of science.

Kind and Osborne (2016) summarize several lines of argumentation under the term of “the citizenship argument”: Students are supposed to learn science in order to understand political decisions, make informed personal decisions (e.g,. in the area of health), be involved in knowledge production, and to understand the technology they use. These claims are all aimed at educating scientifically informed citizens. It can be discussed how many people truly understand how their computers or smartphones work, or how many are engaged in civic science projects. Nevertheless, it is evident that a basic understanding of how science arrives at conclusions is necessary in order to make adequate political and personal decisions. Recently, there has been a new development of antiscience attitude. Liu (2012) calls that “denialism”: by believing in conspiracy theories or fake experts, cherry-picking data, having unrealistic expectations for research, many people turn their back to scientific knowledge. A lack of understanding how science arrives at conclusions plays an important role here. This leads directly to the third argument, the cultural argument. Science has formed the modern world and human‘s view of it as little else has. While people in the Middle Ages would not believe that the earth rotates around the sun or that organisms so small we cannot see them are the cause of illnesses, we have no problem taking these as facts today – though our everyday observations would tell us differently. Kind and Osborne (2016) argue that, as science is the basis of our culture, everyone should have both basic conceptual knowledge as well as an understanding of how this knowledge was derived.

But even when there are good reasons for science education in general, why should preschoolers learn about science? The three answers discussed are: because they can, because they are interested, and because it will help them later.

Developmental research has found that preschool children possess basic abilities to reason scientifically (Croker & Buchanan, 2011; Glauert, 2010; Koerber, Sodian, Thoermer, & Nett, 2005; Piekny, Grube, & Maehler, 2013a, 2013b; Piekny & Maehler, 2013; van der Graaf, Segers, & Verhoeven, 2015, 2016) and that their conceptual knowledge also develops significantly at this age (Inagaki & Hatano, 2004; Mähler & Ahrens, 2003; Opfer & Gelman, 2010). They also already bring first scientific knowledge from home that can be referred to and built upon in preschool (Cumming, 2003).

The second argument is children‘s apparent interest in the world surrounding them, asking “why-questions” on both the living and non-living environment (Lück, 2015). While young children show high interest in science and technology, it decreases when they get older (Gardner, 1998). At the same time, interest is known to have an effect on their learning in science (Sha, Schunn, Bathgate, & Ben-Eliyahu, 2016). Therefore, it makes sense to make use of children‘s interest in science as early as possible.

The third argument is that learning first scientific concepts in preschool will help them later in school. As preconceptions have a large influence on performance in science activities (Hardy, Jonen, Möller, & Stern, 2006; Möller, Hardy, Jonen, Kleickmann, & Blumberg, 2006), teaching science in preschool may not only prepare all children better for primary school, but also help ―close the gap" for disadvantaged children (Camilli, Vargas, Ryan, & Barnett, 2010; Nores & Barnett, 2013). The effects are not necessarily limited to the science domain – Lück (2015) stresses the impact of science activities on children‘s language abilities, which in turn show positive effects on academic achievements in other domains.

The skepticism toward early science education often also stems from the fact that people associate science learning with instruction-based fact learning. While this has been the first approach when science was supposed to be integrated into early education in the seventies, the modern approaches take children‘s developmental stage into account and choose adequate contents and learning goals (Möller, 2002). Möller and Steffensky (2010) list five main learning goals for science in early education: building connectable conceptual knowledge, a beginning understanding of scientific reasoning and inquiry methods, a beginning understanding of science and the scientific method, an interest in thinking about natural phenomena, and self-efficacy concerning one‘s ability to find out and understand. The first three goals focus on both conceptual knowledge and understanding science in general, and they all emphasize that in early education, first structures are set up that can be built upon in later education. The goal of engaging preschool children in science activities is not to make them expert scientists, but ―that their observations can become increasingly more powerful, productive, and scientific in educational settings" (Eberbach & Crowley, 2009, p. 60). The latter two goals – an interest in thinking about natural phenomena and self-efficacy concerning one‘s ability – are not focusing on cognitive, but affective factors, which shall be awakened and reinforced. Many early education researchers emphasize the relevance of affective goals when bringing science into preschool (Andersson & Gullberg, 2012; Fleer, 2013). Self-efficacy and interest in science have also shown to be crucial for children‘s performance in science classes (Sha et al., 2016). On top of that, having a positive attitude toward science is also an important goal of scientific literacy in general (DeBoer, 2000). Thus, it makes sense to consider both cognitive and affective learning goals in early science education.

1.2. Focus of This Study

As mentioned above, research on children‘s scientific reasoning has been expanded to preschool age and has shown promising results: even though preschoolers‘ abilities are still limited in the sense that their performance depends on prior knowledge and context (Croker & Buchanan, 2011; Koerber et al., 2005), they are in general able to choose adequate experiments (Leuchter, Saalbach, & Hardy, 2014; van der Graaf et al., 2015) and interpret simple data sets (Koerber et al., 2005; Piekny et al., 2013a). However, there are some critical voices concerning this line of work: Lehrer and Schauble (2006) argue that relations to science education and the practical contexts of learning sciences were mostly neglected. For example, most studies did not look at different domains of science, but either focused on physics or just assumed that they are measuring domain-general scientific reasoning. This has led to the situation that many science education programs also focus on physics or chemistry as domains. Looking at preschool science programs, there are many programs focusing on physical and chemical topics like water, density, magnetism, or astronomy (e.g., Hecker & Tansaway, 2008). The materials by Lück (2015) exclusively focus on phenomena of the inanimate nature from the domains of physics and chemistry. While many activities in preschools can be counted as being part of the domain of biology, such as collecting leaves or growing plants, they are not conducted in a scientific way. However, children have many questions about biological contexts – about animals, plants, and humans – that could be picked up and investigated together with them.

Another point Lehrer and Schauble (2006) criticize is that developmental research on scientific reasoning has focused on the control-of-variable-strategy in experiments. While the application of this strategy is a good indicator that study participants have understood the logic of experimentation, science learning in practice has many more facets and requirements and depends on more competencies than applying one strategy. In fact, most investigations in science classes are not experiments, but observations (Eberbach & Crowley, 2009). Observation is a complex scientific method with specific demands on the researcher using it (Oguz & Yurumezoglu, 2007). It is a method particularly relevant for research questions in biology, but is also useful in other domains (Kohlhauf, Rutke, & Neuhaus, 2011). While there is research on how domain-specific factors are related to observation competency (Eberbach & Crowley, 2009; Kohlhauf et al., 2011) and how it can be applied in science lessons (Johnston, 2009; Kelly, 2013; Lommen, 2012; Tokarczyk, 2015), little is known about the development and general cognitive influencing factors. With respect to affective factors, Tomkins and Tunnicliffe (2007) found that children are generally interested in observing biological objects, especially animate ones and that these objects trigger their desire to learn more about them. However, concerning children‘s observation competency, there is a lack of studies investigating the impact of affective factors.

The present study addresses these research gaps: firstly, it will focus on biology as a domain, which is specifically interesting for young children and suited for simple investigations with them. Secondly, it will investigate both children‘s biological observation competency as well as their domain-general scientific reasoning, exploring the parallels and differences of these constructs. Other cognitive domain-general and domain-specific influencing factors will also be considered, as well as the relation with affective factors in the situation. Thirdly, a training program for fostering preschoolers‘ scientific observation in biology has been developed, executed and evaluated in order to close the gap between research and practice.

2. Theoretical Background

This thesis is aimed at understanding more about how children's scientific observation competency in biology develops and how it can be fostered. As this competency consists of the epistemic activities of inquiry in general, parallels to the development of scientific thinking – both domain-specific and domain-general – can be expected. This study further investigates the relation with several cognitive factors, both domain-specific and domain-general, as well as the relation of affective factors to children's performance in a biological observation situation. In the following, the current status of research on the central concepts is summarized. In order to foster this competency adequately in preschool, several factors known from early education research have to be taken into account. Therefore, those will be introduced as well.

2.1. Scientific Thinking[1]

When talking about scientific thinking, authors may have several aspects in mind: they either refer to a style of domain-general reasoning, or include domain-specific knowledge in their definition (Zimmerman, 2000). The first strand is research on scientific reasoning, which is defined as domain-general reasoning and problem-solving strategies. The second one

[1] In the developmental psychology literature, the terms scientific reasoning and scientific thinking are often used interchangeably. In this thesis, however, the term ―scientific thinking" will be used to describe the aggregate of domain-general and domain-specific scientific thinking competencies, while the term ―scientific reasoning" will be used for the domain-general aspects.

focuses on domain-specific scientific concepts, i.e. knowledge about a specific field, which develops through conceptual change. Both aspects are relevant for good scientific work, and both are relevant for science education (Sinatra & Chinn, 2012). Accordingly, both are relevant for the development of children‘s biological observation competency, as children will need both general problem-solving strategies and domain-specific concepts for making good observations. Consequently, I will give a short overview of both aspects.

2.1.1. Children’s Domain-General Scientific Reasoning

The first aspect of children‘s scientific thinking is their domain-general scientific reasoning ability. I will discuss the definition and conceptualization of the term before presenting research on the development of scientific reasoning in young children.

2.1.1.1. Conceptualization of Scientific Reasoning

Zimmerman (2007) defines domain-general scientific reasoning 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 the case of fully developed skills, to reflect on the process of knowledge acquisition and change.(p. 173)

This definition is based on the model of Scientific Discovery as Dual Search (SDDS) by Dunbar and Klahr (1989). This model conceptualizes scientific reasoning as a problem-solving strategy that is relevant for two problem spaces, the hypothesis space and the experiment space. The necessary competencies of a researcher are the knowledge of where and how to look for evidence (in the experimental space), and the ability to look for it in terms of hypotheses (in the hypothesis space). The goal of a scientific process is to formulate a hypothesis or theory that can explain as much as possible evidence from the environment as closely as possible (Klahr, 2002). For this purpose, it is necessary to coordinate the search within each space as well as between the hypothesis space and the experimental space by means of methodological knowledge. The coordination of theory and evidence is seen as crucial in this model and lays the ground for most contemporary approaches in research on scientific reasoning, which put the focus on theory-evidence coordination (Kuhn, 2011).

Additionally, knowledge about the application of adequate methods is necessary across the three phases of the inquiry process: hypothesis generation, experimental design, and the evaluation of hypotheses (Dunbar & Klahr, 1989). These three phases of the inquiry process have in general been accepted and used for competence models of inquiry skills (e.g., Hardy et al., 2010).

While Dunbar and Klahr (1989) only distinguish three phases of the inquiry process, Fischer et al. (2014) introduce a model of scientific reasoning including eight epistemic activities: problem identification, questioning, hypothesis generation, construction and redesign of artifacts, evidence generation, evidence evaluation, drawing conclusions and communicating results. These activities are meant to be valid for scientific reasoning in all domains as well as for science education. Depending on the discipline and context, not all activities are equally relevant in all epistemic endeavors.

At the same time, Klahr (2002) himself expects the phases to blend in into each other, and they have been found to be highly correlated (Wilhelm & Beishuizen, 2003). Koerber, Mayer, Osterhaus, Schwippert, and Sodian (2015) could show with Rasch modeling that a unidimensional model fits best for their scientific reasoning test, which contained items on experimental design, data interpretation, and understanding the nature of science.

It should be noted that though the models by Dunbar & Klahr (1989) and by Fischer et al. (2014) are designed to fit several or even all scientific methods, psychological research has focused on people‘s abilities in experimentation (Lehrer & Schauble, 2006). Specifically, the control-of-variables strategy, which is central for scientifically adequate experimenting, has been looked at in detail (Chen & Klahr, 1999; Zimmerman, 2000), while little research examined other scientific methods.

2.1.1.2. Development of Scientific Reasoning Skills

According to the definitions introduced before, scientific reasoning is assumed to be a domain-general problem-solving strategy, but there is a debate on whether it develops because of increasing knowledge, or independently from knowledge (Klahr, 2002). As mentioned before, at least for physics it has been shown that domain knowledge may lead to better strategy use (Penner & Klahr, 1996). Schauble (1996) finds that neither strategies nor beliefs alone can fully explain the performance in an inquiry task, indicating interplay of domain-general and domain-specific factors. Zimmerman (2007, p. 173) also speaks of an –interdependent relationship” of investigation skills and domain knowledge.

The first ideas about children‘s development of scientific reasoning skills come from Piaget (1983). In his theory, children before the formal operations stage (12 years and older) are not able to reflect on their own thinking or draw logical conclusions and therefore cannot reason scientifically. For several reasons, Piaget‘s theory has been refuted in some aspects: for one, he severely underestimated children‘s abilities. Studies have shown that younger children are able to show logical reasoning if the task demands are less high (Bullock & Ziegler, 1999). Furthermore, cognition seems to develop less broadly and domain-generally, but rather specifically in distinct domains (Gopnik, 1996).

Still, researcher‘s view on children‘s scientific reasoning abilities stayed critical: in her review, Kuhn (1989) states that children are not able to differentiate between theory and evidence due to a lack of both metacognitive and strategic skills. For the development of their model of Scientific Discovery as Dual Search (SDDS), Dunbar and Klahr (1989) also tested preschool children. The children were much worse on the test than the adults, failing the overall aim in 90%, making mistakes in hypothesizing, testing and interpreting.

Sodian, Zaitchik, and Carey (1991) were the first to argue in favor of children‘s capabilities in scientific reasoning: They gave children a scenario in which they had to be able to use different strategies depending on whether the goal is effect production or hypothesis testing. A story was presented about two children who wanted to find out whether a mouse was big or small (find-out-condition: hypothesis test) and feed the mouse (feed-condition: effect production). If children are able to distinguish between hypothesis testing and effects production, they should choose different strategies in the two conditions. Over 50% of first-graders and over 80% of the second-graders could apply the right strategy in each situation and differentiate between a conclusive and non-conclusive test. Ruffman, Perner, Olson, and Doherty (1993) could show that 6-year-olds could understand a theory change based on new evidence.

Zimmerman (2007) summarizes the research on the development of scientific reasoning skills in primary and secondary school children. She states that children show an evolving understanding of science, but show problems when their prior belief is refuted by new evidence or when there might be no causal relation. Their mistakes often stem from the tendency to produce positive outcomes or from unsystematic procedure while planning experiments or recording results. However, children are already able to apply experimental strategies correctly if the task assignment is simple enough.

In recent years, the research on scientific reasoning development has been extended to preschool age. Koerber et al. (2005) replicated the results of Ruffman et al. (1993) with children aged between 4 and 6. At the same time, they showed that children had a harder time interpreting covariation data when there was no relation between the presented variables. However, when children were told that there might be no relationship, their performance increased again. In their second experiment, preschool children were asked to evaluate data about a content they had prior beliefs about. Even though this made it harder for them to correctly interpret the results, they were above chance level.

Piekny et al. (2013a) carried out both the task from Koerber et al. (2005) as well as the task from Sodian et al. (1991) with preschool children between 4 and 6 years of age in order to test their abilities in evidence evaluation and experimentation longitudinally. Their results on evidence evaluation mainly confirmed the results from Koerber et al. (2005), though their sample was worse in the interpretation of non-perfect covariation data. There was an increase of the performance from the age of 4 to 6. On the experimentation task, children were much worse, though they also showed an increase in performance between 5 and 6.

Croker and Buchanan (2011) did another experimentation task with their sample of children between 3 and 11, varying the context (good or bad outcome) and the consistency with the prior beliefs of children. They could show that already 4-year-olds choose the scientifically appropriate experiment (applying a control-of-variable strategy), but only if the evidence was consistent with their prior beliefs. This confirms that children‘s ability to coordinate theory and evidence correctly is still highly dependent on the task characteristics.

In order to evaluate preschoolers‘ control-of-variable strategy use, van der Graaf et al. (2015) constructed a test situation in which children could design experiments with a ramp hands-on. The testing was dynamic, adapting to children‘s individual performance level.

Almost 90% of the children (aged 4 to 6) managed to correctly design an experiment with at least two variables. 30 % of the children designed experiments with four variables correctly applying the control-of-variable strategy. These results are stunning, considering the difficulties children usually show when they have to construct the solution themselves instead of choosing the right option (Osterhaus, Koerber, & Sodian, 2015). Once more, specific demands and affordances of the task seem to make a huge difference for children‘s performance. In addition to the specific characteristics of the task, there are also abilities of the children themselves that have are related to children‘s performance in scientific reasoning tasks. They will be discussed in chapter 2.3. in terms of influencing factors.

2.1.2. Domain-Specific Scientific Thinking: Biology Understanding

Biology is the science of living organisms. Important core-ideas in biology are the understanding of living systems, the development of organisms throughout their life-span, and the relation between structure and function of living organisms (Deckelmann & Neuhaus, 2014). Therefore, the first and most important biological concept is the distinction between living (animate) and not living (inanimate) objects. On the one hand, babies already seem to be able to make this differentiation, as they expect different behaviors from animate and inanimate objects (Opfer & Gelman, 2010). At the same time, children have problems with the correct attribution of living to some objects, as even primary school children sometimes still believe that plants are not alive, but that other phenomena like wind, fire or clouds, are alive (Pauen, 1997; Piaget, 1978).

Piaget (1978) introduced the term animism to describe children_stendency to attribute life to inanimate objects. He interpreted this as a sign of children being in the preoperational stage when children are not able yet to reflect logically on their concepts and generalize egocentric views to other objects.

Further work on the topic has shown that children develop a non-animistic biology understanding, but in preschool still have problems differentiating between alive and real/existing/visible, and between dead and inanimate (Pauen, 1997). Two basic mistakes preschool children make is the classification of plants as non-living and natural kinds, such as fire or wind, as living. Reasons for that being discussed are using motion as main indicator, executive functions, and knowledge about the objects or beings.

Movement seems to be the most salient feature of animate objects, and while infants with seven months are able to differentiate between self- and externally initiated movement (Spelke, Phillips, & Woodward, 1995), they make mistakes with plants, sorting them as non-living, and fire or wind, which seem to move on their own and therefore appear alive. Even adults – and biology professors – tend to make these mistakes when asked for a classification under time pressure (Goldberg & Thompson-Schill, 2009).

Movement seems to be a heuristic all humans spontaneously use for the classification of living/non-living. This means that this heuristic has to be inhibited when correctly classifying objects. Inhibition, the ability to control one‘s behavior and impulses, is a core facet of executive functions (Diamond, 2013). This might be one reason why executive functions have proven to have an impact on children‘s non-animistic biology understanding (Zaitchik, Iqbal, & Carey, 2014). The authors found an effect of executive functions on children‘s biological reasoning even when controlling for age and IQ.

At the same time, knowledge deficits are also an important factor for accurate biological reasoning: both children and adults are better with the life status classification of familiar objects like cars than with objects that are further away from their everyday reality like planets (Richards & Siegler, 1986). Carey (1985) interprets children‘s false classification of plants not as a domain-general stage they are in, but as falsely applied domain-knowledge.

While Carey believes that children use their knowledge from naïve psychology and transfer it to animals, Inagaki and Hatano (2004) assume that children have a specific biological understanding, a vitalistic causality. This assumption is supported by results indicating children‘s basic understanding of some biological processes: they understand that living organisms grow and can heal themselves, that they can be ill and die (Inagaki & Hatano, 1996), and that bodily processes can at least partly be controlled (Inagaki & Hatano, 1993). They also have a basic understanding of inheritance and expect related organisms to share features (Johnson & Solomon, 1997). Gelman (2004) found indicators for the concept of essentialism in preschool children. I.e., children seem to understand that organisms carry a potential in them that will develop independently from their environment.

While many studies have investigated children‘s naïve concepts in biology, there are no studies looking into the effect these concepts might have on children‘s approach to scientific inquiry. For the field of physics, however, it has already been shown that children‘s concepts influence their strategies in an inquiry situation (Penner & Klahr, 1996). In general, it has been hypothesized that naïve concepts can either help or hinder people‘s learning and reasoning (Geary, 2008). Understanding that living organisms function differently from non-living objects is crucial for biology so that an impact of a non-animistic biology understanding on biological inquiry skills can be assumed.

2.2. Observation Competency

The focus of the developmental research has been on children‘s experimentation skills. The results derived from studies testing children‘s experimentation skills have been taken to estimate their general reasoning abilities. Other scientific methods were widely neglected. Meanwhile, the research on scientific thinking in biology has only looked at domain-specific conceptual understanding and not on the use of methods and strategies. I will now introduce observation competency as a scientific method, which combines domain-general and domain-specific aspects, being a general strategy but needs domain-specific competencies when applied in biology.

2.2.1. Conceptualization and Relevance of Observation

In scientific practice, there are different scientific methods that may have specific demands on the person applying them (Bybee, 2006; Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002). Wellnitz and Mayer (2011) differentiate the scientific methods experimentation, comparing, and observation, and explain how they differ in terms of the general, underlying themes (question, hypothesis, design, and data). They developed a questionnaire for students in order to check whether the scientific method is better understood as one concept or whether the three methods are different concepts. Their data fitted best to a three-dimensional model, suggesting that the three methods experimentation, comparing, and observation are in fact different concepts that have specific demands and affordances. This speaks for the necessity to investigate the different methods separately.

Looking closer at the concept of observation, there are two different conceptualizations. The first defines observation as a process underlying the scientific method in general (Kosso, 2011; Oguz & Yurumezoglu, 2007). Thus, observation is not only a relevant scientific method in itself; it is also an important process during the application of other scientific methods: when experimenting, the results in the different conditions still have to be observed. Often, observations also stand at the beginning of a scientific process, when an interesting observation leads to a research question or hypothesis that then may be tested (Bybee, 2002). In general, ―all scientific knowledge must be based on observation” (Kosso, 2011, p. 7), and it is an essential strategy throughout the inquiry process (Oguz & Yurumezoglu, 2007).

The second conceptualization defines observation as a unique research method, standing next to other methods of knowledge acquisition, such as experimentation (Kohlhauf et al., 2011; Wellnitz & Mayer, 2011). This method is particularly relevant for biology. Darwin developed his ideas on evolution by “long-continued observation of the habits of animals and plants” (Darwin, 1887, p. 120), basing it on the ideas of inductive reasoning. Ever since, observation has been a central method for biology as a science (Wellnitz & Mayer, 2011). However, the relevance of observation as a research method is not limited to biology. It also plays an important role in, for example, social sciences; both qualitative and quantitative observations of behavior are a typical method for data acquisition (Bortz & Döring, 2013). In chapter 2.3.2.3 the application of observation in the field of psychology for measuring emotions will be introduced shortly.

As a research method, observation is a convenient method to realize with young children, as they already use observations a lot in order to make sense of the world (Rogoff, Paradise, Arauz, Correa-Chávez, & Angelillo, 2003). Of course, these observations are not on the level of scientific observations (Bortz & Döring, 2013; Kohlhauf et al., 2011), but they are already used to implicitly constructing reality by observing their environment (Perner, 1991). This makes observation a good starting point for science teaching, before proceeding with more sophisticated research methods like experimenting. Furthermore, in the practical school context we mainly do not find experiments, but observations (Eberbach & Crowley, 2009). Looking at the suggested inquiry activities in the syllabus for biology in primary schools in Bavaria, only two are real experiments (N. Kümpel, private communication, August 2016). Usually, it is too complicated to conduct an experiment with several conditions in the classroom environment, and it is feasible to investigate many relevant questions with observations in biology education (Wellnitz & Mayer, 2012).

This study follows the latter conceptualization, treating observation as a research method that follows the general steps of scientific inquiry (Fischer et al., 2014; Wellnitz & Mayer, 2011). Observation is a valid and important method on its own, especially in biology. At the same time, the role of observation as a process underlying all scientific methods has to be recognized as well. The results of this study primarily relate to observation as a discrete method but are not limited to this conceptualization.

2.2.2. Observation Competency Models

Observation is a complex research method (Oguz & Yurumezoglu, 2007).The adequate application of this method needs experience, skills and knowledge (Eberbach & Crowley, 2009). In that sense, it is a competency that has to be developed and can be fostered (Kohlhauf et al., 2011). A competency is defined as having the cognitive (and sometimes also social and emotional) prerequisites for mastering a complex task (Weinert, 2001). Hence, having observation competency means to have the necessary skills to make scientifically sound observations. Norris (1984) defines observation competency as the ability to make observations well, report them well, and correctly assess reports of observations. In recent literature, there have been two studies that look into the structure of observation competency (Eberbach & Crowley, 2009; Kohlhauf et al., 2011).

The first is the literature review by Eberbach and Crowley (2009). In order to understand what defines the quality of observations and which skills compose observation competency, they summarized several studies that compare laypersons‘ and experts‘ behavior in an observation situation. They found that the two groups use different strategies in all phases of the inquiry process: they ask different questions, notice, filter, and reason differently. The experts ask more specific questions and go on questioning and noticing details. Meanwhile, laypersons often ask wrong questions, miss important details, and do not document their observations adequately. Based on their findings, the authors formulate four components of scientific observation, which are displayed in Table 1: noticing of relevant objects or circumstances, expectations and coordination of observations and theories, observational records (cognitive, physical or virtual), and productive dispositions – that is the extent of the engagement with the observed object. They also propose that there are three states of observation: everyday, transitional, and scientific.

Table 1. Observation Competency Model by Eberbach and Crowley (2009)

	Noticing	Expectations	Records	Productive Disposition
1 Everyday novice	- Noticing irrelevant features - Describing few features	- Vague expectations - Confuse evidence and beliefs	- No recording of observations	- Incidental observations
2 Transitional	- Noticing more relevant features - Use and describe features	- More explicit expectations - Expectations may be scientific or everyday	- Recording of observations (e.g., personal journal)	- Intentional observations and seeking information
3 Scientific	- Notice, describe and structure relevant features - Chunk observational information	- Explicit hypothesis in line with theory - Coordination of hypothesis and evidence	- using established recording procedures - Organize and analyze recordings	- Persistent, sustained engagement

Adapted from ―From Everyday to Scientific Observation: How Children Learn to Observe the Biologist‘s World," by C. Eberbach and K. Crowley, 2009, *Review of Educational Research, 79(1),* p. 56

Though Eberbach and Crowley (2009) assume that these components and levels are domain-general, domain-specific knowledge and concepts play a major role for observations in their model. The model has been utilized in recent studies on the practical application of observation as an inquiry method in school. These studies mostly focus on tools needed throughout the observation: the implementation of technological tools to record data enhanced students‘ observation behavior (Lommen, 2012; Tokarczyk, 2015). With younger students, the use of drawings helped them retain the information needed for the interpretation of observations (J. E. Fox & Lee, 2013).

Kohlhauf et al. (2011; 2013) constructed a competency model for observation, identifying the following components as important for the quality of observation: describing specific and unspecific details, questioning, hypothesizing, testing, and differentiating between the observations and the interpretation. In line with Eberbach and Crowley (2009), they also differentiated three ascending levels, incidental observation, unsystematic observation, and systematic observation. The model is displayed in Table 2.

In order to validate the model, they analyzed the observation behavior of 110 study participants aged between 4 and 29 years. The results showed that there are actually three dimensions: describing details, scientific reasoning (questioning, hypothesizing, and testing), and interpretation. It is notable that they could not differentiate between questioning, hypothesizing, and testing, while other studies could (Wellnitz et al., 2012). At the same time, the interpretation is its own dimension and not part of an overall scientific reasoning competency as found for example by Koerber et al. (2015).

Table 2. Observation Competency Model by Kohlhauf et al. (2011)

	Describing	Scientific reasoning[2] asking/ assuming/testing	Interpreting
1 Incidental	- Perception without details	- Nonexistent or passive	- Nonexistent or passive
2 Unsystematic	- Describing few details - little focus on specific details	- Partly use of research question, hypothesis, and testing of hypotheses	- Interpretation of Observations - Observation and interference are not separated
3 Systematic	- Describing many details - strong focus on specific details	- Working with research question, hypothesis, and testing of the hypothesis	- Observation and interference are separated

Adapted from ―Influence of Previous Knowledge, Language Skills and Domain-specific Interest on Observation Competency" by L. Kohlhauf, U. Rutke and B.J. Neuhaus, 2011, *Journal of Science Education and Technology, 20(5),* p. 670

Comparing the studies by Eberbach and Crowley (2009) and Kohlhauf (2013), there are both parallels as well as differences in the conceptualization of observation competency. While Eberbach and Crowley (2009) differentiate between noticing and recording observations, this is subsumed in the description of details in Kohlhauf (2013). Vice versa, the differentiation between questioning, hypothesizing, and testing on the one hand and

[2] In this study, both children's domain-specific scientific reasoning and their observation competency will be investigated. Consequently, it is not feasible to call one dimension of observation ―scientific reasoning". In the following, this category will be renamed to ―inquiry".

interpretation on the other cannot be found in Eberbach and Crowley (2009). The only aspect of scientific reasoning represented in their model is the category of expectations, covering the coordination of theory and evidence: while on the novice level of expectations, observational evidence is confused with one‘s beliefs, experts are able to coordinate their theoretical expectations with the new evidence (Eberbach & Crowley, 2009). Finally, the aspect of productive dispositions cannot be found in Kohlhauf (2013), while it is rather important for Eberbach and Crowley (2009). Independent from the examined dimensions, both studies found deficits in children‘s observation competency and showed that lay adults are not necessarily on the highest level of observation competency.

Another difference is the purpose of each of the models. Eberbach and Crowley (2009) reviewed literature from science education, containing many qualitative studies and small interventions, and the model they develop is meant to promote the further development of fostering scientific observation competency. While this is also a long-term goal of Kohlhauf (2013), their main and first purpose is diagnostic: the development of both a model and an instrument for identifying the level of observation competency, ideally across age groups. These different purposes may also have led to the different weighting of factors in their models.

This study will follow the model of Kohlhauf (2013), as the goal is here, too, to identify children‘s observation competency at a specific point in time, and investigate influencing factors. For measuring children‘s engagement in the observation situation, a different measure is used, assessing children‘s level of involvement (see 2.3.2.4). However, the value of the model by Eberbach and Crowley (2009) for other purposes, such as practical science education, has to be acknowledged. There, training of the use of established disciplinary recording procedures to document results seems very promising.

2.3. Factors Related to Children's Scientific Thinking

For both observation competency and domain-general scientific reasoning, there are factors that have been empirically shown or theoretically discussed to be related to these competencies. One goal of this study is to understand the structure of observation competency and domain-general scientific reasoning by identifying both cognitive and affective constructs influencing them. Since most of the studies have correlative designs, the direction of the relation is not clear. However, the direction from general competencies influencing the more specific competencies is usually assumed, meaning that the discussed factors are assumed to have an influence on the scientific thinking competencies.

2.3.1. Cognitive Factors

The main investigated cognitive factor in psychological research is intelligence. While studies usually find relations between academic competency measures and intelligence (e.g., in the PISA study), they can be differentiated conceptually: while intelligence is a general, largely innate ability to solve new problems, competencies are usually for a specific context and can be learned (Hartig & Klieme, 2006). Therefore, it can be expected that children's intelligence helps them making good observations, but that other factors like knowledge, interest, and situational factors are also related to their performance.

The same holds true for domain-general scientific reasoning: while influenced by intelligence, it is a specialized skill that is applied in specific contexts of science and science learning. Bullock, Sodian, and Koerber (2009) found correlations between scientific reasoning abilities and general intelligence in their longitudinal study, but earlier scientific reasoning abilities were a better predictor for later scientific reasoning than intelligence. The data in the study of Mayer, Sodian, Koerber, and Schwippert (2014) fitted best to a model assuming that intelligence and scientific reasoning are two separate, albeit related constructs. At the same time, they found problem-solving skills and spatial reasoning to be predictors for primary school children‘s scientific reasoning.

One problem with the measurement of intelligence is the broadness and variability of definitions (Duggan & Garcia-Barrera, 2015; Sternberg et al., 2000). In this study, I will focus on other, more specific factors that have proven to be crucial for young children‘s cognitive development. Theory of mind, executive functions and language abilities will be introduced and investigated as potential predictors of both general scientific reasoning as well as observation competency. Prior knowledge will be discussed as a specific predictor of observation competency.

2.3.1.1. Theory of Mind

Theory of mind is defined as the ability to attribute mental states like desires, beliefs, feelings or intentions to oneself and to other people (Perner, 1991). It starts developing around the age of 3 to 5 years (Wellman, Cross, & Watson, 2001). When children are capable of social perspective taking, a distinction between reality and appearance is possible. Therefore, theory of mind serves not only the facilitation of the interaction with other people but also helps children with the discovery of their environment (Astington, 2000).

Though there is no evidence relating theory of mind development to scientific observation, there are arguments for expecting a relationship between the two. Scientific observation includes per definition the reflection on the observations (Bortz & Döring, 2013). In order to reflect on observations, however, the observer must be able to represent the observation and at the same time understand that this representation is not the only possible representation, but only one construction of reality (Reich, 2001). Between 1 and 1 ½ years, children become able to construct multiple models of reality, and later their theory of mind also allows them to represent multiple models (Perner, 1991). Consequently, theory of mind can be expected to be a prerequisite for scientific observation.

Theory of mind also plays an important role in the development of domain-general scientific reasoning abilities. Kuhn (1999) connects theory of mind, metacognitive knowledge, and epistemological knowledge: when children do not understand that another person may have a belief the child knows is false (theory of mind/ metacognitive knowledge), they are realists in the sense that, in their perception, beliefs and mental concepts of people are a copy of the actual reality (epistemological knowledge). Thus, it does not make sense for them to question beliefs or test them, which means theory and evidence are the same for them and they are not able to reason scientifically (Kuhn, 1999, 2000). This leads to the assumption that theory of mind must be a prerequisite for scientific reasoning. In fact, Piekny et al. (2013b) could confirm that in their study. They tested children‘s theory of mind and their scientific reasoning with the mouse task from the study of Sodian et al. (1991), both at the age of 4 and 5. They could show that children‘s theory of mind at age 4 predicted their scientific reasoning abilities at age 5, but not the other way round.

2.3.1.2. Executive Functions

Researchers have also related children‘s domain-general scientific reasoning skills to their executive functions. Executive functions are defined as “top-down mental processes needed when you have to concentrate and pay attention when going on automatic or relying on instinct or intuition would be ill-advised, insufficient, or impossible” (Diamond, 2013, p. 135). Usually, researchers subsume the three aspects inhibition, working memory and cognitive flexibility (also known as set shifting) under this term.

There is evidence that executive functioning skills are related to general reasoning abilities (Decker, Hill, & Dean, 2007; van der Sluis, de Jong, & van der Leij, 2007). However, little research investigated the relation to scientific reasoning. From a theoretical point of view, the first two aspects seem to be relevant for scientific reasoning: one has to inhibit one‘s own theories in order to plan an adequate testing (Kuhn & Franklin, 2006); the working memory is necessary for keeping the hypothesis, design, and results in mind. Working Memory has indeed shown to have an impact on students‘ science understanding (Gathercole, Pickering, Knight, & Stegmann, 2004). Mayer et al. (2014) did not find a correlation between their scientific reasoning paper-pencil test and inhibitory control. The authors argue that this might be due to the format of their test, in which children did not have to inhibit their prior beliefs as much as would be necessary for a hands-on science activity with the entire circle of scientific discovery. In fact, van der Graaf et al. (2016) found a relation of both working memory and inhibition with preschoolers‘ performance in an inquiry situation.

Osterhaus, Koerber, and Sodian (2016) also investigated the connections between executive functions, theory of mind, and scientific reasoning. They found a relation, specifically between executive functions and scientific reasoning, while the theory of mind was specifically linked to children‘s nature of science understanding.

2.3.1.3. Language Abilities

Language abilities are a common control variable in studies with young children. After all, many test instruments are language based, and therefore children‘s language abilities will confound their measure of other cognitive variables. This is also true for tests of scientific skills when the instruction contains complicated sentence structure and the outcome measure is derived from verbal answers. However, language is not only a confounding variable but can also be seen as a motor of development. According to Vygotsky‘s sociocultural assumption that ―intermental (social) activity will promote intramental (individual) intellectual development” (Vygotsky, 1978, p. 86), it can be assumed that children with better language abilities experience more and better learning situations to improve their cognitive skills. More recent studies show that humans have further developed social-cognitive skills in comparison to apes, which reinforces the hypothesis that human intelligence is formed culturally (Herrmann, Call, Hernández-Lloreda, Hare, & Tomasello, 2007). In the theory of mind literature, language has proven to be an important instrument in the development of false belief understanding (Lohmann & Tomasello, 2003).

Language as motor of social interaction is also important for developing children‘s observation competency (Johnston, 2009). Eberbach and Crowley (2009) state that children need specific tools in order to make scientifically sound observations. These tools can be physical objects, like measurement instruments, or virtual tools for recording data, but language can also be seen as such a tool, necessary for being able to notice, record and communicate observations properly. Accordingly, Kohlhauf et al. (2011) found an impact of preschoolers‘ language abilities on their observation competency.

The relation between verbal intelligence and scientific reasoning has been found both in cross-sectional and longitudinal designs (Bullock et al., 2009; Mayer et al., 2014). Research on children with language impairment suggests that the understanding of causal connectives is specifically important for the development of scientific reasoning (Matson & Cline, 2012). Other studies have shown specific effects of language abilities on learning about science (e.g., Mercer, Dawes, Wegerif, & Sams, 2004).

2.3.1.4. Prior Knowledge

While in psychology researchers usually investigate the influence of general factors like intelligence, science education focuses on specific requirements for specialized tasks. One basic requirement (and outcome, for that matter) is prior knowledge. Therefore, it is the main factor being discussed to influence observation competency. The review by Eberbach and Crowley (2009) confirms that the quality of an observation is strongly related to the observer‘s knowledge in the domain. It is relevant for all steps of the inquiry process: domain knowledge is necessary for asking the right questions, planning adequate testing situations, documenting meaningful details and drawing the right conclusions from the data (Alberdi, Sleeman, & Korpi, 2000). As we also look at biology understanding in this study, our focus with prior knowledge lies on factual knowledge about the observed object, and not a general understanding of the field. Kohlhauf et al. (2011) found that prior knowledge about the object of investigation had a positive impact on the observation competency.

2.3.2. Affective Factors

Affective states have been found to have an impact on people‘s reasoning across several contexts and domains (e.g., Forgas & Vargas, 2000; Moors, Ellsworth, Scherer, & Frijda, 2013; Pekrun & Linnenbrink-Garcia, 2012) However, there are no studies yet investigating their specific effects on observation. In the following, the relevant terms will be defined and the available methods of measurement discussed. The focus will lie on the instrument for measuring emotional well-being and involvement (Laevers, Kog, & Vandenbussche, 1997), which is specifically suited to measure the affects of young children via observation.

2.3.2.1. Conceptualization of Affective Factors

For measuring affective states, the terms of emotion, mood, motivation, and engagement become relevant. All of these can be seen as processes that move organisms toward action (Bradley & Lang, 2000). Moods are less intense emotional states (Pekrun & Linnenbrink-Garcia, 2012); usually, only positive and negative mood are distinguished. Emotions and moods are often summarized under the term of affects. According to the control-value-theory, affects can be positive or negative as well as activating or deactivating (e.g., Pekrun, 2006). Anger is an example for a negative activating emotion; happiness or excitement are positive activating affective states, sadness and boredom are negative deactivating affective states and relaxation is a positive deactivating affective state (Feldman Barrett & Russell, 1998).

The concepts of emotion and motivation are tightly interwoven. S. S. Tomkins (1962) defined emotions as the primary motivational system for human behavior. Izard and Ackerman (2000, p. 262) argue that "emotion motivates and organizes perception, cognition, and actions". The control-value theory also states that motivational (and cognitive) processes precede emotions, follow emotions, but are also an integral part of emotions themselves (Pekrun, 2006). Thus, while motivation and emotion can be separated conceptually, it is hard to keep them separated empirically.

Engagement is a concept that is often interpreted as a mediator between emotions and achievements (Pekrun & Linnenbrink-Garcia, 2012). It is defined as task participation or enjoyment but also has cognitive and motivational characteristics, such as investment, perseverance, and use of deeper strategies (Fredricks, Blumenfeld, & Paris, 2004). Sinatra and Taasoobshirazi (2011) postulated that it needs motivation for engagement, which then leads to conceptual change.

2.3.2.2. Affects and Reasoning

An influence of emotions on scientific reasoning is generally assumed (Pekrun & Linnenbrink-Garcia, 2012). Several types of emotions can be distinguished that may have an effect on reasoning processes: epistemic emotions, achievement emotions, topic emotions, social emotions, and incidental emotions or moods (Fischer et al., 2014).

Epistemic emotions are emotions that are directly part of epistemic activities during the inquiry process (Pekrun & Stephens, 2010). Surprise and curiosity are the most well known epistemic emotions, but frustration or enjoyment can also be epistemic emotions when they appear throughout the reasoning process (Pekrun & Linnenbrink-Garcia, 2012). Epistemic curiosity can be conceptualized as a state or a trait and is related to the motivational concept need for cognition (for an overview on curiosity see Jirout & Klahr, 2012).

As scientific reasoning processes will have outcomes – either success or failure – that may be judged by us or others, it can be expected achievement emotions emerge as well. Positive achievement emotions are hope (for positive outcomes) and pride; their negative counterparts are anxiety/hopelessness and shame (Pekrun & Linnenbrink-Garcia, 2012). The strength of the achievement emotions depends on the perceived importance of the outcomes (Fischer et al., 2014).

The topic under investigation itself can also trigger emotions. While interest in a topic usually has positive effects on reasoning or learning processes (Ainley, 2006), boredom has negative effects on students‘ performance, and a bad performance can lead to more boredom (Pekrun, Hall, Goetz, & Perry, 2014). In biology lessons, emotions such as fear or disgust can arise when working with live animals or talking about topics like digestion, sexuality or death (Tunnicliffe & Reiss, 1999). Dräger and Vogt (2007) could show that detailed examination and investigation with spiders in the classroom led to reduced disgust and anxiety and increased interest in and even sympathy for spiders.

When the scientific reasoning happens in a group situation, social emotions like love, hate, admiration, envy, contempt, or empathy can occur and influence the reasoning process (Fischer et al., 2014).

Finally, emotions and moods a person already had before they entered the inquiry situation, like stress or happiness, can also have an effect on their reasoning and behavior (Fischer et al., 2014).

There are no studies investigating the specific influence of emotions on observations, but several studies investigate the influence of emotions on science classes, and some of the studied outcome variables, like noticing details or categorization of objects, are relevant for observation as well. Fleer (2013) did a qualitative analysis of the interactions between the teachers and children. They observed that teachers emotionally charged learning situations, e.g., by embedding them in stories, in order to focus children on specific details. This intensified children‘s emotions and led to more scientific noticing on the side of the children.

Although the results are not completely conclusive, positive mood has been linked to higher engagement and achievement (Linnenbrink, 2007). Positive mood has been found to have a specific impact on learning achievements in science lessons (Laukenmann et al., 2000). Murray, Sujan, Hirt, and Sujan (1990) induced either positive or neutral mood and gave their participants a categorization task. Participants in a good mood noticed more details than the participants in a neutral mood, identifying both more similarities and more differences. This result is contrary to a lot of research, where positive mood usually has been linked to holistic thinking, while negative mood would go along with a better reception of details (Pekrun &

Perry, 2014). However, this may be due to the fact that many studies only distinguish between positive and negative affect and not between activation and deactivation. In the case of the study of Murray et al. (1990), it can be assumed that the positive-mood-group was more activated and therefore performed better on the task than the neutral group. As they only compared positive and neutral moods, it is not clear whether a negatively (activated) group might have been even better than the positive (activated) group.

All studies on the relationship of emotional and motivational factors with scientific reasoning or science activities have been with students at least on secondary level, except the study by Fleer (2013), who worked with qualitative research methods. One reason for this is the problems with measuring the emotions of younger children. The available measuring methods and their advantages and disadvantages will be discussed now.

2.3.2.3. Measuring Affective States

The empirical investigation of emotions can be divided into three basic methods: observation of emotions, self-assessment, and physiological measures.

Typical physiological measures are heart rate, EEG, and cortisol levels. While their objective character makes them attractive, the measurements can be difficult to interpret (N. A. Fox & Calkins, 1993). Furthermore, the measurement itself can also limit the mobility and consequently the options of tasks from a practical point of view.

Self-assessments of emotions can take a variety of forms: they can relate to a specific emotion or a list of emotion/moods, they can assess emotional states or traits, and they can aim at measuring the emotion in the moment or in retrospective. The assessment itself can be structured or unstructured, oral or written, qualitative or quantitative with one single item or several items (Pekrun & Bühner, 2014). While for adults and older students, self-assessment has proven to be a valid method, there is no good instrument for measuring younger children‘s emotions via self-assessment. Preschoolers‘ ability to assess their emotions still varies a lot and is often strongly biased toward positive feelings (Fabes, Eisenberg, Nyman, & Michealieu, 1991).

The third option is the observation of facial expressions and behavior. In general, observation can be seen as specifically suitable for measuring preschool children‘s emotional states, as they are not that skilled at internalizing emotional expressions (Holodynski, 2005). As subjectivity can be a problem, especially when observing something as complex as emotions, a good training of the observers and control of interrater reliability is crucial. There are several approaches for measuring emotions by observation. The inference of specific emotions from facial micro-expressions originates from Paul Ekman (Ekman, 1973) and has led to the development of several measurement instruments, using both human raters and computer-based interpretation of facial expressions (for an overview of research results see Keltner, Ekman, Gonzaga, & Beer, 2003). Another approach is to assess a broader range of affects, taking into account facial, verbal and behavioral expressions and cues. An example for this is the Leuven Scales for Emotional Well-Being and Involvement (Laevers et al., 1997). These scales have been developed to measure preschool children‘s affective states and their impact on learning. Thus, they are well suited to be used for measuring affective states in this study and will be introduced in the following.

2.3.2.4. Leuven Scales for Emotional Well-Being and Involvement

Laevers (1993) introduced his deep level learning model, illustrating the theory with results from a study on preschool children‘s understanding of swimming and sinking in an inquiry situation. In this model, context factors like the person of the teacher, the environment, and the activity, influence the children‘s learning process, which then results in an outcome (Laevers, 2000). As the important process variables, Laevers identified involvement and emotional well-being as the relevant mediators. The outcome, emotional well-being, and involvement as conceptualized by Laevers will now be further explained and related to other concepts from research on emotion and motivation. An overview of the model and its relation to other concepts is displayed in Figure 1.

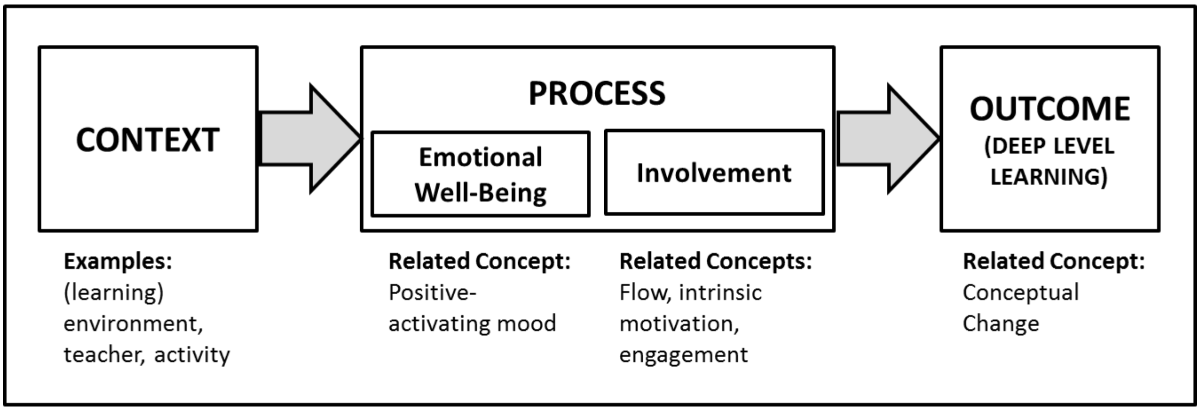


Figure 1. Deep Level Learning Model and its Relations to Other Concepts
Adapted from ―Forward to Basics! Deep-Level-Learning and the Experiential Approach" by F. Laevers, 2000, *Early Year, 20(2),* p. 24.

The ideal and aspired outcome is deep-level learning: ―the development from elementaristic, mechanistic and gross images of the world to more differentiated structures that articulate more elements of reality and their dynamics [and] can be documented for different domains" (Laevers, 1993, p. 57). This definition shows parallels to the idea of conceptual change, which is the idea that learners already have naïve theories and shift from these theories to a new, more sophisticated theory (Carey, 1985). However, Laever‘s conceptualization of deep-level learning is more fundamental and less specific for different domains.

Emotional well-being is conceptualized as ―the degree to which children feel at ease, act spontaneously, and show vitality and self-confidence" (Laevers, 2000, p. 24). Since it is a rather long-lasting state, working in the background, it can be categorized as mood. Applying the control-value-theory (Pekrun & Perry, 2014), in which the dimensions of valence (positive - negative) and activation (activating – deactivating) are distinguished, emotional well-being can be classified as positive (*feel at ease*) activating (*show vitality*) mood. In line with the results of prior research, a positive impact on performance in reasoning tasks can be expected (Laukenmann et al., 2000; Linnenbrink, 2007; Murray et al., 1990).

The description of involvement includes concentration, intensity, and endurance during the task (Laevers, 2000). The author himself relates it to both intrinsic motivation and flow. Intrinsic Motivation is the desire to explore and seek out learning situations in order to extend one's knowledge and competencies (Ryan & Deci, 2000). Flow is defined as a state during which people experience deep enjoyment, creativity, and involvement in the task (Csikszentmihalyi & Csikszentmihalyi, 1992). There are also parallels between the conceptualization of involvement and engagement: both include enjoyment, motivation to persevere and cognitive characteristics, such as the use of deeper strategies (Fredricks et al., 2004). Additionally, engagement has also been assumed to lead to conceptual change (Sinatra & Chinn, 2012). As engagement is a mediator between emotions and achievements (Pekrun & Linnenbrink-Garcia, 2012) it can be expected that there is a mediation between emotional well-being and performance through involvement. Laevers (2000) himself describes well-being as a prerequisite for involvement.

Based on his theory, Laevers et al. (1997) developed an observation tool to measure children‘s emotional well-being and involvement. This instrument has mainly been used to identify environmental factors needed for children to be involved so that deep level learning can take place. Goldspink, Winter, and Foster (2008) conducted several studies on students‘

emotional well-being and involvement from preschool up to 12-year-olds, using both observation and self-report measures with the older children. Their results show that educators‘ pedagogical attitudes and the quality of their relationship with the children are related to children‘s engagement in learning. Furthermore, the pedagogical philosophy and overall quality of the school also had an impact on the level of children‘s involvement and well-being. Declercq (2014) applied the instrument in preschool settings in South Africa. The overall values were low. The main reason for that was that too few adults supervised children. However, some schools showed much higher values for well-being and involvement, despite the difficult circumstances, implying that the quality of the teaching and learning environment can make a difference. Aydo an, Farran, and Sa s z (2015 observed teachers and children in 45 classrooms. They found that the teachers‘ instructional practices and a positive emotional tone in the classroom had large effects on children‘s involvement.

There is also one study investigating the effects of involvement on the learning outcome. Pascal, Bertram, Mould, and Hall (1998) assessed children‘s involvement as well as their school grades and found that involvement explained 16% of the variance in the grades.

So far, the Leuven scales have shown to be reliable (Declercq, 2014; Laevers et al., 1997) and valid instruments for assessing children‘s affective states (Goldspink et al., 2008). The fact that they mix up several facets of emotion, mood, and motivation can be seen critical. However, this is probably the reason why they work so well in practical research, with good values for interrater reliability and correlations with self-assessment instruments (Goldspink et al., 2008). Since it is difficult to differentiate the different theoretical concepts empirically (Pekrun, 2006), it makes sense to measure them in one combined instrument.

2.4. Interventions in Early Education

> All true education in training and instruction should, therefore, at every moment, in every demand and regulation, be simultaneously double-sided – giving and taking, uniting and dividing, prescribing and following, active and passive, positive yet giving scope, firm and yielding. (Fröbel, 2012, p. 14)

In 1840, Friedrich Fröbel founded the first *Kindergarten* in Germany – before, comparable institutions were seen as places where children were kept safe, but not as places where knowledge and skills were developed (Grell, 2013). Since then, a lot has happened concretizing the formal education of children in preschool contexts. Since 2006, all German states have educational guidelines (*Bildungspläne*) for preschool (Diskowski, 2009). Preschool[3] has become a relevant learning environment, especially since most children nowadays attend it: in March 2015, over 95% of children in Germany between 3 and 6 visited day care facilitations (Statistisches Bundesamt, 2016).

This chapter will look at the general effectiveness of early education interventions, the specific content topics of these interventions, and their design and pedagogical style. As this study has taken place in Germany, the focus will lie on German literature, though important international results will be integrated as well.

---

[3] Across nations and educational systems, the use of the terms kindergarten and preschool differs. In this thesis, I will use the term preschool for children from 3-6 before they start school, as this is the phase children in Germany go to kindergarten or comparable institutions.

2.4.1. Effectiveness of Early Education Interventions

Before discussing specific intervention programs, the question arises whether interventions in early childhood show any effects at all, and how long these effects can be found. Preschool intervention programs aim at both creating equal opportunities for children from socially deprived backgrounds (Nores & Barnett, 2013), as well as improving all children‘s academic competencies (Anders & Roßbach, 2013).

There is some research on the effectiveness of going to preschool in general. Camilli et al. (2010) conducted a meta-analysis of all studies investigating preschool intervention effects from 1960 to 2000. They also find a significant, stable, positive effect of early education, especially on cognitive skills, but also on social skills and their school progress. Overall, studies find a consistent positive effect of both the duration and the quality of children‘s attendance to preschool, as Roßbach, Kluczniok, & Kuger (2009) summarize in their review paper. This effect can be found for all children, which also means that preschool does not directly have a compensatory effect, but that for compensation of risk factors specific and more intensive interventions are needed. Studies differ in their results on how long the positive effects can be found, as some find it until the end of primary school while others only measured short-term effects. Different reasons for this can be discussed: either the effect really vanishes, or it is undermined by bad quality of education in schools (Sammons et al., 2009). Either way, the effects of home and family remain crucial: Stutz (2013) showed that 13 years later the main factor for school success was not the participation in preschool interventions, but the education level of the parents.

Concluding, it can be said that preschool interventions are not a wonder weapon for erasing differences between children or boosting their development over their whole school career. However, literature indicates that they have positive effects on children's cognitive development into the first years of primary school.

2.4.2. Topics of Early Education Interventions

Looking at the educational guidelines for early education in Germany, there are rarely defined norms children have to reach at the end of the preschool phase (Diskowski, 2009). Instead, they formulate specific topics and themes children can and should be fostered in. The Bavarian guideline, which is the most extensive and detailed, formulates both basic competencies of children (personal, social, learning related, and change related competencies) and general topics of education (values, language, maths and sciences, arts, and sports) (Staatsinstitut für Frühpädagogik München (IFP), 2012). A more detailed list of the competencies and topics can be found in Table 3.

Looking at the section describing science and technology in more detail, there are specific learning goals for this field (Staatsinstitut für Frühpädagogik München (IFP), 2012). As mentioned in the introduction, the focus lies on the domain of physics, with learning goals revolving around understanding electricity, magnetism, time, or gravity. Some goals, however, focus on biological contexts: children are supposed to collect, sort, organize and describe nature materials like leaves, fruit, and blossoms. Another goal is to observe processes in nature and derive questions from these observations. Additionally, there are also goals aimed at domain-general inquiry skills: children are supposed to develop a basic understanding of measurement methods, conduct experiments, and formulate hypotheses and test them adequately. These activities are in line with researchers' ideas about how to foster observation competency: the learning goals include the components of observation competency describing details, questioning and hypothesizing (Kohlhauf, 2013), have a focus on organizing materials (Johnston, 2009) and on the use of measurement tools (Eberbach & Crowley, 2009). One crucial aspect of observation competency that is missing in these learning goals is interpretation, specifically the differentiation between observation and inferences (Kohlhauf, 2013).

Table 3. Competencies and Topics in the Bavarian Educational Guidelines

Basic Competencies	
Personal	- Self-Perception - Motivational Competencies - Cognitive Competencies - Physical Competencies
Social	- Social Competencies - Values and Orientation - Responsibility - Participation
Learning Related	- Methodological Competencies
Change Related	- Resilience
Topics	
Values	- Values and Religion - Emotions and social relations/ conflicts
Language	- Language and Literacy - Media and Communication Technology
Maths and Sciences	- Maths - Science and Technology - Nature and Environment
Arts	- Art and Culture - Music
Sports	- Moving, Rhythm, Dancing, Sports - Health

(Staatsinstitut für Frühpädagogik München (IFP), 2012)

These educational guidelines are given to the kindergarten teachers and are the basis for their everyday work in kindergartens. For some topics, however, there are specific training programs, which are being used in the preschool sector and have also been evaluated. For these programs, Hasselhorn and Schneider (2016) distinguish between domain-general and domain-specific competencies that can be fostered.

Domain-general competencies are metacognition, self-regulation, and working memory. The intervention program “Red Light, Purple Light”, for example, uses well-known playgroup games, but increasing their difficulty step by step by introducing new rules, like doing the opposite of what has been said (Tominey & McClelland, 2011). Research has shown that this 16-week intervention works very well with preschool children and shows positive effects not only on their executive functions, but also on their academic skills (Schmitt, McClelland, Tominey, & Acock, 2015).

The focus of research on domain-specific competencies lay on prerequisites for learning reading, writing, and arithmetic. The fostering of reading and writing skills is usually subsumed under the term of literacy, defined as “the activities and skills associated directly with the use of print[,] primarily reading and writing” (Snow, 1983, p. 166). Programs are either code-focused or meaning-focused (Connor, Morrison, & Slominski, 2006) and show good effects on children‘s literacy skills (Wasik, Bond, & Hindman, 2006; Wolf, Schroeders, & Kriegbaum, 2016). Children‘s mathematical competencies have also been trained successfully with learning programs, especially when the task difficulty meets children‘s developmental status (e.g., Krajewski, Renner, Nieding, & Schneider, 2009).

Scientific skills have become more important in the recent years as well. It is not only part of the educational guidelines for preschool in German states (e.g., the Bavarian one, see Table 3), but there are also nationwide initiatives aiming at improving science education in kindergartens. One of the most well known examples is the “Haus der Kleinen Forscher” (*house of little scientists*) (Hecker & Tansaway, 2008), a foundation that offers instructions for experiments as well as training for preschool teachers. The instructions for the different activities and experiments are story-based and follow a general structure in order to facilitate the use for the preschool teachers. The activities are not sorted into classical science domains, but according to the topics astronomy, health, acoustics, communication, living spaces, light and colors, air, magnetism, mathematics, sustainability, carbon dioxide, electricity, technology, water, and time (Stiftung Haus Der Kleinen Forscher, 2016). While the program focuses on experimentation, the activities do not all qualify for being called experiments; there are experiments, observations, and general activities, but all aimed at fostering children‘s science understanding. The first results from the running evaluation are promising, though the general approach naturally leads to differences in how preschool institutions apply the materials (Anders, Ballaschk, & Tietze, 2014).

Another big science program for German preschools was ―Vom Klein-Sein zum Einstein" (from being small to being Einstein) (Pauen, 2009). The concept mainly consisted of trainings for preschool teacher teams over the course of a year, including visits to the preschools by supervisors and the establishment of science learning workshops in the preschool institutions. The topics of the trainings were mathematics, chemistry, astronomy, and physics. What makes this program special is that it included a thorough evaluation, with pre- and posttesting and a control group and data from both preschool teachers and children. The results show that the participation in the program improved the educators‘ perceived self-competence in science and had positive effects on children‘s scientific knowledge. In the domains physics and chemistry, children‘s domain-general inquiry skills were also specifically fostered, and the researchers found training effects in the posttest.

Another large study investigating the effects of a science training in preschool is the SNAKE study. Steffensky, Lankes, Carstensen, and Nölke (2012) developed a 3-session training centering on the topic water, with one session on each freezing/melting, boiling/evaporation, and solubility in water. They evaluated it with a sample of 245 children. The training shows positive effects on children‘s scientific knowledge and knowledge about inquiry methods, but only if the training contained both experiments and the application to everyday situations. The authors argued that, while the experiments show the phenomena best so that the children can remember them, the application to everyday contexts is needed for the transfer to and recognition of the phenomena in other contexts.

Additionally, there exist a lot of scientific concepts and materials for preschool that have usually been developed by experts but not evaluated so far. Michalik (2010) differentiates two approaches in early scientific education: didactical models that focus on instruction and experimentation on the one hand, and concepts that put children‘s self-education in the center of attention. The first approach is best represented by the work of Gisela Lück (Lück & Demuth, 1998; Lück, 2004, 2013, 2015). Her concept focuses on experimentation with children, as experimentation does not only activate cognitive dimensions and train the senses but also show positive effects on social and language competencies (Lück, 2015). The materials for teachers and parents usually contain a precise structure, with the presentation of a problem, the instructions for the conduction of the experiment, and an interpretation phase, in which the actual knowledge growth is supposed to happen (Lück, 2015). With the topics, she consciously focuses on inanimate nature and therefore physics and chemistry. The experiments deal with phenomena around air, fire, water, and food products (e.g., eggs, fruits, or tea) and are sorted in degrees of difficulty.

Another approach is presented by Gerd Schäfer (Rosenfelder, 2006; Schäfer, 2001, 2009), who criticizes prestructured experimentation and emphasizes the relevance of children‘s self-regulated and holistic encounter with nature. In his concept ―Lernwerkstatt Natur” (*learning workshop nature*) (Schäfer, 2009), children are supposed to make their own experiences in nature, which will lead not only to the development of scientific or biological understanding, but also further their artistic, motoric, and emotional development. He also highlights the importance of observations with all senses. The suggested activities usually take place outdoors or in a workshop, and while general activities are given, such as collecting, investigating, or making, the specifics are left to the children, who are supposed to further their own theories individually.

When summarizing the learning goals for science education, usually observation is listed as well: Möller and Steffensky (2010) mention it as a relevant scientific method, and in the Bavarian education guidelines, *observing processes and changes in nature* are in the list of educational goals (cf. Staatsinstitut für Frühpädagogik München (IFP), 2012, p. 262). While the above-described interventions and designs normally include observation as an inquiry activity and/or additional learning goal, they do not explicitly train observation. Johnston (2009) specifically investigated children‘s observation and categorization skills and formulates factors for fostering children‘s observation behavior, with a focus on the relevance of social interaction with adults and peers. She identified affects as an important starting point for observations, which will then usually move from being broad to then being more specific. Monteira and Jiménez-Aleixandre (2015) conducted a 5-month project with snails in preschool institutions. The children observed the snails together with their preschool teachers, collecting evidence for their own research questions. The qualitative results of the study show promising effects of the intervention: the children showed purposeful observation behavior and sophisticated dealing with evidence. The authors also concluded that the instructional support by the preschool teacher was crucial for the quality and depth of children‘s investigation.

Kohlhauf (2013) developed game materials for training children‘s observation competency in an everyday context in preschools. The games were developed according to the dimensions of the observation competency model, covering the skills describing, inquiry, and interpreting. They were meant to be easily adapted into everyday activities in kindergarten and were tested by preschool teachers who gave positive feedback on the usability and effects of the materials. It seems promising to follow up on these first findings.

2.4.3. Design and Style of Early Education Programs

Having a second look at the two approaches Michalik (2010) identified for science education in preschool, it becomes clear that they do not only differ in content, but in the general style. While the materials by Lück (2015) do stimulate children‘s self-activity and problem-solving, the procedure and the goal of the experiment is structured and regulated by an adult. In the activities proposed by Schäfer (2009), the learning goal is open to children – for him, it is important that the children experience nature, and they will learn more about it in the process. This debate is common in early childhood education. Both approaches are based on the idea of constructivist learning but differ in the degree of self-organization they expect from learners. The theoretical approach of constructivist learning portrays the learner as an observer, who can only interpret the world based on their own individual experiences, and not –objectively”. This leads to the conclusion that learning is a self-organized process, based on the learner‘s prior knowledge, experiences, interests and motivation. Therefore learning processes have to be self-regulated, and should include authentic contexts that are relevant to the learner (for a detailed summary see Schüßler, 2004). This approach also became popular because research showed that learners often do not show the desired ability to transfer or generalize their knowledge: they often accumulate inert knowledge they do not use for different contexts (Renkl, Mandl, & Gruber, 1996). The idea is that situated learning leads to better transfer abilities. In fact, research has shown that when working with problems in realistic, ill-defined contexts, the ability to transfer or generalize increases (for a review see Hmelo-Silver, 2004). However, in practice, the problem arises of how self-regulated learners are able to work, especially with young children. Self-regulation has high demands on metacognition (e.g., Schraw, Crippen, & Hartley, 2006), which is not as fully developed in preschoolers as it is in older students (e.g., Veenman, Hout-Wolters, & Afflerbach, 2006). Critics of self-regulated learning in preschool point out that the fact that children are generally interested in everything in their environment should not lead to less, but more guidance by adults: otherwise, their learning processes run the risk of being random and superficial, and children with less good developed learning strategies suffer and are left behind (Grell, 2010).

So how should learning in preschool look like? Experts and researchers suggest different factors that are relevant for constructivist learning in preschool. Three important concepts – instructional support, differentiation, and learning through play– shall be described in more detail here.

Fthenakis (2009) bridges the conflict between too little or too much regulation with the concept of co-construction: putting an emphasis on the social aspect of learning processes, he calls for a partnership of children and teachers in the learning process. Basing on the idea of the –zone of proximal development" by Vygotsky (1978), the teacher takes up children's questions and ideas and helps them enlarging their knowledge. Results from developmental studies show that children often have problems with activating knowledge or strategies they already have (e.g., Gentner & Namy, 1999). Here it is the job of the teacher to help them activate their concepts and learning strategies by supporting them with instructions. Gentner and Namy (1999; 2006) investigated how low or high instructional support in constructivist learning environments effects primary school children‘s learning and conceptual change. While both groups did learn about swimming and sinking through the intervention, one year later the children who had received high instructional support outperformed the children from the group of low instructional support. This indicates that though self-activity helps learners, instructional support is necessary to help children structure and embed their knowledge.

Secondly, Leuchter, Saalbach, and Hardy (2010) stress the importance of differentiation in early education. To differentiate here means to consider heterogeneous learning conditions in the lesson planning and instructional design (cf., Scherres, 2013, p. 22). It is known that children enter the classroom with different prior knowledge and concepts and that this will influence their learning process (Jonen, Möller, & Hardy, 2003). Krajewski et al. (2009) found in their intervention study that their materials led to a different performance for different age groups, the younger children profiting more from the easier tasks and the sophisticated tasks specifically enhancing the skills of the older children. They also relate these results back to Vygotsky‘s concept of the ―zone of proximal development" (1978 : only if the task is minimally more difficult than the child‘s current competency level, learning, and conceptual change can take place. Practically, this demands adaptively usable materials with several levels of difficulty and/or depth, so that both teachers can plan differentiation in their classroom and learners can adapt the level themselves while on the task (Leuchter et al., 2010).

The last relevant concept is learning through play-activities. Oerter (2012) emphasizes the importance of incidental learning by play. This does not necessarily mean that the children individually pick their learning goal and form the learning processes on their own, but can be guided by adults with adequate materials, which stimulate specific learning processes. When speaking about playing, one has to differentiate between free play, when children learn completely self-regulated and incidental (e.g., in role-playing activities) and learning games, where the learning goal is intended (Mogel, 2008). Hauser, Vogt, Stebler, and Rechsteiner (2014) tested if preschool children learn mathematical skills better in a structured program with direct instruction, or with learning games that have been specifically developed to foster mathematics. The learning games group showed a significantly higher increase of mathematical skills than the control group, while the direct instruction group did not differ significantly from the control group. This reinforces the idea that playing is a central and effective learning mechanism for preschool children. The games met several important facets of constructivist preschool education: all children were actively engaged by playing the games, and the level of difficulty could be met within the games by means of differentiation.

To sum up, it can be said that there is a general agreement on the fact that learning in preschool should be constructivist, i.e., that children‘s prior experiences are taken into account and that they are active themselves throughout the learning process. However, this does not mean that teachers have no responsibilities and fade into the background: they need to activate children‘s resources, take up their ideas, find the ideal level of task difficulty, and help them focus and structure their learning experience so that children‘s levels of competency are met accordingly. With preschool children, play activities have proven to be an effective way of meeting these requirements.

3. Research Questions and Hypotheses

As stated in the introduction, the research goal of this study is two-fold: for one thing, the aim is to understand scientific observation competency and its development in early childhood by investigating potentially related factors. The other aim is to find out whether and how preschoolers‘ observation competency can be trained. According to these two research questions, there are more specific hypotheses that will be introduced in the following.

Research Question 1: How does children’s observation competency develop, and what is it related to?

Domain-specific knowledge has proved to be crucial for children‘s observation competency (Eberbach & Crowley, 2009; Kohlhauf et al., 2011). For the domain of biology, a non-animistic biology understanding develops in preschool age and has an impact on children‘s view of biological contexts (Carey, 1985; Inagaki & Hatano, 1993). Though effects for domain-specific conceptual knowledge have only been shown for the context of physics before (Penner & Klahr, 1996), children‘s non-animistic biology understanding can be expected to be related to their observation competency. At the same time, there is evidence for a development of a domain-general scientific reasoning skill (Osterhaus et al., 2015; Piekny et al., 2013b, 2013b) that is based on the understanding and mastery of general epistemic activities throughout the inquiry process. Since observation as a scientific method bases on the understanding and correct application of these activities (Fischer et al., 2014; Wellnitz & Mayer, 2011), a relation between domain-general scientific reasoning and observation competency can be expected as well. Therefore, the first hypothesis is:

Hypothesis 1a: Both biology understanding and domain-general scientific reasoning are associated with high observation competency.

For both observation competency and domain-general reasoning, several factors have been discussed to have an effect on the development of these competencies. Language abilities have shown to be related to both (Kohlhauf et al., 2011; Mayer et al., 2014; Piekny et al., 2013b; van der Graaf et al., 2016), while executive functions and theory of mind have so far only been related to scientific reasoning (Mayer et al., 2014; Piekny et al., 2013b; van der Graaf et al., 2016). Prior knowledge in the domain has shown to improve observation competency (Eberbach & Crowley, 2009; Kohlhauf et al., 2011). This leads to the second hypothesis:

Hypothesis 1b: Language abilities, executive functions, and theory of mind are associated with both children's domain-general scientific reasoning and their observation competency, and prior knowledge about the observed objects is related to observation competency.

Though children‘s interest is expected to play a role for the quality of their observations (Johnston, 2009), little attention has been on the role of emotional aspects during observation. In general, it is assumed that emotions influence scientific reasoning (Fischer et al., 2014). Children‘s emotional well-being and involvement are prerequisites for deep level learning and are assumed to have an impact on their behavior in inquiry situations (Laevers, 1993). They are therefore also expected to be crucial for children‘s observation competency. The third hypothesis is:

Hypothesis 1c: Emotional well-being and involvement in the situation are related to children’s observation competency.

All factors expected to be related to children‘s observation competency are summarized in Figure 2. While interrelations between the influencing factors themselves can be expected, they are all separate factors that are related to different aspects of observation competency. Therefore, the last hypothesis related to this research question is:

Hypothesis 1d: All Influencing Factors – Scientific Thinking Factors, Cognitive Factors, and Affective Factors – have a particular relation with Observation Competency, even when looking at them at the same time.

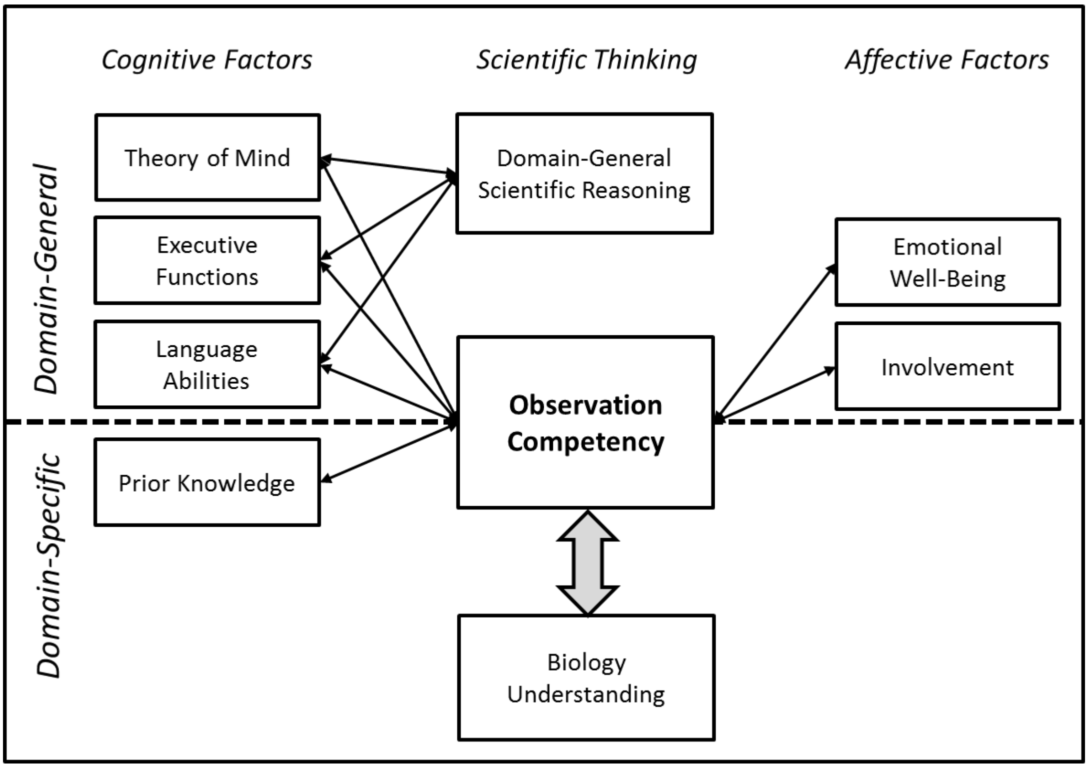


Figure 2. Influencing Factors on Observation Competency (Research Question 1)

Research Question 2: Can children's observation competency be fostered with a structured, game-based training in the last year of kindergarten?

Research has shown that interventions in preschool are particularly effective when they are game-based (Hauser et al., 2014; Mogel, 2008; Oerter, 2012) and give options for differentiation (Jonen et al., 2003; Krajewski et al., 2009; Leuchter et al., 2010). For the training materials used in this study, these factors have been taken into account. They have also been developed according to the observation competency model by Kohlhauf (2013). Consequently, they are expected to be assessed as adequate materials for fostering children‘s observation competency by the trainers. The hypothesis is:

Hypothesis 2a: The developed materials show good usability for preschool children and the specific aim of fostering children's observation competency.

Children‘s scientific reasoning skills develop significantly during preschool age (Piekny et al., 2013a; van der Graaf et al., 2015). Observation competency also shows an increase from childhood to adulthood (Kohlhauf et al., 2011; Kohlhauf, 2013). At the same time, as a competency it should per definition be possible to learn and improve it by training (Hartig & Klieme, 2006; Weinert, 2001). First studies show promising results for teaching preschoolers basic science concepts (Anders et al., 2014; Kohlhauf, 2013; Pauen, 2009). Therefore, both developmental and intervention effects are expected to be found. This leads to two hypotheses:

Hypothesis 2b: All children's observation competency increases from the pretest to the posttest (developmental effect).

Hypothesis 2c: The children of the training group show a higher increase of their observation competency compared to the control group (intervention effect).

4. Method

4.1. Sample

The data for the main study was collected in five preschools. Two of these were in an urban environment, three in a rural area. Three of the preschools were run by municipal authorities, one by the church and one by parent initiative. All preschools had several core groups for the children but also group-overarching activities.

In the pretest, 83 children participated, who were in their last year before starting school. For the analysis of the pretest results, eight children were excluded from the analyses because their language abilities were so low that the testing could not be run with them as it was with the other children. The cutoff for excluding them from the sample was their performance in the language test. If their results fell into the area of –special educational needs”, their performance was not analyzed any further. The age of the final sample of 75 children ranged from 4;9 to 6;7; the mean age was 5;6 (65.56 months, SD = 4.67). 38 (51%) of the children were female, 37 (49%) were male.

Of this sample, 5 children dropped out before the posttest because they moved away, left the kindergarten or were on holiday for the whole period of the posttesting. The final sample for the pre-post-comparison consists of 70 children whose mean age at the pretest was 65.62 months (SD = 4.77); 35 (50%) were female and 35 (50%) male. 40 of these children had participated in the intervention and 30 children were part of the control group.

The children‘s parents or legal guardians had been informed about the study beforehand and had given their consent. The letter can be found in the appendix. They had the possibility to withdraw their consent at any time and ask for the deletion of already recorded data. The children themselves also had the possibility to cancel the testing or participation in the training sessions at any time. Parents also had the opportunity to ask for their own children‘s test results. The letter to the parents with the consent form can be found in the appendix.

4.2. Design and Procedure

4.2.1. Design and Procedure of the Main Study

Data for research question 1 and 2 were collected in one main study. The experimental design consists of an experimental group and a control group. In four of the five kindergartens, complete random assignment of the children to the two groups was possible, in one attention had to be paid to the group membership of the children. While the experimental group participated in the weekly training, the control group did not receive any training.

There was one measurement before the start of the training phase (pretest) and one measurement at the end of it (posttest) with a phase of five months in between. In these five months, the 12 training sessions took place in an almost weekly rhythm, with exceptions due to holidays. At the pretest, we measured children‘s observation competency, domain-general scientific reasoning, biology understanding, theory of mind, executive functions, language abilities, and prior knowledge. Additionally, children‘s emotional well-being and involvement were assessed by a coding of the observation situation. At the posttest, only observation competency, domain-general scientific reasoning, and biology understanding were tested again. The whole design is displayed in Figure 3. The upper part of the figure shows the testings and intervention that took place in the kindergarten, while the lower part shows the analysis of the collected data.

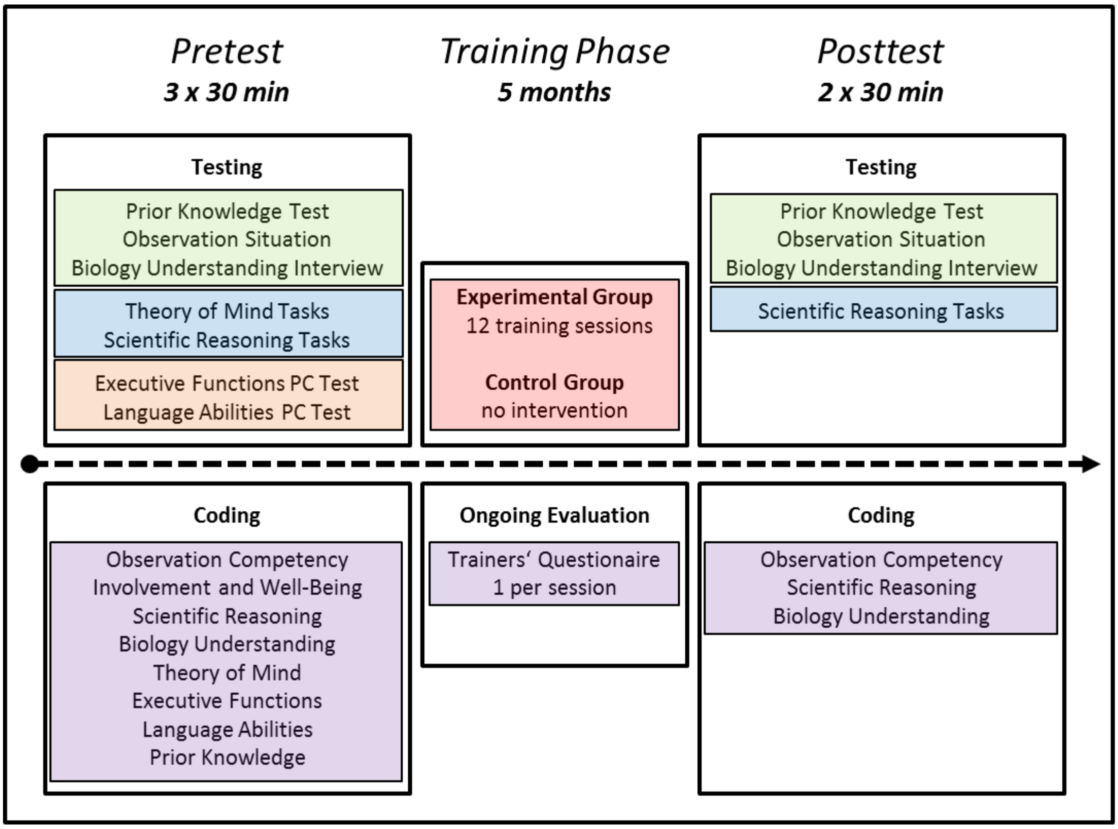


Figure 3. Design of the Study

Table 4. Testing Blocks at the Pretest and Posttest

	Pretest	**Posttest**
First Block	Prior Knowledge Observation Competency Biology Understanding	Prior Knowledge Observation Competency Biology Understanding
Second Block	Domain-General Scientific Reasoning Theory of mind	Domain-General Scientific Reasoning
Third Block	Executive Functions Language Abilities	---

The testing took place in the preschools in a separate room. We tested the children individually in three test blocks that each took up to 30 minutes. Usually, each child was not tested twice on a day; if that did happen, we made sure that they had at least two hours leisure time in between. The testing took either place at a computer or in interview form and was recorded on videotape. If the child did not want to be tested alone, one of the preschool teachers would come along to the testing. Table 4 shows the different test blocks for the pre- and posttesting.

During the intervention phase, the trainers completed questionnaires on the participation of the children in the training; parallel to the posttest, feedback was obtained from the parents of the children in the experimental group.

4.2.2. Pilot Studies

4.2.2.1. Piloting of Scientific Reasoning Test

There were two tasks used in this study for measuring scientific reasoning, the mouse-house-task (Sodian et al., 1991) and the cake-task. Both will be described in chapter 4.4.3. in more detail. While the mouse-house-task has been used several times before, the cake-task was newly developed for this study in order to have more variance when testing children‘s scientific reasoning skills. This task, which was constructed as a parallel to the mouse-house-task, was then piloted on a sample of 51 preschool children (26 girls & 25 boys , ranging from 3;10 to 7 years (mean age: 5;6 years . Children‘s performance on the task correlated with their age ($r = .42$, $p < .01$), as it does for the mouse-house-task. Their answer patterns of their justifications for their decisions were also analyzed, being coded in “no answer”, “Wrong/irrelevant”, “wrong but consistent with prior beliefs”, and “correct”. The results showed that children often relied on their prior beliefs, meaning that when they had a hypothesis about which ingredient makes a difference, they did keep this belief. This result that prior belief effect children‘s scientific reasoning abilities was found before with other scientific reasoning tasks (Croker & Buchanan, 2011; Koerber et al., 2005). Children reported that they liked participating in the task and the experimenters optimized the exact wording of the task after the pilot study.

4.2.2.2. Piloting of Training Materials

While Kohlhauf (2013) had developed the materials in a way so that kindergarten teachers could apply them in the everyday life in kindergartens, the aim was to conduct a more structured and regulated intervention in order to find out if children‘s scientific observation can be fostered. Therefore, we conducted a pilot study in order to find out whether the games and activities can be transferred to a more scheduled format. Lesson plans for four sessions were developed and implemented in a kindergarten with a group of seven children during the course of three weeks. The games and activities of Kohlhauf‘s work were grouped together under an overarching theme for each session, like plants or birds.

The intervention worked well, with positive feedback from the preschool teachers who participated in the sessions. However, several critical points could be identified: firstly, children often varied in how much time they needed for a specific task. While this does not matter too much in an everyday scenario in kindergarten when children can just go on with another activity, in the structured group sessions of this study it would be important to always have backup tasks in order to enable differentiated instruction. Secondly, it became clear that children profited more from some of the games when they already knew the rules beforehand and could, therefore, concentrate on the content.

In order to help children gain the most from the games, these games were integrated into more than one session of the final training. Thirdly, some important conclusions about organizational structures in German kindergartens could be drawn. With their individual timetables, curricular activities and celebrations in the annual calendar, it became obvious that it is necessary to make individual agreements with each group in order to be able to conduct all training sessions as planned. Finally, the pilot was important for gathering some practical experiences with the activities and their difficulty level, which were mainly in line with the results from Kohlhauf (2013).

All of these results flowed into the development of the new materials and the planning of the training sessions for the intervention.

4.3. Intervention

4.3.1. Process of the Intervention

A total of twelve training sessions were held. Each of the meetings was scheduled for up to 90 minutes, but could usually be kept shorter.

The training took place about once a week – with intermissions due to school holidays – over the course of five months. The sessions were conducted in a separate room in the kindergarten by the same trainer. Each child collected their drawings, materials and worksheets in a researcher booklet they could keep in the end as a keepsake.

In total, there were four different trainers, who had all contributed to the development of the materials and the exact articulation of the learning goals. The trainers also met once a month to discuss both the progression and possible problems in the previous sessions as well as the detailed planning for the upcoming sessions. For each session, they developed lesson plans that were structured according to the articulation of a school lesson with the phases introduction, elaboration, backup, and closing. There were specific learning goals for each training session as well as material lists. Examples of the lesson plans can be found in the appendix.

4.3.2. Materials for the Intervention

In total, 46 different games and activities were used in the training sessions. These materials were developed by the trainers or taken over from the previous studies of Kohlhauf (2013), some of which were slightly modified or optimized in their application.

The games and activities were based on the empirical model for biological observation skills by Kohlhauf (2013, p. 91). As the idea of the intervention is to foster children‘s observation competency in the three dimensions *describing details*, *scientific inquiry,* and *interpreting*, there were games for each of these dimensions, fostering specifically its enrollment. Depending on its demands, each activity can be classified either as fostering unsystematic or systematic observation.

Table 5 displays how the games and activities can be assigned to the dimensions and difficulty levels. It should be noted that several games refer to more than one dimension. If they could train observation competencies on both difficulty levels, they are sorted into the higher category.

Table 5. Activities and Games Sorted According to Dimension and Difficulty Level

	Describing details	Scientific Inquiry	Interpreting
Systematic Observation	Drawing the bean Comparing soils Animal Pantomime Memory of smells Flower Pictures Forest Camera Search list Presenting Findings Bird Quartet Bird Puzzles Systematizing Feathers Soil animal Dice Game Detail Pictures Soil animals Originals Soil Animal Pantomime Tactile box Tactile Parcours Observing Hands Finger Prints Nature Findings	Planting the bean Drawing the bean Observing the bean Observing the tree Flight characteristics of feathers Differentiating surfaces	Animal Pantomime Dice with facial expressions Memory of smells Sorting birdsong Soil animal Dice Game Detail Pictures Soil Animal Pantomime Tactile box Tactile Parcours Observing Hands Animal Prints Animal gaits Complete the animal Looking for prints Finger Prints Guessing the fruit Sorting fruit and juice
Unsystematic Observation	Error pictures Bird Memory Bird Puzzles Complete Bird Drawings Describing Birdsong Imitating flying Tracing feathers Modeling soil animals Differentiating surfaces Shoe salad Human footprints Animal tracks Animal gaits Complete the animal Guessing the fruit	Imitating flying Observing the woodlouse	Woodlouse Story

4.3.3. Themes in the Intervention

Thematically, the training sessions were divided into five blocks with different biological topics.

The first block "Bean" focused on planting and observing the growth process of a bean plant. In addition, the sessions included several short games that also focused on plants and necessary aspects for their growth. The topic was carried out in 3 training sessions.

In the second block "tree", the children engaged in the observation of trees, its surroundings, and its products. While the first session concentrated on trees and its surroundings in nature, with mainly visual observation outdoors, the second training session put fruit into the focus and included observation by smelling and tasting.

The third block "bird" dealt with the characteristics of birdsong and feathers of various, mostly local birds. The sessions included auditory observation of birdsong and its imitation, several games with photographs of various birds, finishing drawings of birds using a photograph as a model and investigating the texture and functions of different bird feathers. This topic also lasted for two sessions.

In the fourth block "soil life~~,~~the children engaged in the observation of the appearance and behavior of small soil animals, such as insects, beetles, worms, and spiders.

The two training sessions dealt with the observation of woodlice and their behavior, observing detailed characteristics of different soil animals on enlarged photos, a dice game with a focus on combinations of specific details of soil animals, modeling insects using play dough and toothpicks and more short games that also promote observing.

The fifth and final block "hands and feet" enabled the intensive involvement of the children with their own hands and feet, and its comparison to other living beings. The first session centered on tactile perception with a tactile box, a tactile Parcours and crafting rubbings of various surfaces. The other two sessions concentrated on footprints and animal tracks, with games with children‘s own slippers and feet, human footprints, animal tracks and fingerprints.

4.4. Instruments

In the following, I will introduce the test instruments used in this study. The German testing materials for all instruments can be found in Appendix C, except for the language test and the theory of mind scale. For these, the exact versions of the cited instruments were used.

4.4.1. Observation Competency

For testing children‘s observation competency, we used Kohlhauf‘s competency test (Kohlhauf et al., 2011). Three animals (fish, snails, and mice) were presented to the children in turn. As in the original study, a hand puppet named "Emil" was used in order to encourage the children to participate more and facilitate communication between the experimenter and the child at eye level (Kohlhauf et al., 2011). To find answers to their own research questions, subjects were allowed to use a stopwatch, a ruler, a scale, a magnifying glass and a thermometer as tools. Those instruments were introduced to the children in the warm-up phase.

In the experimentation phase, the hand puppet closed his eyes and asked the children to describe the animal to him. After that, the children could observe the three test animals freely. If they did not come up with research questions, hypotheses and testing ideas themselves, they were prompted or helped to do so by the experimenter. If they still showed no observation behavior, the experimenter and the hand puppet would conduct the step for the children. In the interpretation, the children could again first act freely before the experimenter would prompt them to relate the results back to the hypothesis. The whole interaction was videotaped and, later on, coded for children‘s observation competency.

We first coded children‘s behavior in the same way Kohlhauf et al. (2011) did. In their analysis, there were 15 items: per animal one each for details, questioning, hypothesizing, testing and interpreting. In each case, the behavior was directly coded to be on level 0, 1 or 2. This was sufficient for their sample with an age range from preschoolers to students but proved to be too imprecise for this sample. The children showed bottom effects and it was not possible to reach satisfactory interrater reliability. We, therefore, developed a more specified coding scheme with both more items and more gradations. The final coding scheme consisted of 39 items: 13 per animal, and per animal three on details, two on questioning, two on hypothesizing, two on testing, and four on interpreting. Depending on the items, there were three to four different grades to code children‘s behavior. The list of items with the gradation can be found in Table 6; the German version of the coding scheme is attached in the appendix.

A second rater coded 10% of the data and the Spearman correlations were all above .6, for the subscales inquiry and interpreting they were all above .9.

While the overall scale was reliable ($\alpha = .74$) and the subscale for details was also reliable ($\alpha = .72$), the values of the subscales inquiry ($\alpha = .63$) and interpretation ($\alpha = .40$) were not sufficient. When treating inquiry and interpretation as one scale, satisfactory reliability ($\alpha = .76$) was reached again.

4.4.2. Involvement and Emotional Well-Being

The videos from the observation situation at the pretest were also used to code children‘s involvement and emotional well-being (Laevers & Heylen, 2003) in that situation. The scoring was done separately for each animal the child observed. The adaptation of the observation sheet looked at the signals of involvement and well-being (Laevers et al., 1997) with 9 items for involvement (concentration, energy, creativity, facial expression, persistence, precision, reaction, verbal utterances, satisfaction) and 8 items for emotional well-being (openness, flexibility, self-confidence, assertiveness, vitality, inner peace, enjoyment, feeling at ease). A description of each item can be found in Table 7, the German coding sheet can be found in the appendix. For each item, a score of 1 (no signs), 2 (some signs), 3 (clear signs) or 0 (missing) was given.

The coding was done by a different coder than the coding for the observation competency and the coder prepared for the coding with the materials and DVD from Laevers et al. (1997).

Table 6. Items and Gradation of the Observation Competency Coding

Dimension		Item	Gradation
Describing Details		Dimensions	Number of mentioned dimensions
		Unspecific Details	Number of mentioned details that are not related to the question
		Specific Details	Number of mentioned details that are related to the question
Inquiry	Questioning	Research Question	• Spontaneous • Prompted • With help • No question
		Use of Question	• Child‘s question was used • Emil‘s question was used
	Hypothesizing	Spontaneous Hypothesis	• Spontaneous • Not Spontaneous
		Prompting	• Prompted • With help • No hypothesis
	Testing	Activity	• Autonomously • Help with either idea or implementation • Idea and implementation by experimenter
		Quality	• Real observation • No observation, confirmation bias
Interpreting		Summary of results	• Autonomously • Prompted • None/wrong
		Spontaneous relation to hypothesis	• Spontaneous • Not Spontaneous
		Prompted relation to hypothesis	• Correctly when prompted • None/wrong
		Differentiation between observation and inferences	• Present • Not present

Table 7. Description of Coding Items for Involvement and Emotional Well-Being

Scale	Item	Description
Involvement	concentration	turning body to animal or experimenter, no digressing, full concentration on the object
	energy	Happy on the task, energy related to the task
	creativity	Introducing new, own ideas; if low: child just following instructions from the experimenter
	facial expression	Attention to the object; relaxed but also excitement in facial expression
	persistence	Not distracted, fully concentrated permanently
	precision	Precise describing; precise, meticulous work with the tools; precise observing
	reaction	Following the instructions of the experimenter; good use of given prompts
	verbal utterances	Fluency in language, coherent phrases
	satisfaction	Fascinated facial expression; positive exclamations; silent satisfaction
Emotional Well-Being	openness	Responding to experimenter; telling private stories; open attitude
	flexibility	Responding to task; use of tools; linguistic competence
	self-confidence	Natural behavior, freely interacting with objects and materials
	assertiveness	Voicing own opinions and ideas, knowing what one wants
	vitality	Emitting vitality; reasonable desire to move
	inner peace	Seeming relaxed, if low: signs of nervousness
	enjoyment	Laughing / smiling; positive exclamations; satisfied facial expression
	feeling at ease	Relaxed posture; relaxed facial expression

4.4.3. Scientific Reasoning

We used two tasks to measure children‘s scientific reasoning abilities: the mouse task by Sodian et al. (1991) and the cake task, which was developed in parallel to the mouse task. Both tasks were told to the children in the form of a story, supported with pictures. Children could point at the pictures to answer but also had to verbally justify their answers. If the justification showed a wrong concept or no justification was given, the answer was coded as wrong. For the mouse task, there were control questions on children‘s understanding of the task. If the children answered these wrong, their data was coded as missing.

The mouse task: In this task, the children were told the story of two boys who have a mouse in their cellar. The boys had never seen the mouse and therefore did not know if it was big or small. In the first step, they wanted to feed the mouse and had to choose one of two houses (one with a small entrance, one with a big entrance) to put cheese for the mouse in. In the second step, they wanted to find out if the mouse is big or small and again chose one of the two houses to put cheese in. In a third step, the big house was shown, saying the cheese is missing and asking the children if they now know if it is a big or a small mouse.

The cake task: In this task, a mother baked a cake with two new ingredients and her 3 children liked the cake a lot. In the first step, the mother wanted to bake the cake again for a birthday party and the children made suggestions what she should do. Child A suggested to put only one of the ingredients into the new cake, child B suggested to put both ingredients into the cake (right answer), and child C suggested to bake a cake in a square form instead of a round one. In the second step, the mother wanted to find out which of the ingredients is the one to make the cake so tasty because the ingredients were rather expensive and she only wanted to have to buy one. Child A suggested to bake one cake with both ingredients and one cake without both ingredients, Child B suggested to bake one cake with the first and one cake with the second ingredient (right answer), and child C suggested to bake one round and one square formed cake. In the third step, the family had decided to try out Child A's suggestion and the test instructor asked the children if they now found out which ingredient makes the cake tasty.

As it was only relevant to analyze children's understanding of testing and not that of producing an effect, children's answers on the first question (producing an effect) were not considered. Therefore, there were answers to two questions per task, one on the selection of the right answer and one on the additional posthoc question. Thus, children could score 0, 1 or 2 points on both scientific reasoning tasks. Table 8 shows children's frequency scores in the two tasks.

Table 8. Frequencies of Scores in the Scientific Reasoning Tasks

	Cake Task	Mouse Task
0	31	36
1	16	18
2	19	12

Children's performance on the two tasks was significantly correlated ($\tau = .38$, $p < .01$), even after language and age had been partialed out ($r = .31$, $p < .05$). Because of these correlations, the two scores were aggregated to a single scientific reasoning score.

4.4.4. Biology Understanding

For testing children‘s biology understanding, an animism interview was conducted to test children‘s concept of what it means to be alive. The first version of this kind of interview was developed by Piaget (1978) and has been adapted since (e.g., Carey, 1985). We used the version of the interview Zaitchik et al. (2014) used in their study. The interview started with several open-ended questions: “What does it mean to be alive?”, “Can you name some things that are alive?” and “Can you name some things that are not alive?” Next, the experimenter asked the child to judge a list of things on whether they are alive or not. This list included the categories animals (e.g., a cat), plants (e.g., a tree), natural kinds (e.g., fire) and artifacts (e.g., a lamp). For one item of each of these four categories, children were asked for a justification of their judgment (e.g., “Why do you think a tree is alive / not alive?” .

This interview was always done directly after the observation situation and also videotaped. Children‘s answers were later transcribed and then coded according to the scheme Zaitchik et al. (2014) developed, which is also published online as an additional attachment to their paper. Children could score up to 19 points. The reliability analysis reached $\alpha = .56$ in the pretest and $\alpha = .63$ in the posttest. Children had significantly improved from pre- to posttest ($M_{pre} = 5.01$; $M_{post} = 5.99$; $T = -2.60$; $p < .05$).

4.4.5. Theory of Mind

For the theory of mind, we chose the two hardest tasks from the German adaptation theory of mind scale (Kristen, Thoermer, Hofer, Aschersleben, & Sodian, 2006), as the children are already rather old and we could expect ceiling effects for the easier tasks. The two tasks were the content false belief task (“Smarties-Task” and the real-apparent emotion task (“Emotions-Task” .

In the –Smarties-Task”, the experimenter showed the child a smarties role and asked what they believe to be in it, the desired response being "smarties". Once the child had answered with "smarties", the box was opened and revealed that there was a toy pig inside. After the experimenter showed amazement that it was a pig, the box was closed again and the child was asked a memory question: "What is in the box?”. If the child did not answer correctly, the content was shown again, until the right answer was given. Then the experimenter introduced a previously not present toy figure, saying: "This is Lukas. Lukas has never seen what is in this box.” Here followed the test question: –What does Lukas think is in the box?" When the children had answered to this question, they were asked the memory question whether Luke already had looked in the box.

The situation was videotaped and children‘s answers later transcribed and coded. If children answered with –smarties” to the test question, they passed this test. If they answered –pig”, they did not. If they gave another answer or answered wrongly on one of the memory questions, their data was coded as missing.

In the second task, the –Emotions-Task”, the experimenter begins with introducing three smiley faces: a sad, a neutral and a cheerful face. The experimenter explains that they will now tell a story of a boy who in the course of the story, could be sad, happy or "in between". Before the story starts, it is checked whether the child could assign the right emotions to the expressions. It was pointed out that at the end the child will be asked how the boy really feels inside (the experimenter illustrates this by pressing their hands to the chest), and which facial expression he shows (the experimenter points to their cheek), explaining that these two states could be the same or different. Next, the story of Tim is told: Tim‘s aunt had come back from vacation. She had previously promised to bring him a toy car. Instead, she now brought a book as a gift. However, Tim did not like books. To avoid that the aunt would never bring him gift again, he did not want to show her how he really feels. After that follow two memory-control questions: "What did the aunt bring Tim?" and" What would Tim‘s aunt do if he would show her how he really feels? ". If the child could not answer correctly, the story was told one more time. The testing questions were: "How does Tim really feels when the aunt gives him the book?" and ―Which face does Tim make?" and the children could point at the smileys for answering both questions.

Children‘s answers were later transcribed and coded. If children picked a sadder smiley for the question about Tim‘s inner feelings than for the face he made, they passed this test. If they picked the same smiley or a happier one for the outside face, they did not. If they did not pass the memory questions, even after they had heard the story a second time, their data was coded as missing.

4.4.6. Executive Functions

For measuring executive functions, children completed the Heart and Flowers task (H&F), which has been developed by Adele Diamond (2013). In this computer-based task, children‘s inhibition, set shifting, and memory is tested.

In the testing situation, children sat down in front of a laptop with an external keyboard. There were two keys marked on the keyboard, which the children had to press in specific events. The children were supposed to use only these two keys, while the experimenter used the laptop keyboard for the navigation in the program.

After an introduction of the computer and the keyboard, the rules of the game were explained to the children. In the first round, the congruent condition or heart condition, children had to press the key on the same side as the heart that appeared. After a training round, children had to press the keys on 20 test items. In the second round, the incongruent condition or the flower condition, children had to press the key on the opposite side of the flower that appeared on the screen. They had to react to 20 test items after some training items. In the final round, the mixed condition, either a heart or a flower could appear on the screen. In the case of a heart, the child had to press the key on the same side, in the case of a flower they had to press the key on the opposite side. There was no training round before the 33 items had to be reacted to.

The program directly exported children‘s results into an excel file. It did not only recorded children‘s responses, but also their reaction times. However, as response times have proved not to be a reliable factor when measuring children‘s executive functions (Diamond, Barnett, Thomas, & Munro, 2007), only the response patterns were further considered.

Children‘s responses showed acceptable to good reliability in all three conditions. However, the solving rate shows that children are almost at the ceiling in the first condition (see Table 9). Furthermore, only in the mixed condition, all three aspects of executive functions are tested, as children have to keep both rules in mind (working memory), shift between the rules and inhibit the tendency to press the key on the same side in incongruent items. Consequently, we only used this condition for the analyses, as others have also done before (e.g., Zaitchik et al., 2014).

Table 9. Solving Rates and Reliability of the Executive Functions Test

Condition	Solving Rate	Reliability (Cronbach‘s alpha
Congruent	91%	.67
Incongruent	79%	.80
Mixed	69%	.85

4.4.7. Language Abilities

We assessed children‘s German language abilities with the computer-based test CITO, Version 3 (Duindam, Konak, & Kamphuis, 2010). Based on a language test from the Netherlands (Van Els & Van Hest, 1992), this test is constructed to assess the language abilities of children prior to starting school. It has been used as an official tool for judging children‘s language level before starting school in several German states. The test can be used with children between 4;3 and 6;11 years.

Children do that test on a computer using a mouse. The mouse is introduced in the beginning in a short training phase. In the testing tasks, children have to click on one of two or three displayed pictures. A clown guides the children through the test, explaining the tasks and motivating them in between. The testing time is about 25 minutes.

Four components of language abilities are measured: The component *passive vocabulary* tested children‘s recognition of nouns of verbs with 45 items. In *cognitive terms*, children had to show their understanding of the concepts of colors, forms, sizes, numbers, positions and relations of objects. This component comprises 46 items. In *phonological awareness,* children heard two words and have to decide whether they heard the same word twice or two different words. There were 20 word pairs presented to them. The component *text comprehension* tested children‘s understanding and memory of short stories (4-5 sentences). Overall, they answered 20 questions in this component.

According to the norms from the validation study (Duindam et al., 2010), children‘s results can be rated as “good”, “no special educational needs”, and “special educational needs”. Using this categorization, children were excluded from the sample who fell into the category of “special educational needs”. The computer program also summarized children‘s results in an explanatory sheet, which were offered to parents if they were interested.

For investigating the relationship of language to other cognitive measures, we only used the subscale *passive vocabulary* because vocabulary tests have shown to be a good predictor of verbal IQ without overlapping with children‘s executive functions measures (Zaitchik et al., 2014). This subscale showed a good reliability of $\alpha = .89$ for the sample ($\alpha = .91$ in the norming sample by Duindam et al. (2010)).

4.4.8. Prior Knowledge

As we were also measuring biology understanding in this study, it was the aim for the prior knowledge test to only measure children‘s factual knowledge on the animals used in the observation situation. Therefore, we conducted the same test as Kohlhauf et al. (2011) in their study. The questionnaire consists of 18 questions about the three animals that are part of the observation situation. The children answered these questions verbally. Their answers were written down by the experimenter and recorded on video. The prior knowledge test always took place directly before the observation situation so that children did not yet have new information gathered during the observation.

We found floor effects on several items, as the questionnaire originally had been developed to be used on a sample with a much broader age range. After the deletion of some items, the final scale had 10 items and reached a satisfactory reliability (α = .58). The final items are listed in Table 10.

Table 10. Items of Prior Knowledge Test

Question
How do fish breathe?
Which function has a swim bladder for fish?
How are the small plates called that cover the body of most fish?
How many tentacles do snails have?
Where exactly are the eyes of a snail?
How do snails breathe?
Which material is the house of the snail made of?
How do snails chop their food?
What do newborn mice feed off?
What helps mice not to run against objects in the dark?

4.4.9. Evaluation Instrument for Trainers

Right after each training session, trainers filled an evaluation instrument to evaluate the materials and monitor children‘s participation. The evaluation instrument consists of two parts. The first part is used for the evaluation of the training session and the specific games and activities. On the basis of logbook entries, the application and implementation of activities and games was documented in the three categories ―That worked well”, ―That did not work well”, and ―Things I changed”. Trainers had to assign the activities to a category and then add more detailed descriptions and explanations.

The second part is used for the evaluation of children. It consists of three items that trainers rated on a 5-point-Likert-scale. The items were answered for each child after every session:

- Item 1: ... has actively participated in the program (activity)
- Item 2: ... had difficulty to follow the contents (difficulty)
- Item 3: ... showed observing behavior today (observing behavior)

4.5. Data Analysis

For most analysis, the software SPSS (IBM SPSS Statistics 23) was used. Correlations, regressions, and ANOVAs were conducted.

In order to understand the relations between some of the influencing factors better, mediation analyses were conducted. A mediation analysis makes it possible to estimate the extent to which variable X influences outcome Y through one (or more) mediator variables. For this, the program PROCESS (Hayes, 2012) was used, which functions as a free plugin in SPSS and integrates all the necessary steps of a mediation analysis. Additionally, that program puts out not only the results of the Sobel test but also a confidence interval for the indirect effect of X on Y in the mediation.

In the pre-post-comparison of observation competency and biology understanding, we wanted to check for retest effects. Therefore, we conducted a Rasch DIF (Differential Item Functioning) Analysis with the scales. The idea of the evaluation of differential item functioning (DIF) is to make sure that a test instrument works the same for two different samples or, as in the case of this study, for one sample at two time points. The basic principle of Rasch Analysis to have estimates for both person abilities and item difficulties makes this possible. In the case of DIF, the person estimates are anchored and the item difficulties for each group are estimated. These different item difficulties can then be compared. Finding differences in the item difficulties in pre- and posttest would mean that for some items the difficulty compared to the test overall has changed. In that case, such items can be handled as different items in pretest and posttest (for more information see Boone, Staver, & Yale, 2014). The Rasch Analyses were conducted with WINSTEPS (Linacre, 2006).

5. Results

The first research question deals with the relation between children‘s observation competency and their scientific thinking, additional cognitive factors, and affective factors. Those will first be analyzed separately and then in a joined analysis.

The second research question centers on the intervention. After the presentation of the qualitative analysis of the materials, I will present the results of the pre-post-comparison of children‘s observation competency and biology understanding.

5.1. The Relation Between Observation Competency and Cognitive Factors

First, the relation of observation competency with scientific thinking and the additional cognitive factors will be analyzed. After the presentation of the descriptive results, I will report the results of correlational and regression analyses.

5.1.1. Descriptive Results

In Table 11, the means and standard deviations for the measurements for children‘s cognitive skills are displayed.

Table 11. Means and Standard Deviation of Cognitive Measures

Measure	*M*	*SD*
Observation Competency	15.19	4.13
Domain-General Scientific Reasoning	.37	.33
Biology Understanding	5.47	2.74
Theory of Mind	1.11	.75
Executive Functions	.70	.19
Language Abilities (Vocabulary)	.93	.08
Prior Knowledge	.13	.15

N = 75

5.1.2. Relations of Children's Cognitive Skills with Observation Competency

The correlation analysis in Table 12 shows the intercorrelations between the cognitive measures. We also checked for correlations with age, but only found one with prior knowledge. Domain-general scientific reasoning, biology understanding, language abilities, and prior knowledge are significantly correlated with observation competency. However, most constructs are moderately correlated to children's language abilities. In order to control for the shared influence of language abilities, a partial correlation analysis was conducted, controlling for the language abilities (see Table 13). Though the correlations decrease slightly, all three constructs – domain-general scientific reasoning, biology understanding, and prior knowledge – are still significantly correlated with observation competency.

Table 12. Correlations of Cognitive Measures

Variable	1	2	3	4	5	6	7
1 Observation Competency	1						
2 Domain-General Scientific Reasoning	.51**	1					
3 Biology Understanding	.56**	.32**	1				
4 Theory of Mind	.22	.32**	.27*	1			
5 Executive Functions	.16	.32**	.25*	.29*	1		
6 Language Abilities (Vocabulary)	.38**	.32**	.37**	.30**	.43**	1	
7 Prior Knowledge	.44**	.42**	.36**	.14	.17	.32**	1
8 Age	.19	.21	.21	.02	-.06	.19	.31**

Note: *p < .05, **p < .01; 70 < N < 75

Table 13. Partial Correlations of Cognitive Measures, controlling for Language Abilities (Vocabulary)

Variable	1	2	3	4	5	6
1 Observation Competency	1					
2 Domain-General Scientific Reasoning	.32*	1				
3 Biology Understanding	.43**	.39**	1			
4 Theory of Mind	.18	.33**	.15	1		
5 Executive Functions	.18	.37*	.35*	.31*	1	
6 Prior Knowledge	.41**	.33**	.29*	.08	.11	1
7 Age	.33*	.14	.16	-.03	-.16	.26*

Note: *p < .05, **p < .01; 70 < N < 75

The first hypothesis was that both domain-general scientific reasoning and domain-specific biology understanding are related to observation competency. In order to test this hypothesis, the next step was a regression analysis with observation competency as the dependent variable and domain-general scientific reasoning, biology understanding, and language abilities as predictors.

The regression explains 40% of the variance in the data ($R^2 = .40$, $F = 15.38$, $p < .01$). Biology Understanding is the largest influencing factor, followed by domain-general scientific reasoning. Language abilities prove to be not significantly influential. The results are fully displayed in Table 14.

Table 14. Regression Analysis Summary for Predicting Children's Observation Competency With Scientific Thinking and Language

Variable	*B*	*SE B*	*β*	*t*	*p*
Language Abilities	5.06	5.39	.10	.93	.35
Biology Understanding	.58	.16	.39	3.64	<.01
Scientific Reasoning	3.50	1.43	.28	2.46	.02

N = 75

Language abilities are also correlated with scientific reasoning and Biology Understanding (see Table 12), so an obvious assumption is that scientific reasoning and biology understanding mediate the influence of language abilities on observation competency. The mediation analysis with PROCESS (Hayes, 2012) found an indirect effect of language abilities on observation competency (b = .31, 95% CI [0.20, 0.44]). The Sobel Test for the mediation was significant for both scientific reasoning ($p < .05$) and biology understanding ($p < .01$). An overview of the mediation model is displayed in Figure 4.

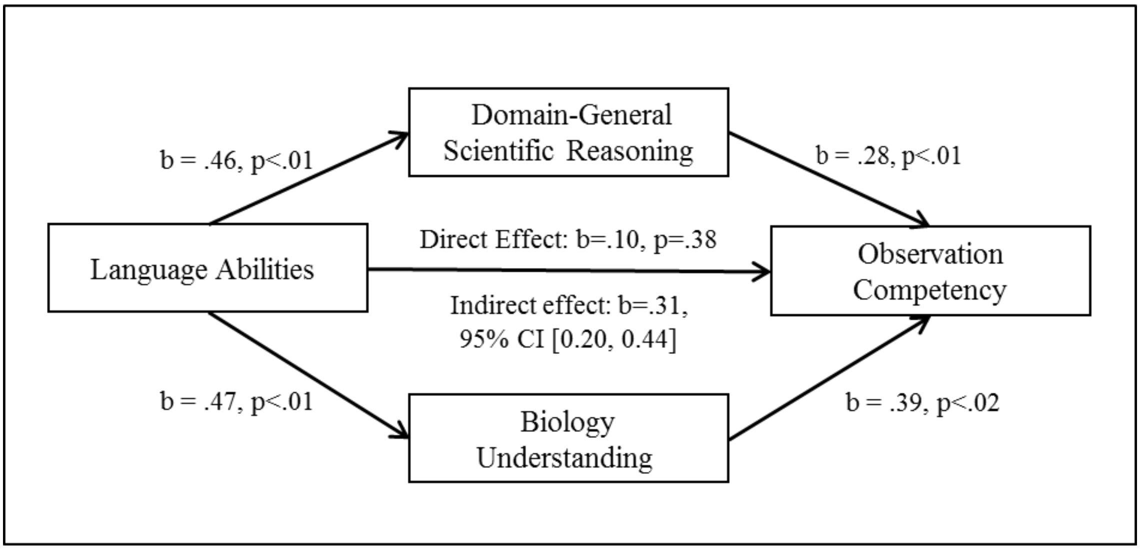


Figure 4. Mediation Analysis for the Influence on Language Abilities on Observation Competency Through Scientific Reasoning and Biology Understanding

The second hypothesis deals with the relation to the additional cognitive factors. Only language abilities and prior knowledge correlated with observation competency. When looking at them in a separate regression analysis, 26 % of the variance is explained and both factors are significant predictors of observation competency (see Table 15). A combined analysis of scientific thinking and additional cognitive factors will follow in the summary of all related factors below.

Table 15. Regression Analysis Summary for Predicting Children's Observation Competency With Cognitive Factors

Variable	*B*	*SE B*	*β*	*t*	*p*
Language Abilities	13.54	5.39	.27	2.51	.01
Prior Knowledge	10.00	2.99	.36	3.34	<.01

We did not find correlations between observation competency and executive functions or theory of mind. However, these two factors correlate highly with domain-general scientific reasoning, even when controlling for language abilities (see Table 13). In order to better understand this relation, another regression analysis was conducted with domain-general scientific reasoning as the dependent variable and executive functions, theory of mind and language abilities as predictors.

The regression model is significant ($F = 14.82$, $p < .001$) with an explained variance of 39% ($R^2 = .39$). Language abilities, executive functions, and theory of mind are all significant predictors for children‘s domain-general scientific reasoning skills, with executive functions being the largest factor. The results are fully displayed in Table 16.

Table 16. Regression Analysis Summary for Predicting Children’s Domain-General Scientific Reasoning With Cognitive Factors

Variable	*B*	*SE B*	*β*	*t*	*p*
Theory of Mind	.11	.04	.25	2.49	.02
Executive Functions	.54	.18	.31	2.95	<.01
Language Abilities	1.08	.42	.27	2.55	.01

N = 75

5.2. Relations Between Observation Competency and Affective Factors

5.2.1. Involvement and Emotional Well-Being

Table 18shows the descriptive results for the items of the involvement and well-being scale and Figure 5 displays the frequency distribution of the two measurements. It can be seen that children, in general, scored high on both instruments, especially on well-being with the means of the items always above 2.5 (on a scale from 1 to 3) and also the values of the overall measure above 2 for all children. The scores for involvement are scattered more broadly with item means between 1.75 and 2.77 and children‘s overall scores between 1.5 and 3.

Table 17. Means and Standard Deviations for Involvement and Well-Being Items

Scale	Item	*M*	*SD*
Well-Being	Openness	2.93	.22
	Flexibility	2.51	.50
	Self-confidence	2.84	.32
	Assertiveness	2.53	.64
	Vitality	2.77	.38
	Inner peace	2.67	.41
	enjoyment	2.99	.06
	Feeling at ease	2.97	.12
Well-Being Overall		2.78	.21
Involvement	Concentration	2.69	.35
	Energy	2.42	.42
	Creativity	1.75	.67
	Facial expression	2.67	.35
	Persistence	2.56	.40
	precision	2.28	.54
	Reaction	2.62	.41
	Verbal utterances	2.55	.55
	satisfaction	2.77	.31
Involvement Overall		2.43	.35

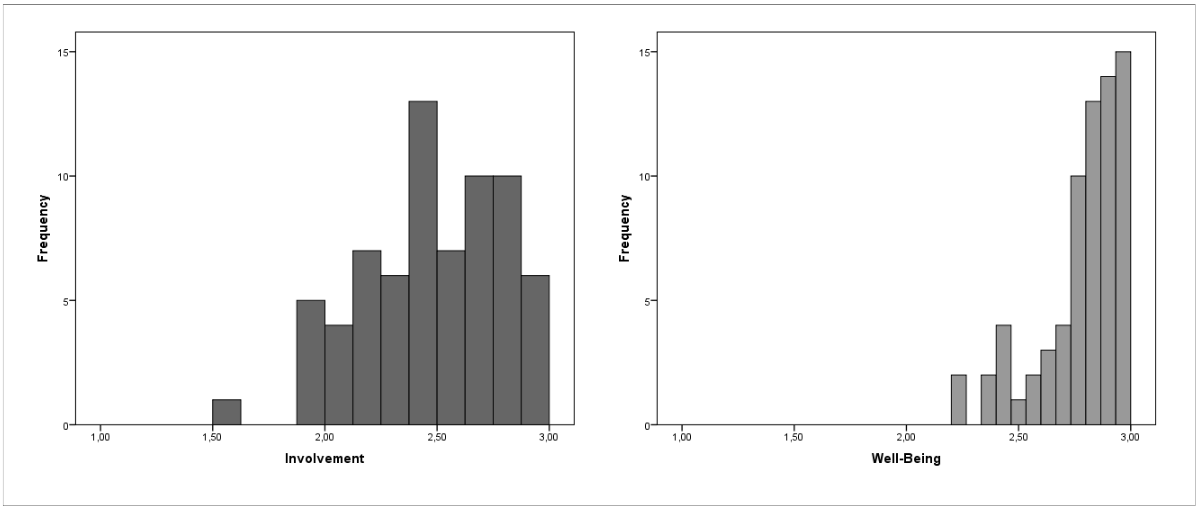


Figure 5. Frequency Distribution of Involvement and Well-Being

5.2.2. Relations of Involvement and Well-Being with Observation Competency

Table 18 shows the correlations of involvement and well-being with observation competency. With values of .58 and .59, these can be interpreted as large. To investigate the relations between the affective measurements and children‘s observation competency further, we conducted a regression analysis. Both predictors were significant (see Table 19) and together explained 46% of the variance ($R^2 = .46$, $p < .001$).

Table 18. Intercorrelations for Observation Competency, Involvement, and Wellbeing

Variable	Observation Competency	Involvement	Well-being
1. Observation Competency	-		
2. Involvement	.58**	-	
3. Wellbeing	.59**	.46**	-

N = 70

Table 19. Regression Analysis Summary for Involvement and Well-Being Predicting Observation Competency

Variable	*B*	*SE B*	*ß*	*t*	*p*
(Constant)	-21.74	5.41		-4.02	.00
Involvement	5.01	1.30	.39	3.85	.00
Wellbeing	8.73	2.16	.41	4.04	.00

N = 70

We then conducted a mediation analysis to see if involvement mediates the effect of emotional well-being on observation competency. There was a significant indirect effect of emotional well-being on observation competency through involvement ($b = 0.18$, BCa CI [0.08, 0.30]). The Sobel test was significant ($p < .01$). Figure 6 shows an overview of the results of the mediation analysis.

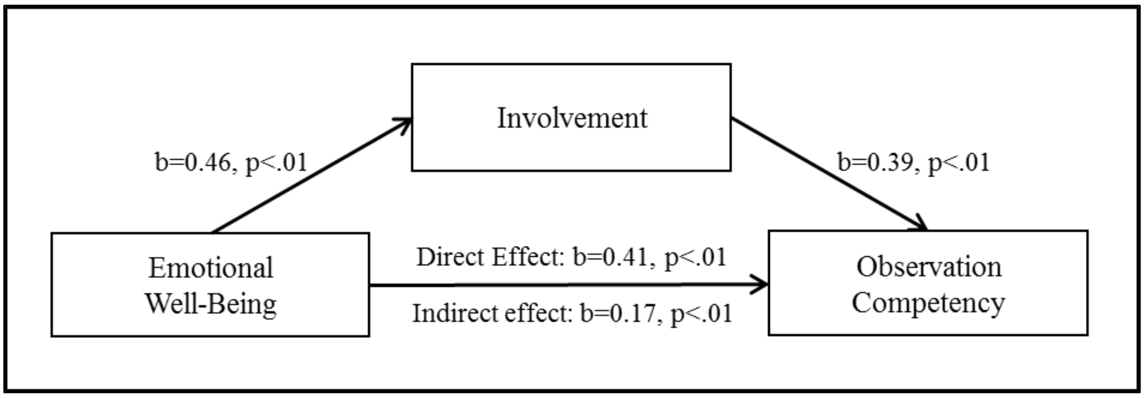


Figure 6. Mediation Analysis of the Influence of Well-Being on Observation Through Involvement

5.3. Summary of all Related Factors

In order to understand the relevance of all investigated factors, a hierarchical regression analysis was conducted. The factors were included in the order of their theoretical closeness to observation competency. In model 1, only scientific thinking competencies, domain-general scientific reasoning and biology understanding, were included. In model 2, the cognitive factors prior knowledge and language abilities were included, as these two correlated with observation competency. In model 3, the two affective factors, emotional well-being and involvement, were included as well.

Table 20. Models in Hierarchical Regression Analysis of Factors Influencing Observation Competency

Step	Model	Total R^2	Incremental R^2	Significance of Change in R^2
1	Scientific Thinking Factors	.40	.40	<.01
2	Scientific Thinking Factors and Cognitive Factors	.41	.05	.08
3	Scientific Thinking Factors, Cognitive Factors, and Affective Factors	.62	.17	<.01

N = 69

Table 20 shows the changes in the explained variance across the models. While the addition of the cognitive factors did not significantly change the explained R^2, the addition of the affective factors led to a significant increase of the R^2 from 40 % to 62 % explained variance. When looking at the particular predictors (see

Table 21), domain-general scientific reasoning and biology understanding were significant predictors in every of the three models. Prior Knowledge was only a significant predictor in Model 2, but not in Model 3. Language Abilities were no significant predictor in either of the models. Involvement and Emotional Well-Being became both significant when entered as predictors in Model 3, with Emotional Well-Being having the highest beta-weight of all predictors.

Table 21. Hierarchical Regression Analysis Summary Predicting Observation Competency

Variable	*B*	*SE B*	*ß*	*t*	*p*
Model 1					
Domain-General Scientific Reasoning	4.45	1.35	.35	3.30	<.01
Biology Understanding	0.57	0.16	.39	3.59	<.01
Model 2					
Domain-General Scientific Reasoning	3.13	1.47	.25	2.13	.04
Biology Understanding	0.48	0.16	.32	2.98	<.01
Prior Knowledge	6.80	3.24	.22	2.10	.04
Language Abilities (Vocabulary)	3.55	5.34	.07	.66	.51
Model 3					
Domain-General Scientific Reasoning	2.56	1.24	.20	2.06	.04
Biology Understanding	0.29	0.14	.20	2.09	.04
Prior Knowledge	3.67	2.81	.12	1.30	.20
Language Abilities (Vocabulary)	4.23	4.56	.09	.93	.36
Involvement	2.49	1.24	.19	2.01	.04
Emotional Wellbeing	7.44	1.93	.35	3.86	.00

N = 69

5.4. Evaluation of the Training Activities

5.4.1. Evaluation of the Sessions and Theme Blocks

For the evaluation of the sessions, each trainer filled out a questionnaire on children‘s participation during the session. For each child, they scored the child‘s activity in the participation of the session, the difficulties the child had with following the instructions and the observation behavior he/she showed (N (children): between 22 and 39; N (trainers): 5). As Figure 7 displays, there is a rather high activity throughout the program, an overall increasing observation behavior, and slightly decreasing display of difficulties. However, there are also quite some up and downs. E.g., the fifth lesson seems to have been problematic, with higher difficulties and consequently less activity and observation behavior of the children.

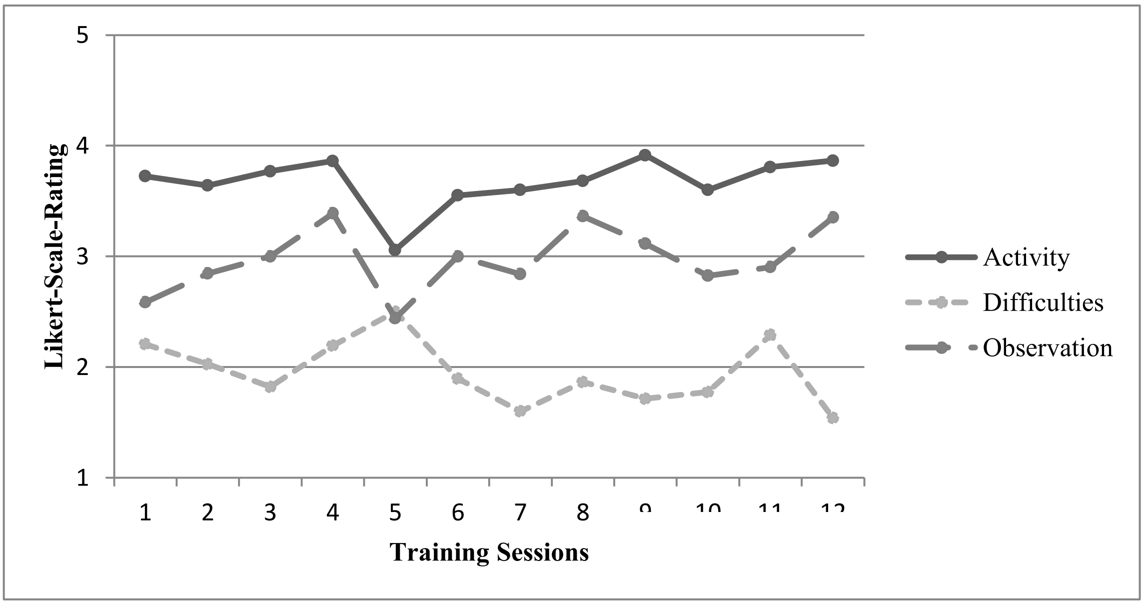


Figure 7. Evaluation of Children's Participation in the Sessions by Trainers

However, there is also a lot of missing data because many children missed several sessions due to illness or being on holidays with their families. In order to find out if children‘s observation behavior really increased while the difficulties decreased, we took the data for each child‘s individual first two and last two sessions and compared the means. Children who attended less than 5 sessions were excluded from the analysis (final $N = 42$). The results show a significant increase of observation behavior ($t(41) = -4.56$, $p < .001$ and a decrease of children‘s display of difficulties ($t(41) = 3.28$, $p < .01$).

Taking a look at the theme blocks (see Table 22), the activity was rather stable across the blocks and always above 3.5 except for the birds theme. The lower activity here could be explained with the increased difficulty, as the birds theme block seems to have been the most difficult. The tree and the soil sessions, in contrast, were easier than the others. At the same time, those two show to have been most effective in triggering children‘s observation competency.

Table 22. Means and Standard Deviations for Children's Performance in the Theme Blocks

Item		Themes				
		Bean	Tree	Birds	Soil	Hand and Feet
Activity	M	3.72	3.81	3.32	3.65	3.69
	SD	.92	.68	.74	.75	.74
Difficulty	M	2.14	1.68	2.24	1.79	2.01
	SD	.73	.57	.81	.83	.79
Observation	M	2.91	3.20	2.72	3.03	2.84
	SD	.97	.80	.86	.89	.89

$22 < N < 39$

5.4.2. Evaluation of the specific activities

After each of the 12 sessions, trainers had written down comments about the session, sorted by *that worked well*, *that did not work well*, *that has been changed by me*. For these comments, a category system was developed based on the methods of Grounded Theory (cf. Strauss & Corbin, 1994). For this purpose, the data was read repeatedly and overarching categories formulated based on recurring similarities and differences in wording and meaning of the comments. Finally, 12 categories were identified. Descriptions of the categories and examples of comments are displayed in Table 23.

Two of the categories, *organization,* and *locality*, were based on specifics of the concrete situation and are not relevant for the general assessment of the games.

The categories *difficulty*, *discipline*, *well-being*, *creativity* and *knowledge/ideas* cover the fundamental aspects of a good game or activity and are therefore relevant for the assessment of the activities.

We identified the categories *verbalizing*, *teamwork*, *concentration*, *motivation* and *observation* as highly relevant for the assessment as a good game or activity. They form the most sophisticated skills and competencies that are crucial learning observation.

Table 23. Description of and Examples for Categories in the Qualitative Analysis of the Materials

Rele vance	Category	Description	Example *that worked well*	Example *that didn't work well*
0	organization	Preparation and execution of the sessions: materials, process of the session	Bird Memory: Has to be guided closely (whose turn is it?), but then it worked well	Guessing the fruit: difficult to blindfold all children fast enough
	locality	Location (indoors or outdoors); spatial conditions, such as the size of the available space, and the properties, such as the equipment the location entails	Flower Pictures: children enjoyed to do this activity outside	Animal Pantomime: too little space in the kindergarten (renovation). Stopped activity because no concentration (distractions outside the room)
+	difficulty	Difficulty / demands of an activity; problems children have in following the instructions or in the use of new materials; also possibility for differentiation (individually adapting the level of difficulty)	Soil Animal Dice Game: slowly increasing difficulty by first using two dice, then three	Animal Prints: children were fascinated by animals, but had difficulties with identifying the relevant characteristics (tracks were rich in detail, outline shape not clear)
	discipline	Behavior of the children during the training sessions; willingness and ability of the children to follow set rules and show desired behaviors	Woodlouse Story: motivation, active participation, follow rules gladly	Flight Characteristics of Feathers: Children were too vigorous, some feathers were lost, were thrown around
	well-being	Well-being and satisfaction of children with the situation, influenced by the environment, the task and the types of interaction	(no comment)	Observing Hands: blindfolding scared some children

	creativity	Children‘s creativity in the execution of the tasks and use of fine motor skills to represent the observed (e.g., when drawing or sculpturing)	Complete Bird drawings: some children managed very accurate and beautiful drawing and coloring	Drawing the Bean: didn‘t observe properly but draw what they already believe (especially roots)
	knowledge/ ideas	prior knowledge of children they can use to name examples or make suggestions for conducting experiments	Planting the Bean: already knew a lot about the bean and conditions for planting	(no comment)
++	verbalizing	Children's ability and opportunities to put the observed and learned in their own words	Observing the bean: children described roots and germs; spontaneous comparisons	Memory of Smells: hard to describe characteristics of smells
	teamwork	Cooperation of children, social skills, willingness to interact with others and contribute their own opinion, following a common goal	Error Pictures: good cooperation of children, working in teams	Looking for prints: children only want to find their own fingerprint and not sort / find pairs
	concentration	Capability of children to turn their attention to an activity; Ability to execute an activity in a way that a given goal is achieved	Footprints: good concentration on finding differences, described many differences	Search list: No occupation with the things that have been found, going from one to the next
	motivation	Willingness of children to participate in activities clear signs of interest, enjoyment, curiosity and excitement	Observing the bean: Children were excited to see the bean again	Drawing the bean: ―not again…”
	observing	how purposefully, attentively and precisely children perceive and process the details and features of the materials	Soil Animal Originals: Naming of colors / properties and comparing to originals, accurate observing	Tactile Parcours: At the end, some children still did not have the impression that you can feel better with no socks

0 = not relevant for the general assessment, + = relevant for the general assessment; ++ = highly relevant for the general assessment

Taking into account the number and type of comments, each activity could be assigned to one of the groups "very good", "good", "some change needed" and "many problems". The evaluation of all the activities is displayed in Table 24, with a mention of the relevant categories and, where applicable, a short description of problems or changes.

The first thing that is noticeable is the fact that most of the inquiry activities are very good or good activities. Observing the bean, observing the woodlouse and differentiating surfaces all had very good effects on children‘s observing and verbalizing of the observations. Activities in that manner seem to be a good opportunity to foster children‘s observation competency. At the same time, many of the activities that did not work well or showed many problems were activities meant to foster interpretation. Usually, the problems occurred because the level of difficulty was not ideal for the children. In most cases when there were negative comments about the level of difficulty, this also had an impact on children‘s motivation. Therefore, especially the activities for fostering interpretation skills should be checked in order to find the ideal level of difficulty, and find ways to differentiate the tasks according to children‘s different performance levels.

Another interesting finding is the fact that motivation and verbalizing seem to be linked. When the atmosphere is stimulating and the children are motivated, this seems to open up opportunities for them to verbalize their thoughts and exercise describing observations. This shows once more how important motivational factors are for children‘s learning processes. Considering the fact that the motivation is closely linked to the right degree of difficulty of the tasks, this emphasizes once more the importance of meeting children‘s performance level by differentiation.

Finally, the specification of a task and close guidance by the instructor seems to be just as important. While the general animal pantomime showed some problems, the specific versions soil animal pantomime and animal gaits showed much better effects. Likewise, some of the games that need changes mainly need a more close guidance by the trainer. This is true for both the memory of smells and the fingerprints game, where in both cases the activities only worked when the trainer discussed the material with the children in several steps instead of letting them try to sort the materials on their own. Overall, it can be said that most of the activities worked well and that for most problems there are already good ideas how to solve them.

Table 24. Evaluation of the Activities According to Categories

Group	Activity	Characteristics
Very good	Observing the Bean	Good for motivation, verbalizing
	Observing the Woodlouse	Good for Verbalizing, Motivation, ideas and teamwork
	Woodlouse Story	Good for Observing, concentration, high motivation; but much space needed
	Tactile Box	Good for observing, high concentration and motivation
	Differentiating Surfaces	Good for observing, creativity, motivation and verbalizing
	Human Footprints	Good for observing and verbalizing, good level of difficulty, therefore high concentration
	Nature Findings	Good for observing and verbalizing, high motivation and concentration
Good	Planting the bean	Good for Motivation
	Drawing the bean	Good for Motivation, creativity
	Error pictures	Good level of difficulty, motivation, teamwork, verbalizing
	Flower Pictures	Good for Verbalizing, Concentration, motivation, and creativity; but has to be guided closely, otherwise very difficult
	Observing the tree	Good for verbalizing, concentration and discipline

Good	Forest Camera	Good for motivation and observing, but very dependent on stimulating environment
	Comparing soils	Good for Verbalizing, Teamwork
	Search list	Good potential for differentiation; if task has adequate level of difficulty, good for motivation, concentration, and discipline
	Bird Memory	Good for observing and verbalizing; some children not motivated because game is so well-known
	Complete Bird Drawings	Good for observing and creativity; sometimes problems with motivation/ well-being because children feel pressure to replicate original photos
	Describing Birdsong	Good for verbalizing and motivation
	Sorting birdsong	Good for verbalizing and motivation
	Tracing feathers	Good for observing and creativity, but also rather difficult
	Systematizing Feathers	Good for Verbalizing, some problems with discipline
	Flight characteristics of feathers	Activating knowledge/ ideas
	Detail Pictures	Good for Verbalizing and Motivation, but the trainer has to make sure that all children contribute
	Soil Animal Pantomime	Good for observing and motivation, but sometimes too easy to guess animal (therefore difficulties with describing before interpreting)
	Tactile Parcours	Good teamwork and motivation, verbalizing, but not all children make observations
	Observing Hands	Good for observing and concentration; some children did not feel well being blindfolded (instead just closing eyes as an option)
	Animal Prints	Good for observing and verbalizing, stimulates knowledge/ ideas, highly motivating, but rather difficult for some children
	Animal Gaits	Good for verbalizing and motivation; needs lots of space and good organization
	Looking for prints	Good for observing and activating knowledge, high motivation, sometimes difficult to clearly recognize the feet in the pictures
	Guessing the fruit	Good for verbalizing, motivation and teamwork, problems with blindfolding (well-being and organization)

Good	Sorting fruit and juice	Good for verbalizing, motivation and teamwork
	Soil animals Originals	Good for Observing and motivation, in the beginning, high concentration but decreases quickly for some children
Some changes needed	Memory of smells	High motivation but difficulty with the description of the smells → individual discussion of smells necessary (instead of free sorting by children)
	Bird Quartet	Very good for observing but high difficulty can lead to lack of motivation and verbalizing →Talk about birds in the group before, let children get used to characteristics
	Bird Puzzles	Motivating, but puzzle pieces have very different levels of difficulty → purposeful distribution of pieces by the trainer necessary for differentiation
	Soil animal Dice Game	Very different in the degree of difficulty for children (―some slow, some bored" → having them play in more homogenous groups
	Modeling soil animals	Very good for creativity and high motivation, but very difficult → big models of insects instead of originals, so that children can touch them and understand forms and proportions better
	Fingerprints	High motivation but very difficult when sorting all prints at the same time → always just comparing one print to all the others
Many Problems	Animal Pantomime	High difficulty, location (much space needed)
	Dice with facial expressions	High difficulty, therefore problems with motivation/ discipline/ concentration
	Imitating flying	No positive comments, problems with discipline
	Shoe salad	Task was too easy, therefore no describing or justifications, just sorting
	Complete the animal	Activating knowledge, but too easy, therefore children were quickly bored

5.5. Training Effects on Observation Competency and Biology Understanding

For investigating the training effects, children‘s observation competency in the pretest and posttest was compared, as well as their biology understanding. First, a Rasch analysis of both scales was conducted, then variance analyses, correlations, and regressions were calculated in order to understand children‘s development over time and the effect of the treatment. As a first step, t-tests were calculated in order to check for group differences between training- and control group already before the intervention. As we are assuming the null-hypothesis is correct (no group differences), the critical p-value is set to .2. Doing that, only the test on biology understanding shows group differences ($t = 1.59$, $p < .2$), with the control group being significantly better than the experimental group. All values can be found in Table 25.

Table 25. Group Differences Between Control Group and Experimental Group Before the Intervention

Measure	Control group		Experimental group			
	M	*SD*	*M*	*SD*	*t*	*p*
Observation Competency	14.75	4.08	15.57	4.28	-.82	.41
Biology Understanding	6.00	2.83	4.98	2.66	1.59	.12
Scientific Reasoning	.39	.32	.33	.33	.78	.44
Prior Knowledge	.15	.18	.12	.12	1.03	.31
Theory of Mind	1.00	.67	1.17	.80	-.99	.33
Executive Functions	.70	.19	.71	.19	-.19	.85
Language (Vocabulary)	.92	.09	.93	.07	-.44	.66

N = 75

5.5.1. Observation Competency

5.5.1.1. Rasch Analysis

The data was Rasch-analyzed in order to take into account the difficulty of the items, and especially control for potential changes in difficulty from pretest to posttest. The Rasch analysis showed no anomalies or problems with the instrument. The mean measures for both pretest and posttest are beneath zero, suggesting that the test was still rather hard for children.

Therefore, a DIF Analysis of the pre- and post-measures was conducted, whose results are shown in Figure 8. The two lines show the range of the 99%-confidence-interval around the middle line. Items lying within this range have been of the same difficulty level pretest compared to posttest. Items above the upper line have been relatively harder in the posttest compared to the pretest. Items below the lower line have been relatively easier in the posttest compared to the pretest.

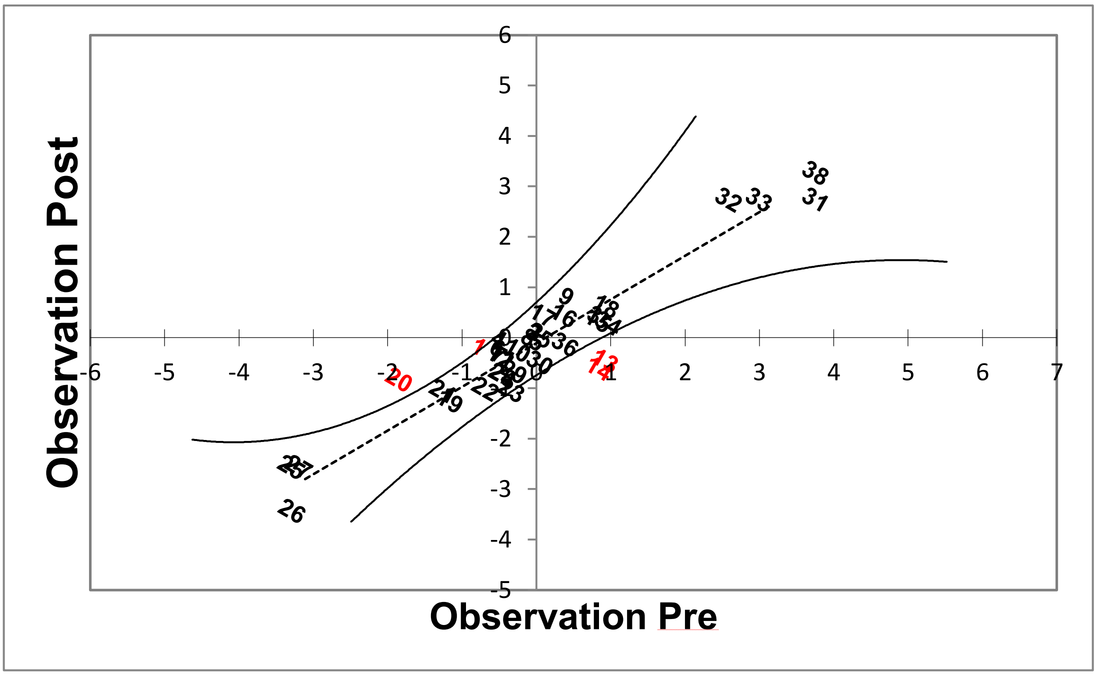


Figure 8. DIF Analysis of Observation Comparing Pretest and Posttest

The analysis showed that almost all items behaved similarly in pre- and posttest and did not change in their relative level of difficulty. Two items became harder in the posttest compared to the pretest (item entry 1, 20), two items became easier (item entry 13, 14). These items were therefore floated, meaning they were handled as different items in the pre- and posttest, allowing them to potentially mark different locations on the trait in pre- and posttest.

All further analyses have been conducted with these final measures of children‘s observation competency.

5.5.1.2.Pre-Post-Group-Comparison

Table 26 displays the mean values for the control group and experimental group in the pretest and the posttest. The first further analysis of the data was an ANOVA with repeated measures (see Table 27). It showed a significant growth of children‘s performance over time, but no effect of the training groups. The results are also displayed in Figure 9.

Table 26. Means and Standard Deviation In Pretest and Posttest, Observation Competency

	Control group	Experimental group	Total
N	30	40	70
Pretest			
M	-.73	-.66	-.69
SD	.57	.49	.52
Posttest			
M	-.50	-.31	-.39
SD	.43	.42	.43

Table 27. ANOVA with Repeated Measures, Observation Competency

	df	*Sum of Squares*	*Mean Squares*	*F*	*p*	η^2
		Between subjects				
Treatment	1	.59	.59	1.64	.20	.02
Error 1	68	24.60	.36			
		Within subjects				
Time	1	2.85	2.85	31.84	<.01	.32
Treatment x Time	1	.13	.13	1.46	.23	.02
Error (time)	6.087	68	.09			

N = 70

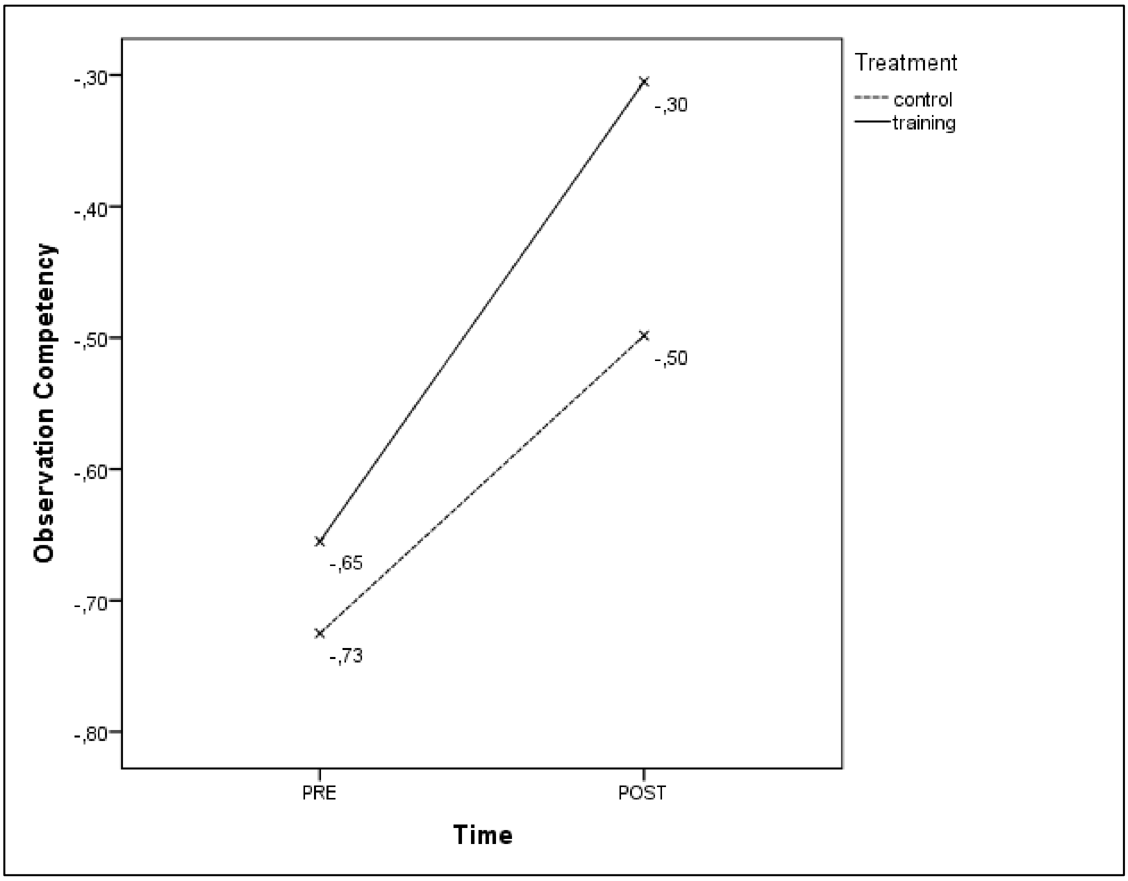


Figure 9. ANOVA with Repeated Measures, Observation Competency

Table 28. Correlations Between Pre, Post, Change Measures and Executive Functions

Variable	*M*	*SD*	1	2	3	4
1. Observation Pre	-0.69	0.52	–	.62**	-.60**	.04
2. Observation Post	-0.39	0.43		–	.26*	.24*
3. Observation Change	.30	.42			–	.19
4. Executive Functions	1.55	.95				–

N = 70

For further analysis, correlations between the pre- and post-measures, as well as the change in the observation competency (post-measure minus pre-measure) and the executive functions, were calculated. The results, displayed in Table 28, show that the observation pre-measure is negatively correlated with the change in observation competency, meaning that children who were worse in the beginning showed more improvement than children who were better at the pretest. This compensation effect probably also lead to the reduced standard deviation from pretest to posttest.

5.5.1.3. Differential Effects

In order to find out whether the training had different effects on children depending on their prior abilities, we did a median split for the observation competency in the pretest and calculated ANOVAs for both the better and the worse half. Table 29 displays the groups resulting from the median split, the mean values of these new groups can be found in Table 30.

Table 29. Groups in Median Split

		Treatment		Total
		control	training	
Median split	Worse 50%	17	18	35
	Better 50%	13	22	35
	Total	30	40	70

Table 30. Means and Standard Deviations of Observation Competency, Median Split

		Pretest		Posttest	
		Control	Training	Control	Training
Worse 50%	*M*	-1.11	-1.09	-.62	-.57
	SD	.40	.31	.44	.33
Better 50%	*M*	-.23	-.30	-.34	-.09
	SD	.31	.26	.38	.35

For the worse 50%, there is only an effect of time, but no effect of treatment or the interaction (see Table 31). The effect of time has a large effect size (η^2 = .74 . In contrast, for the better 50%, there is only a significant effect of the treatment, meaning only the training group improved over time. This effect only has a small effect size (η^2 = .14 .

Table 31. ANOVA with Repeated Measures, Observation Competency, Median Split

		df	*Sum of Squares*	*Mean Squares*	*F*	*p*	η^2
			Between subjects				
Worse 50%	Treatment	1	.02	.02	.09	.76	<.01
	Error 1	33	7.53	.23			
			Within subjects				
	Time	1	4.40	4.40	93.30	<.01	.74
	Treatment x Time	1	.00	.00	.08	.78	<.01
	Error (time)	33	1.56	.05			
			Between subjects				
Better 50 %	Treatment	1	.13	.13	1.00	.33	.03
	Error 1	33	4.18	.13			
			Within subjects				
	Time	1	.04	.04	.49	.49	.02
	Treatment x Time	1	.44	.44	5.27	.03	.14
	Error (time)	33	2.75	.08			

The results are also displayed in Figure 10, where it can be seen that the growth of the worse 50% has a higher slope than the better 50%, but still, the worse 50% do not even reach the starting level of the better 50%.

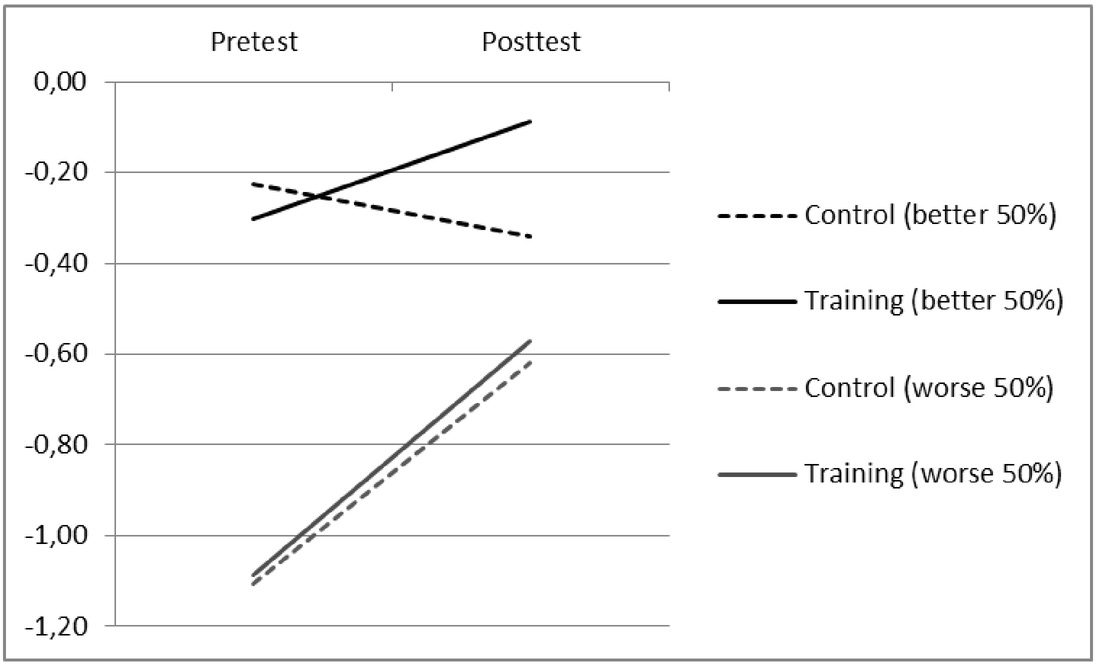


Figure 10. ANOVA with Repeated Measures, Observation Competency, Median Split

5.5.1.4. The Role of Executive Functions

Children's executive functions (measured at pre) were not correlated with their performance in the pretest for observation competency, but with their performance in the posttest (see Table 28).

Table 32. Regression Analysis Summary for Predicting Children's Observation Competency Post

Variable	*B*	*SE B*	*β*	*t*	*p*
Treatment	0.15	0.08	.17	1.85	.07
Observation Pre	0.49	0.08	.60	6.53	.00
Executive Functions	0.09	0.04	.21	2.27	.03

N = 70

In a regression analysis, we then investigated how pretest, treatment, and executive functions predict children's posttest measure. All predictors together explained 46% of the variance ($R^2 = .46$, $F = 18.36$, $p < .01$). While both observation competency in the pretest and executive functions proved to be significant predictors, the treatment did not get significant (see Table 32).

5.5.2. Biology Understanding

Correspondingly to the procedure with the observation competency, we conducted a Rasch analysis with the data of the biology understanding test for the pre-post-comparison.

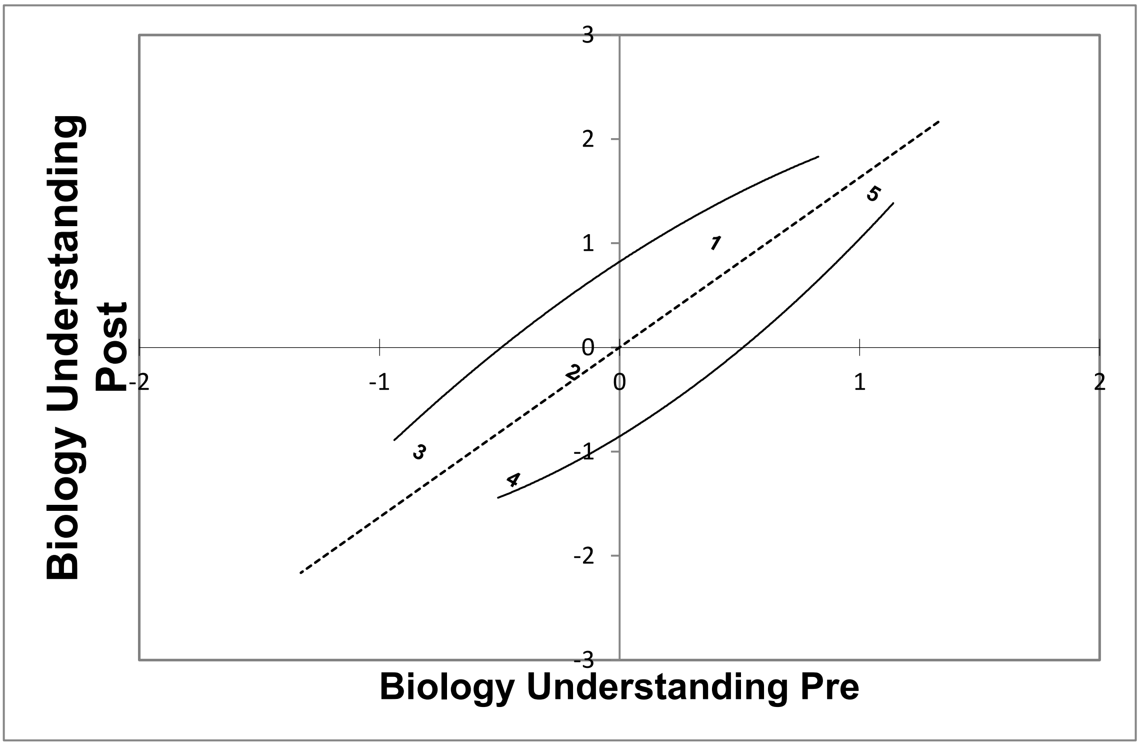


Figure 11. DIF Analysis Pretest Compared to Posttest of Biology Understanding

The plot of the DIF-analysis (see Figure 11) shows that all five items have the same difficulty in the posttest compared to the pretest, meaning the instrument worked the same way at both measurement points.

Table 33. Means and Standard Deviation in Pre- and Posttest, Biology Understanding

	Control group	Experimental group	Total
N	30	40	70
Pretest			
M	5.93	4.85	5.31
SD	2.90	2.57	2.75
Posttest			
M	6.40	6.05	6.20
SD	2.79	3.51	3.20

Table 34. ANOVA with Repeated Measures, Biology Understanding

	df	*Sum of Squares*	*Mean Squares*	*F*	*p*	η^2
			Between subjects			
Treatment	1	17.61	17.61	1.30	.26	.02
Error 1	68	920.13	13.53			
			Within subjects			
Time	1	23.81	23.81	5.66	.02	.08
Treatment x Time	1	4.61	4.61	1.10	.30	.02
Error (time)	68	285.93	4.21			

N = 70

Table 33 displays the mean values for the control group and experimental group in the pretest and the posttest. The next step was an ANOVA with repeated measures. It showed a significant growth of children‘s performance over time, but no effect of the training groups. The results are displayed in Table 34 and Figure 12.

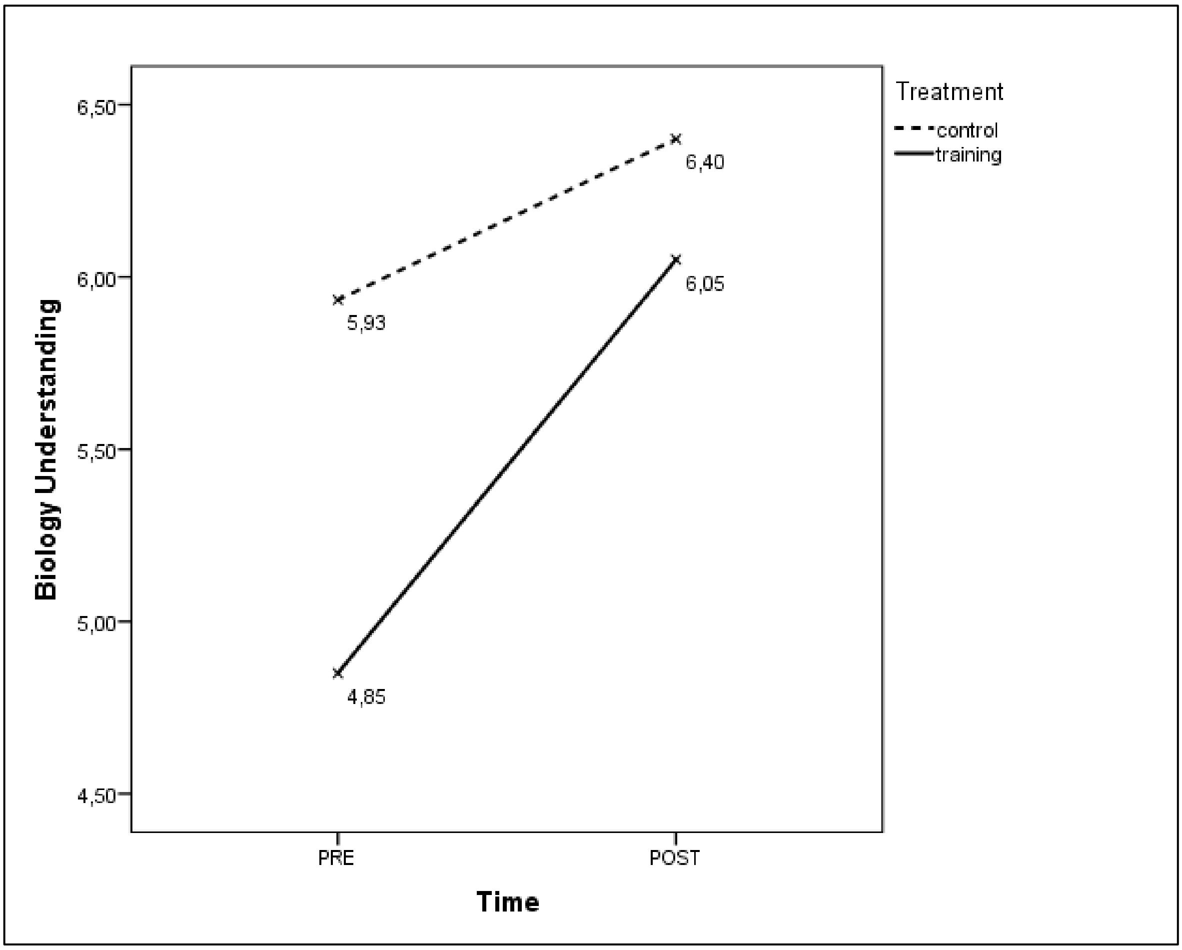


Figure 12. ANOVA with Repeated Measures, Biology Understanding

For further analysis, the correlations between the pre- and post-measures, as well as the change (post-measure minus pre-measure) and the executive functions, were analyzed. The results, displayed in Table 35, show that the pre-measure is negatively correlated with the change in biology understanding, meaning that children who were worse in the beginning showed more improvement than children who were better at the pretest. However, the post-measure is also highly correlated with the change and there is no decrease in the standard deviation as it was the case for the observation competency. Therefore, the results cannot be interpreted as a compensation effect and the data was not further analyzed grouped by a median split for this measure.

Table 35. Correlations between Pre, Post, Change Measures and Executive Functions

Variable	*M*	*SD*	1	2	3	4
1. Biology Understanding Pre	5.31	2.75	–	.53**	-.36**	.15
2. Biology Understanding Post	6.20	3.20		–	.60**	.32**
3. Biol. Understanding Change	.89	2.90			–	.21
4. Executive Functions	1.55	.95				–

N = 70

As was the case for the observation competency, there is a positive correlation of executive functions with the post-measure. Further analysis with a multiple regression should show the importance of executive functions, pre-measure, and treatment on the biology understanding.

The regression explained 38% of the variance ($R^2 = .38$, $F = 14.64$, $p < .01$). Both the pre-measure and the executive functions are significant predictors, while the treatment once again proves not to be significant (see Table 36).

Table 36. Regression Analysis Predicting Biology Understanding Post

Variable	*B*	*SE B*	*β*	*t*	*p*
Treatment	0.05	0.62	.01	0.09	.93
Biology Understanding Pre	0.56	0.11	.49	4.98	.00
Executive Functions Pre	4.80	1.65	.28	2.91	.01

N = 70

6. Discussion

After the presentation of the results, this chapter will now discuss the study, summarizing the results and setting them in relation to prior research, as well as looking critically at the methodology of the study and reflecting the implications for both further research and practice.

6.1. Summary of the Results

The results will be presented in order of the hypotheses derived from the two main research questions: First, I will summarize the results on the factors related to observation competency, discussing the relation with scientific thinking, cognitive factors, and affective factors. Secondly, I will go over the main points of the intervention, recapitulating the usability of the materials and the general effectiveness of the training.

6.1.1. Relations Between Observation Competency and Scientific Thinking

The research on scientific thinking differentiates between two facets, domain-general scientific reasoning and domain-specific science understanding. In the case of this study, a non-animistic biology understanding is the relevant domain-specific facet of scientific thinking. The hypothesis was that both domain-general scientific reasoning and biology understanding are related to children‘s observation competency, as both are needed for relevant observations. This hypothesis can be confirmed, as not only both factors correlated with observation competency but also proved to be significant predictors in a regression analysis. In fact, when putting them in a regression together with language abilities, language is not a significant predictor anymore. A mediation analysis showed that language abilities predict domain-general scientific reasoning and domain-specific biology understanding, which in turn predict the observation competency. This also implicates that the relation between domain-general scientific reasoning and biology understanding with observation competency is more than just the shared variance with language abilities: they do both have a specific, independent relation to children‘s performance in the observation task.

Looking at the structure of observation competency, the relation to domain-general scientific reasoning becomes clear: several general epistemic activities are part of observation, and both observation and scientific reasoning require a basic understanding of inquiry. Only when children understand the value of ―finding something out", does it make sense that they try to find something out by observation. The relation between domain-specific biology understanding and children‘s observation competency has not been shown in research yet, but has been conceptualized: Geary (2008) has posed the idea that children‘s naïve concepts and their tendency to focus on goal-direction (Evans, 2008) would be factors that could hamper children‘s performance in reasoning situations. More sophisticated knowledge about the domain would lead to better, unbiased reasoning. It has to be noted, though, that the data in this study does not give information about the direction of the relation: it is also possible that children who are better observers already gained a better understanding of their environment and have better developed concepts. A circular interaction of both processes is also possible: the domain-specific understanding has an impact on the correct use of strategies, and the correct use of strategies has an impact on the development of the understanding in the domain (Schauble, 1996). A recent trend in science education research is to stress the importance of domain-specific abilities for the quality of reasoning processes (Sinatra & Chinn, 2012). This study underlines that both factors are relevant for observation competency – at least for preschool age when the understanding of theory and evidence still develops (Kuhn, 2000; Piekny & Maehler, 2013). When children are still lacking the basic understanding of inquiry, it seems not feasible to engage them in inquiry tasks. At the same time, specific understanding of the domain should help them make meaningful observations.

6.1.2. Relations with Cognitive Factors

The second hypothesis focused on the relations of cognitive measures with children‘s observation competency as well as with their domain-general scientific reasoning. Both were expected to be effected by children‘s language abilities, prior knowledge, executive functions, and theory of mind. This hypothesis can be only partly confirmed. Language abilities had an impact on both variables, and prior knowledge was significantly correlated with children‘s observation competency. The impact of concrete prior knowledge and language abilities on observation competency is a direct replication of the results by Kohlhauf et al. (2011) and confirms the general trend in research on scientific observation to find an impact of prior knowledge (Eberbach & Crowley, 2009). However, when looking at all related factors together (model 3 of the hierarchical regression analysis), those two factors are no significant predictors of observation competency anymore. It seems that scientific thinking skills and affective factors are more important for the observation process. However, it cannot be ruled out that covariations of the variables are responsible for these results, as most variables were correlated with children‘s language abilities, and both scientific thinking components correlated with prior knowledge.

Executive functions and theory of mind did not correlate with observation competency, but they were linked with children‘s domain-general scientific reasoning. This result could be further strengthened by a regression analysis showing that language abilities, executive functions, and theory of mind are all independently relevant predictors of scientific reasoning. This result is in line with prior research on scientific reasoning, as studies have found language abilities, theory of mind , and executive functions to be related to children‘s scientific reasoning abilities (Mayer et al., 2014; Piekny et al., 2013b; van der Graaf et al., 2016).

6.1.3. Relations with Affective Factors

The third hypothesis was about the relation of affective factors to children‘s observation competency, expecting both children‘s emotional well-being and their involvement to have a significant impact on the showed observation behavior. There was indeed a significant prediction of the observation competency through emotional well-being and involvement, with a rather large effect. A mediation analysis showed that the effect of emotional well-being is partly mediated by involvement. This is in line with the assumption of the control-value-theory that engagement, which is a similar construct to involvement, is a mediator between emotions and performance (Pekrun & Linnenbrink-Garcia, 2012). In general, the results underline the importance of affective factors and situational interest on observation – as Johnston (2009) puts it: ―children observe only what interests them” (p. 2513). The results are not only interesting for research on observation competency, though, but also relevant for the Leuven Well-Being and Involvement Scales. While prior studies mainly investigated factors that influence children‘s well-being and involvement (Declercq, 2014; Goldspink et al., 2008; Laevers & Heylen, 2003), this study confirms that these affective facets do in fact have an impact on children‘s cognitive performance. Thereby, the results confirm the deep level learning model (Laevers, 2000), that expects that the environment effects children‘s well-being and involvement, which in turn has an impact on their performance on learning tasks. However, as these results of the study consist only of correlational data, one has to be careful with causal interpretations. While this direction of the effect is plausible, it is also possible that the factors influenced each other the other way round: children who better understood what they had to do also enjoyed participating more and were more perseverant than those who were overwhelmed by the assignments. As the effects of affects are assumed to run reciprocally (Pekrun & Linnenbrink-Garcia, 2012), a mixture of these mechanisms would also be plausible.

6.1.4. Summary of Related Factors

The final hypothesis about the scientific thinking, cognitive, and affective factors was that they all have an individual impact on observation competency when looking at them together. This hypothesis can only partly be confirmed. The relations between observation competency and the investigated factors are summarized in Figure 13. The cognitive factors, prior knowledge and language abilities, did not significantly increase the explained variance when adding them to the scientific thinking components. However, the addition of the affective factors led to a significant increase of the explained variance. In the final model, both scientific thinking components (domain-general scientific reasoning and domain-specific biology understanding) and the two affective factors (emotional well-being and involvement) are significant predictors of observation competency. This result strengthens the position of observation competency as a scientific thinking skill with both domain-specific and domain-general aspects. It also emphasizes the importance of affective processes for reasoning processes (Fischer et al., 2014).

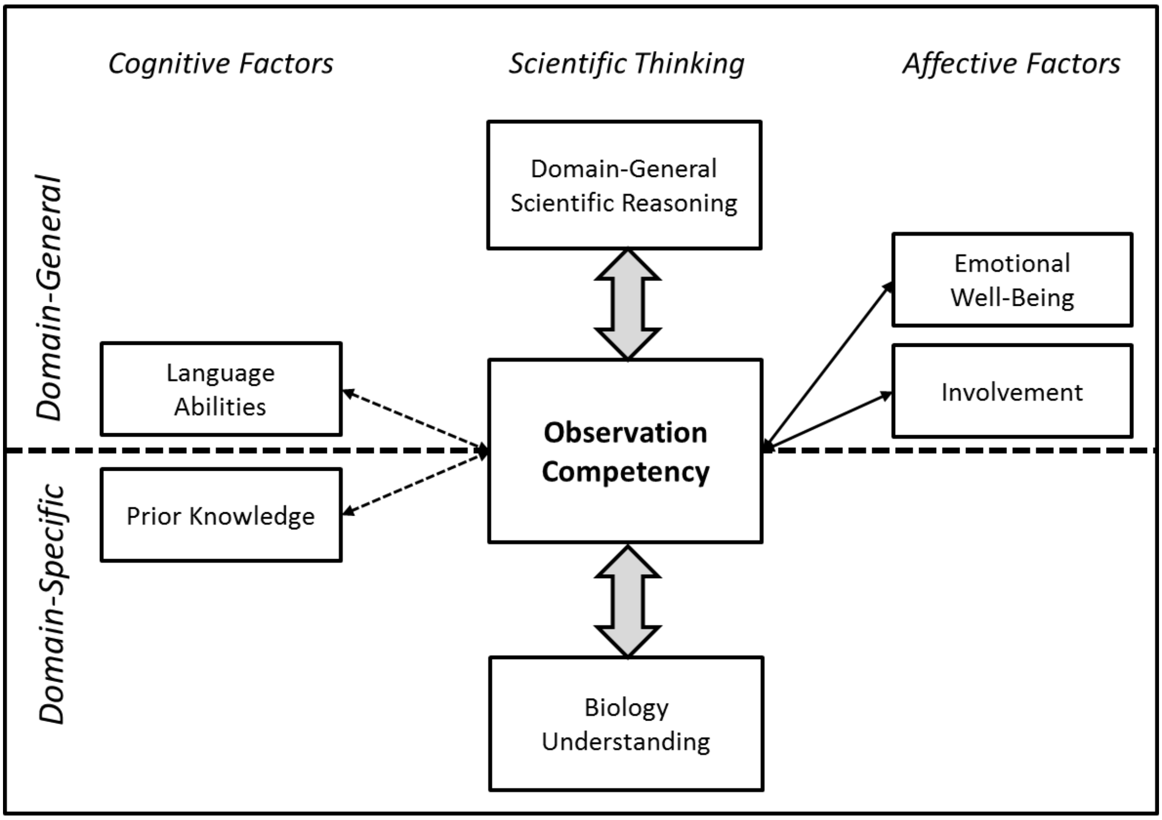


Figure 13. Summary of Results on Factors Related to Observation Competency

6.1.5. Training Materials

The feedback from the trainers on each of the intervention sessions was used to analyze whether the materials are suitable for training observation competency of preschool children. Trainers had filled out a questionnaire with two scales on their evaluation of training sessions and materials. One was a 3-item-scale on children‘s behavior, the other three open questions dealt with the activities of the session. The data from the rating of children‘s individual behavior showed high activity for almost the complete program, an increasing observation behavior, and slightly decreasing display of difficulties.

Overall, the hypothesis that the materials are suitable for training preschool children could be confirmed. Most of the games and activities proved to be good or very good for fostering children‘s observation competency. However, it has to be noted that the activities of the inquiry dimension had the best results, while the activities for training children‘s interpretation skills often showed problems with the difficulty level. For the activities that were sorted as problematic or changes needed, the trainers gave good ideas in the feedback on how the tasks could be changed or adapted.

As Scherres (2013) argues, differentiation and thereby meeting children‘s performance potential with the task difficulty leads to higher motivation and performance. The results of this study are in line with these assumptions: When there were problems in the sessions, they could be explained by problems with the difficulty of the tasks or the motivation on the children‘s side. The trainers also specified that children often need close guidance in the activities in order to concentrate on the given tasks. This reinforces the argumentation by Grell (2010) that children need adults to help them guide their learning processes in order to make them thorough and profound.

6.1.6. Training Effects and Longitudinal Analysis

The last two hypotheses were related to the intervention: both developmental and intervention effects were expected for children‘s observation competency and biology understanding. While there was a rather robust developmental effect, the intervention effect could not be found for the whole sample. To understand the differential effects of the intervention on children‘s observation competency, the data was analyzed again after a median split according to their pretest performance: the results showed that for the worse 50% of the sample there was a large improvement over time, independent of training. For the better 50%, we could not find this development effect, but an intervention effect: here, only the observation competency of the training group improved significantly over time. It seems that general development does have an impact on observation competency on a low level, while the higher, more sophisticated observation skills need to be fostered. It can also be interpreted as a ―Matthew effect" (Walberg & Tsai, 1983): only children on a specific level of competency profit from a training, while for the others the tasks are too hard or the newly learned cannot be integrated into their knowledge structure yet. It is also possible that the basic parts of observation competency cannot be learned, but develop independently from instruction, while the advanced aspects of the competency can be fostered. One hint into this direction is the correlation of executive functions at the pretest with the posttest performance, which will be discussed below.

For children‗s biology understanding, there were no effects of the training, but a developmental effect: all children‗s biology understanding significantly improved from the pretest to the posttest. This indicates that the understanding of the concept of life develops independently and could not be fostered, at least not by a training on children‗s observation competency.

Further analysis showed that executive functions at the pretest, and not the participation in the training, had an impact on both children‗s observation competency and their biology understanding at the posttest. It has to be noted that the observation competency and the biology understanding do not correlate with executive functions at the pretest, indicating a delayed developmental effect. In contrast to the results of this study, Zaitchik et al. (2014) found a direct correlation of executive functions and biology understanding measuring both at the same time point. They explain this effect with conceptual change: in order to develop from naïve biology to a vitalistic view of biology, conceptual change is needed. Conceptual change meanwhile has high requirements on children‗s executive functions: several concepts have to be represented at the same time in the working memory, false conceptualizations have to be inhibited, and finally it has to be shifted to the correct concept. It is possible that in this study children‘s executive functions at the first measurement influenced their conceptual change in the next five months. This would have led to a biology understanding at the second measurement, which is dependent on the executive functions five months earlier. It is possible that Zaitchik et al. (2014) found the correlation at the same point in time because the executive functions in preschool are rather stable. It is also possible that the sample of the current study was not large enough to find the effect at the first measurement. Further research should be conducted to investigate the exact relation between executive functions and biological competencies of preschoolers.

6.2. Limitations

The results of this study are mainly in line with prior research and may have interesting implications for both theory and practice. However, it has to be kept in mind that there are factors limiting the explanatory power of the results. These factors will now be discussed, sorted by limitations concerning the sample, the intervention, and the instruments.

6.2.1. Limitations of the Sample

Firstly, there were some constraints due to the sampling of the study. Although it was large enough to find specific effects, the small sample size leads to a rather small test power. Consequently, the results of the influencing factors investigated in this study should be reinforced with more studies, ideally direct replications of this study‘s design. Additionally, we could not calculate more complicated models. With a larger sample from more different preschools, it would be possible to calculate structure equation models, looking at different models as well, e.g., differentiating between the components of observation competency and identifying specific factors effecting these components.

For the intervention, it is possible that we did not find a (small) effect of the training on all participating children because of the sample size. Another problem is that the data was in fact not completely independent but nested in subgroups. The study took place in five different preschools, each of which had a training and a control group. There were four different trainers conducting the training. It would have been interesting analyzing the data taking into account the nested structure of the data. However, these analysis methods – e.g., multi-level-analysis – are not feasible to calculate with only four or five subgroups (Woltman, Feldstain, MacKay, & Rocchi, 2012). Therefore, classical analyses were used, reducing measurement error by having both a control and a training group in each kindergarten. These groups also showed not to differ on most assessed variables, indicating that the groups were homogeneous enough for the following analyses.

6.2.2. Limitations of the Intervention

The feedback on the intervention was mainly positive. However, there were also some difficulties with it. The materials had probably too much focus on the describing of details, too little tasks on fostering inquiry, and some activities for interpreting were not working out as expected. Ideas on solving these problems have already been described in detail in the results on the materials (see chapter 5.4). The training only showed effects for the children who already started on a higher level of observation competency, suggesting it might have been too demanding for some children. This problem should also be solved by a better differentiation in the individual activities, but it could also be considered to have an even more basic introductory level with tasks fostering children‘s general inquiry skills or their biology understanding, as these proved to be important for children‘s observation competency.

The aim of the structure of the intervention study was to balance internal and external validity. While internal validity asks for comparability and controlled variability, external validity has the goal of having as realistic conditions as possible. In order to have controlled conditions, there were structured manuals with clear instructions and learning goals for the trainers, who were trained themselves before and supervised throughout the implementation. In order to have realistic conditions, the intervention took place on an almost weekly basis over five months instead of having one blocked training. However, it might have had a negative impact on the results that the interventions were short but stretched out over a rather long period. Children may not have connected the learned content as they should, and many children missed several sessions due to illness. Another point hurting the external validity is the fact that not the preschool teacher but external trainers conducted the training. While this was important for controlling for the correct conduction of the sessions, this changed the learning situation. Preschool teachers usually know the competency level of the children in their group better than the trainers did, which makes differentiation easier. They could also have integrated the training into a daily routine instead of having the rather artificial situation of a weekly training. Still, in order to secure the general effectiveness of the training, a study with controlled conditions is necessary. Ideally, the training study would be repeated with improved materials in such a design, before further studies could investigate the applicability of the materials in the daily life of preschool.

6.2.3. Limitations of the Instruments

One side goal of the study was to replicate the usability of the observation competency test by Kohlhauf (2013). This only worked partially: while the situation itself proved to be fitting for generating observation behavior, the coding scheme was not fine-graded enough to represent the ability levels of the children in this sample. Therefore, a new scheme was developed with more items and a more grades per items. As the coding was changed, the results cannot be sorted into competency levels as they could with the original test. However, it was more important for this study to investigate the interindividual differences between the children due to influencing factors or training participation. This worked well with the new coding, which had good results concerning interrater reliability, reliability, and difficulty and can be recommended for measuring children‘s observation competency.

For the measurement of the affective states of the children, there are two limiting factors: firstly, the instrument measured emotional well-being and involvement, which are rather broad concepts, instead of looking at the impact of specific emotions or motivational facets. However, this was due to the age of the participants, who are not yet able to reliably report their own emotions (Fabes, Eisenberg, Nyman, & Michealieu, 1991). Therefore, an observational data collection method was selected for this study, which is not suitable for measuring more specific constructs like intrinsic motivation. This also leads us to the second limitation, which is that both measurements, children‘s affects and their observation competency, were derived by observation. Although two different coders analyzed the two constructs, it is still possible that spill-over effects at least partly caused the high correlations. However, if that were the case, a higher correlation between emotional well-being and involvement would be expected because the same person coded those. In fact, emotional well-being and involvement both showed higher correlations with observation competency than with each other. Consequently, it can be assumed that spill-over effects could be kept to a minimum.

The instrument for measuring scientific reasoning consisted of two tasks, measuring children‘s basic understanding of scientific inquiry. It showed good performance, looking at the fact that there are specific relations to theoretically related constructs. However, the test battery consisting of only two tests was rather small. Instruments for measuring children‘s scientific reasoning in preschool age are still scarce, but ideally, future studies would use a larger test battery covering several aspects of scientific reasoning.

Both biology understanding and children‘s concrete prior knowledge were measured. In order to keep the two tests distinctive, the prior knowledge test was only about the animals used in the observation situation. The fact that the tests only correlated moderately with each other (as they both also did with other constructs) indicates that two different constructs were actually measured. The prior knowledge test, which had originally been used for a sample with a wider age range, had some items that were too hard for the children of this sample and had to be deleted. While the remaining items showed satisfactory reliability and variance between children, more items should be developed when using it for future studies on preschool children samples.

Though children did the complete language test CITO, it was later decided to only use the scale “vocabulary” for measuring children‘s language abilities in order to make a clear differentiation from other cognitive measures possible. In fact, for a good performance on the subscale “text comprehension” children probably need more working memory than actual vocabulary. While it is unfortunate that rather a lot of testing time of the children is not used for the results of this study now, it was helpful to use the overall score for the exclusion of children with poor language abilities from the study. Additionally, the complete test results could be used to give interested parents feedback on their children‘s language development.

The analyses include a hierarchical regression analysis summarizing all related factors. One has to be careful, however, with the interpretation of the results with regard to the different measurement methods. While children‘s cognitive abilities were measured with separate tests, the values both for observation competency and for well-being and involvement were derived by observation of children in the same scene. The model with both types of factors in it may overestimate the relation with the affective factors or underestimate the relation with the cognitive factors.

For the evaluation of the training materials, the trainers filled out feedback forms after each session. Additionally to the three items per child, they wrote down feedback on the individual activities. Thereby, several categories were derived by qualitative analysis. While we found many important indications for potential problems and ideas for solving them, this approach also has its downsides. There was some missing data as not every trainer always gave feedback on each activity, and not every trainer gave feedback in all of the categories. However, by this procedure it was ensured that the trainers could give all the feedback they assumed relevant and not only on primarily set items. In future studies, the now found categories could be formed into items in order to get more systematic and complete feedback on all of the activities in all categories.

Overall, it can be said that this study had both strengths and limitations. Specifically the sample size restrains the scope of the results. At the same time, the study connected several research fields, exploring new questions, which makes the results valuable. The limitations could be overcome by new research, building on and extending the results of the present study.

6.3. Implications

Although observation competency is a central scientific method, research on its structure is still rather scarce. This study is the first looking at the role of developmental factors and the interplay of domain-general and domain-specific factors in the functioning of observation competency. Consequently, the results are exploratory and preliminary, and further research has to be conducted to reinforce the effects and improve the knowledge about the development of observation competency. For practice, this study already provides some important results, which will be discussed after the implications for further research.

6.3.1. Implications for Further Research

This study has shown the effect of scientific reasoning abilities and domain-specific biology understanding on observation competency for preschool children. However, no conclusions can be made for other age groups. Is there always a constant relation with both factors, or are they of different relevance for different age groups? The domain-general scientific reasoning measurement in this study was basic, as it tested whether children understand the logic of inquiry at all. It is possible that there is a minimum requirement and as soon as children have crossed this threshold, domain-general factors become less important. The other possibility is that the comprehension of epistemic activities always helps to make systematic scientific observations. More research is needed here to understand the interplay of domain-specific knowledge and domain-general reasoning over time and development.

Similarly, this study looked particularly at the scientific method observation, which has specific requirements on the person using it (Wellnitz & Mayer, 2011). However, many aspects of observation competency are general inquiry skills, e.g., hypothesizing and interpreting observations, so it could be assumed that the results are similar for other methods. More research could help find out how generalizable the results are.

This study concentrated on cognitive and affective factors of the children themselves. However, there are also many environmental factors that have proven to have an impact on children‘s competencies. Both structural factors of the preschool institution in general and the specific interactions of the preschool teacher have shown to have an impact (Kuger & Kluczniok, 2009), as well as the family environment (Niklas, 2015). For literacy development, the importance of the home learning environment is already well researched (Niklas & Schneider, 2013). Little to nothing is known about the impact the home might have on children‘s scientific literacy. As a lot of learning still happens in the family in preschool age, this might be an interesting facet to look at in more detail.

The training study did not show the desired effects, at least not for all children. However, we found the interesting longitudinal relation between executive functions and observation competency and biology understanding. This result has not been hypothesized beforehand and can therefore only be interpreted as exploratory results with a limited validity. However, the result is in line with prior research and theoretically plausible. Thus, it seems worthwhile to look into this relation more deeply. Ideally, a longitudinal study would measure children‘s executive functions and their biological competencies several times over the span of the last preschool year and the first grade of primary school. Thereby, the impact of executive functions on biological competencies could be investigated more deeply: is it a better predictor of biological competencies at the same time point or is there a developmental gap? How does the change of the learning environment (preschool to school) change the relationship between executive functions and biological competencies? Such a study could also help decide whether it is feasible to train observation competency already in preschool or if the improvements in that age are still dependent on development.

Since there were also some difficulties with the training itself, there are several potential improvements in follow-up studies. Firstly, more games for some competency levels and dimensions are needed. The evaluation of the materials showed that there is a need for more activities fostering children‘s interpretation skills on an easy level. For the dimension of inquiry, more activities are needed in general.

Secondly, executive functions have shown to be relevant for children‘s development of observation competency. One idea would be to not only train children‘s observation competency, but also their executive functions in order to facilitate their conceptual change. The observation training could be combined with an already existing executive functions training, like the red light purple light intervention, which also uses variations of already known games for preschoolers to foster their executive functions and self-regulation (Schmitt et al., 2015; Tominey & McClelland, 2011). Interaction effects of the combination of the two trainings could be investigated with a study design comparing four groups: a control group, an observation training only group, a self-regulation training only group, and a combined training of observation and self-regulation group. The results could inform about the specific effects both have on the development of biological understanding and scientific skills, as well as the added value of a combination. If executive functions play a role in the development of observation competency, the training of executive functions alone should already show effects. If the self-regulation skills are only needed for the active participation in the observation training, only the combined training should be effective. In both cases, larger effects of the combined training than of the observation training alone would be expected.

A third constraint with the intervention was that it was conducted once a week by a trainer coming to the preschool. From a practical point of view, this is not representative of how children usually learn in preschool. One long-term goal would be to train preschool teachers to do the training and find ways to integrate the activities into the daily routine or react to spontaneous ideas of children. As preschool teachers know their children very well and could tailor the activities to children‘s current level of ability and concentration, this could lead to better differentiation and therefore better learning on the children‘s side. On the other hand, the teachers‘ performance would also have a strong influence on the effects of the training and would be difficult to control. This also leads to a completely new field of research questions, as we know little about preschool teachers‘ knowledge, skills, and attitudes in the field of science, and how these are related to their science teaching. For elementary teachers, it is known that they often have low self-efficacy in science teaching, related to lack of knowledge in science (Bleicher & Lindgren, 2005). Similar or even worse results could be expected for preschool teachers, as science is only a limited part of their professional training. As self-efficacy and knowledge are both known to have an impact on science teaching (Guo, Piasta, Justice, & Kaderavek, 2010), this would be relevant factors to investigate in order to understand preschool teachers‘ ability to foster children‘s observation competency. After finding out which cognitive and motivational factors play a role in science teaching in preschool, these factors could be controlled for in an intervention study where the preschool teachers conduct the activities themselves.

Additionally to investigating science education interventions in preschool, new research could also focus on learning biology in non-formal educational institutions, such as museums or zoos. Eberbach and Crowley (2005) showed that in a museum exhibition the type of representation or model plays a role in children‘s learning process, but also parents‘ scaffolding of children‘s learning processes. With a closer focus on observation competency, studies could look into the role epistemic activities play throughout a learning process in a museum, what kind of prior knowledge is important, and how adults help children notice relevant details.

6.3.2. Implications for Practice

One goal of this study was to find ways to foster children‘s scientific skills in preschool. The results show several aspects that are important to keep in mind when doing science activities with children in preschool.

When we want to foster children‘s science skills, there will always be an interplay between their domain general scientific reasoning and their domain-specific understanding, and both is needed for good performance in inquiry situations, and therefore for expanding both their knowledge and reasoning skills. Thus, when we do science with children, we have to keep both in mind and activate both in children.

This study also emphasizes the importance of motivational factors for children‘s learning. It is not enough to provide children with materials and instruments, but it has to be ensured that their emotional well-being is sufficient so that they can be involved in the situation. Only then can they observe adequately and learn more about the topic. Prior research has shown that both instructional support and the emotional tone of the teacher are relevant for children‘s engagement Aydo an et al. (2015 . This also became eminent in this study: in the feedback comments of the trainers, children‘s motivation and their verbalizing were linked. They needed a stimulating and motivating learning situation in order to be able to verbalize thoughts and describe observations. The comments also showed that for being motivated, children needed the task to be neither too hard nor too easy. It is crucial to meet children‘s individual level of competency and support them in their learning process in order to foster the development of their competencies adequately.

Our intervention with game materials to foster observation skills showed positive results, but only for children who already had basic skills at the beginning. At least for children who are already capable, it does seem to make sense to do these game-based activities with them in order to enhance their observation skills. The child‘s individual competency level has to be considered in order to reach good results.

The materials to train children‘s observation competency were successful overall, though some changes and adaptations could enhance effects: firstly, more activities for training children‘s inquiry competency are needed, as they proved to be especially effective. At the same time, the activities for the dimension of interpretation need to be adapted to children‘s performance level in order to show good usability. As discussed above, the training could be combined with tasks from a self-regulation training, as executive functions proved to be crucial for children‘s development of scientific observation skills. Finally, the ideal way of learning in preschool is not program-based, but either project-based or ―en-passant” (Oerter, 2012). For these reasons, it would be desirable that preschool teachers take up our materials for their everyday program. As they know the children best, they could adapt the tasks even better to children‘s level of competency.

The BIKE study by the institute for early education in Bavaria (IFP) investigated interaction quality in Bavarian preschool institutions. While the emotional and organizational support of the preschool teachers is high, the learning support is rather low. Only support of language development is sufficient, cognitive support, in contrast, is low (Wertfein, Wirts, & Wildgruber, 2015). When differentiating between situations in the preschool, it shows that the learning support is lower during free-play phases than during moderated activities (Wirts, Wertfein, & Wildgruber, 2016). The materials developed for this study open new possibilities for cognitive stimulation throughout both moderated activities and free-play phases. In general, the activities from the training are not so different from activities that already exist in preschool anyway. Playing Memory, planting a plant – this is nothing new for both children and teachers. It is the affordances of the material and how the process is accompanied by the preschool teachers that make a difference. Are the children aware of the research question? Are they encouraged to form hypotheses before beginning the activity? Is there a discussion about results that do not match the hypothesis, or when some children arrive at different conclusions? Are the children stimulated to describe details exactly? In this manner, observation skills can already be promoted playfully in preschool, and when the children later come into contact with experiments in school, it will be easier for them to test theories critically.

6.4. Conclusion

In the introduction, I discussed the relevance of science education in general and its specific role in preschool education. The already existing arguments for science education in preschool were that children are both able to and interested in learning about science and that the early introduction to the topic helps them later during their school career. I will now take a look at these arguments under the light of the results of this study.

This study confirms the idea that preschool children show first abilities for domain-general scientific reasoning and observation competency. At the same time, this study also shows that these skills depend on more general cognitive skills like language abilities or executive functions. At the same time, the children really were highly motivated to learn about animals and investigate their behavior, as their involvement in both the observation situation and the training showed. Children who had some basic skills in the beginning showed improvement in their observation competency. Observation competency as a basic scientific method can be seen as a prerequisite for learning more sophisticated inquiry methods, and it can be expected that higher observation competency does help children with later science classes, though this study did not investigate this.

Overall, this study confirms that science education in preschool – in this case a training of observation competency – can be beneficial, as long as it takes children‘s specific developmental levels into account. This study showed how important differentiation was for the success of the intervention activities, and how children‘s concentration span played a major role for the duration of their active participation. Consequently, it does not seem feasible to have science education in preschool in the style of school lessons. Instead, co-constructivist activities, in which teacher and children develop their ideas together, and where the children are stimulated according to their competency level, look much more promising. When science is taught in this way, preschool teachers do not need to feel uncomfortable with their own science knowledge and skills: it is encouraged that they find out things together with the children, and maybe learn new things themselves on the way. The lesson that knowledge can be gained by investigation is crucial for children‘s learning about inquiry. It is probably the one that will stick with them throughout further science education.

Literature

Ainley, M. (2006). Connecting with learning: Motivation, affect and cognition in interest processes. *Educational Psychology Review*, *18*(4), 391–405. doi:10.1007/s10648-006-9033-0

Alberdi, E., Sleeman, D. H. & Korpi, M. (2000). Accommodating surprise in taxonomic tasks: The role of expertise. *Cognitive Science*, *24*(1), 53–91. doi:10.1207/s15516709cog2401_2

Anders, Y., Ballaschk, I. & Tietze, W. (2014). *Wissenschaftliche Untersuchungen zur Arbeit der Stiftung „Haus der kleinen Forscher "* (Vol. 6). Retrieved from http://www.haus-der-kleinen-forscher.de/de/wissenschaftliche-begleitung/ergebnisse-publikationen/schriftenreihe/

Anders, Y. & Roßbach, H. G. (2013). Frühkindliche Bildungsforschung in Deutschland. In Stamm, Margrit and Edelmann, Doris (Ed.), *Handbuch frühkindliche Bildungsforschung* (pp. 183–195). Wiesbaden: Springer VS.

Andersson, K. & Gullberg, A. (2012). What is science in preschool and what do teachers have to know to empower children? *Cultural Studies of Science Education*, 1–22. doi:10.1007/s11422-012-9439-6

Astington, J. W. (2000). *Wie Kinder das Denken entdecken*. München: Reinhardt.

Aydo a n, C., Farran, D. C. & Sa s z , G. (2015 . The relationship between kindergarten classroom environment and children's engagement. *European Early Childhood Education Research Journal*, *23*(5), 604–618. doi:10.1080/1350293X.2015.1104036

Bleicher, R. E. & Lindgren, J. (2005). Success in science learning and preservice science teaching self-efficacy. *Journal of Science Teacher Education*, *16*(3), 205–225. doi: 10.1007/s10972-005-4861-1

Boone, W. J., Staver, J. R. & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Berlin, Heidelberg: Springer.

Bortz, J. & Döring, N. (2013 . *Forschungsmethoden und Evaluation*. Berlin, Heidelberg: Springer-Verlag.

Bradley, M. M. & Lang, P. J. (2000). Emotion and motivation. In Cacioppo, John T. and Tassinary, Louis G. and Berntson, Gary (Ed.), *Handbook of psychophysiology (2nd edition)* (pp. 602–642). Cambridge: Cambridge University Press.

Bullock, M., Sodian, B. & Koerber, S. (2009). Doing experiments and understanding science: Development of scientific reasoning from childhood to adulthood. In Schneider, Wolfgang and Bullock, Merry (Ed.), *Human development from early childhood to early adulthood. Findings from the Munich Longitudinal Study* (pp. 173–197). Mahwah, NJ: Erlbaum.

Bullock, M. & Ziegler, A. (1999). Scientific reasoning: Developmental and individual differences. In Weinert , Franz E. and Schneider, Wolfgang (Ed.), *Individual development from 3 to 12: Findings from the Munich Longitudinal Study* (pp. 38–60). Cambridge: Cambridge University Press.

Bybee, R. W. (2002). Scientific inquiry, student learning, and the science curriculum. In Bybee, Rodger W (Ed.), *Learning science and the science of learning* (pp. 25–35). Arlington, VA: NSTA.

Bybee, R. W. (2006). Scientific inquiry and science teaching. In Flick, Lawrence B. and Lederman, Norman G. (Ed.), *Scientific inquiry and nature of science* (pp. 1–14). Dordrecht, the Netherlands: Springer. doi:10.1007/978-1-4020-5814-1_1

Camilli, G., Vargas, S., Ryan, S. & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*, *112*(3), 579–620.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, Massachusetts: MIT press.

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. doi:10.1007/978-1-4020-5814-1_1

Connor, C. M., Morrison, F. J. & Slominski, L. (2006). Preschool instruction and children‘s emergent literacy growth. *Journal of Educational Psychology*, *98*(4), 665. doi:10.1037/0022-0663.98.4.665

Croker, S. & Buchanan, H. (2011). Scientific reasoning in a real-world context: The effect of prior belief and outcome on children‘s hypothesis-testing strategies. *British Journal of Developmental Psychology*, *29*(3), 409–424. doi:10.1348/026151010X496906

Csikszentmihalyi, M. & Csikszentmihalyi, I. S. (1992). *Optimal experience: Psychological studies of flow in consciousness*. Cambridge university press.

Cumming, J. (2003). Do runner beans really make you run fast? Young children learning about science-related food concepts in informal settings. *Research in Science Education*, *33*(4), 483–501. doi:10.1023/B:RISE.0000005254.53876.6e

Darwin, C. (1887). *The Autobiography of Charles Darwin*. New York: Barnes & Noble Publishing.

DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching*, *37*(6), 582–601. doi:10.1023/B:RISE.0000005254.53876.6e

Deckelmann, N. & Neuhaus, B. J. (2014). Basiskonzepte und problemorientierte Kontexte im Heimat- und Sachunterricht der Grundschule. *Erkenntnisweg Biologiedidaktik*, 105–117.

Decker, S. L., Hill, S. K. & Dean, R. S. (2007). Evidence of construct similarity in Executive Functions and Fluid Reasoning abilities. *The International Journal of Neuroscience*, *117*(6), 735–48.

Declercq, B. (2014). Levels of well-being and involvement of young children in centre-based provision in the Free State Province of South Africa. *South African Journal of Childhood Education*, *1*(2), 64–80. doi:10.4102/sajce.v1i2.85

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. doi:10.1146/annurev-psych-113011-143750

Diamond, A., Barnett, W. S., Thomas, J. & Munro, S. (2007). Preschool program improves cognitive control. *Science (New York, NY)*, *318*(5855), 1387–1388. doi:10.1126/science.1151148

Diskowski, D. (2009). Bildungspläne für Kindertagesstätten—ein neues und noch unbegriffenes Steuerungsinstrument. In Roßbach, Hans-Günther and Blossfeld, Hans-Peter (Ed.), *Frühpädagogische Förderung in Institutionen* (pp. 47–61). Wiesbaden: Springer.

Dräger, M. & Vogt, H. (2007). Von Angst und Ekel zu Interesse. *Erkenntnisweg Biologiedidaktik*, *6*, 133–149.

Duggan, E. C. & Garcia-Barrera, M. A. (2015). Executive functioning and intelligence. In Goldstein, Sam and Princiotta, Dana and Naglieri, Jack A. (Ed.), *Handbook of intelligence* (pp. 435–458). New York: Springer. doi:10.1007/978-1-4939-1562-0_27

Duindam, T., Konak, Ö. & Kamphuis, F. (2010). *Sprachtest - Wissenschaftlicher Bericht*. Butzbach: Cito Deutschland GmbH.

Dunbar, K. & Klahr, D. (1989). Developmental differences in scientific discovery processes. In Klahr, David and Kotovsky, K (Ed.), *Complex information processing: The impact of Herbert A. Simon* (pp. 109–143). Hillsdale, NJ: Erlbaum.

Eberbach, C. & Crowley, K. (2005). From living to virtual: Learning from museum objects. *Curator: The Museum Journal*, *48*(3), 317–338. doi:10.1111/j.2151-6952.2005.tb00175.x

Eberbach, C. & Crowley, K. (2009). From everyday to scientific observation: How children learn to observe the biologist‘s world. *Review of Educational Research*, *79*(1), 39–68. doi:10.3102/0034654308325899

Ekman, P. (1973). Cross-cultural studies of facial expression. In *Darwin and facial expression: A century of research in review* (pp. 169–222). Los Altos, CA: Malor Books.

Evans, E. M. (2008). Conceptual change and evolutionary biology: A developmental analysis. In Vosniadou, Stella (Ed.), *International handbook of research on conceptual change* (pp. 263–294). New York, NY: Routledge.

Fabes, R. A., Eisenberg, N., Nyman, M. & Michealieu, Q. (1991). Young children‘s appraisals of others‘ spontaneous emotional reactions. *Developmental Psychology*, *27*(5), 858–866. doi:10.1037/0012-1649.27.5.858

Feldman Barrett, L. & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*(4), 967–984. doi:10.1037/0022-3514.74.4.967

Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., … others. (2014 . Scientific reasoning and argumentation: Advancing an interdisciplinary research agenda in education. *Frontline Learning Research*, *2*(3), 28–45. doi:10.14786/flr.v2i2.96

Fleer, M. (2013). Affective Imagination in Science Education: Determining the Emotional Nature of Scientific and Technological Learning of Young Children. *Research in Science Education*, *43*(5), 2085–2106. doi:10.1007/s11165-012-9344-8

Forgas, J. P. & Vargas, P. T. (2000). The effects of mood on social judgment and reasoning. In Lewis, Michael and Haviland-Jones, Jeanette M. (Ed.), *Handbook of emotions (2nd edition)* (pp. 350–368). New York, NY: Guilford Press.

Fox, J. E. & Lee, J. (2013 . When Children Draw vs When Children Don‘t: Exploring the Effects of Observational Drawing in Science. *Creative Education*, *4*(7A1), 11. doi:10.4236/ce.2013.47A1002

Fox, N. A. & Calkins, S. D. (1993). Multiple-measure approaches to the study of infant emotion. In Lewis, Michael and Haviland, Jeannette M. (Ed.), *Handbook of emotions* (pp. 167–184). New York, NY: Guilford Press.

Fredricks, J. A., Blumenfeld, P. C. & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59–109. doi: 10.3102/00346543074001059

Fröbel, F. (2012). *The education of man*. Mineola, NY: Dover Publication.

Fthenakis, W. E. (2009). Ko-Konstruktion: Lernen durch Zusammenarbeit. *Kinderzeit*, *3*, 8–13.

Gardner, P. L. (1998 . The development of males‘ and females‘ interests in science and technology. In Hoffmann, Lore and Krapp Andreas and Renninger K. Ann and Baumert, Jürgen (Ed.), *Interest and learning* (pp. 41–57). University of Kiel.

Gathercole, S. E., Pickering, S. J., Knight, C. & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, *18*(1), 1–16. doi:10.1002/acp.934

Geary, D. C. (2008). An evolutionarily informed education science. *Educational Psychologist*, *43*(4), 179–195. doi:10.1080/00461520802392133

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences*, *8*(9), 404–409. doi:10.1016/j.tics.2004.07.001

Gentner, D. & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development*, *14*(4), 487–513. doi:10.1016/S0885-2014(99)00016-7

Glauert, E. (2010). Erkundungen und Erklärungen zur Elektrizität. Zum Sachverstehen und Sachlernen im Vorschulalter. In Fischer, Hans-Joachim and Gansen, Peter and Michalik, Kerstin (Ed.), *Sachunterricht und frühe Bildung* (pp. 123–138). Bad Heilbrunn: Klinkhardt.

Goldberg, R. F. & Thompson-Schill, S. L. (2009 . Developmental ―roots" in mature biological knowledge. *Psychological Science*, *20*(4), 480–487. doi:10.1111/j.1467-9280.2009.02320.x

Goldspink, C., Winter, P. & Foster, M. (2008). Student engagement and quality pedagogy. In *European Conference on Educational Research in Goteborg* (pp. 10–12). Retrieved from http://www.learningtolearn.sa.edu.au/tfel/files/links/Student_Engagement_and_Qu_1.pdf

Gopnik, A. (1996). The post-Piaget era. *Psychological Science*, *7*(4), 221–225. Retrieved from http://www.jstor.org/stable/40062949

Grell, F. (2010). Über die (Un- Möglichkeit, Früherziehung durch Selbstbildung zu ersetzen. *Zeitschrift Für Pädagogik*, *56*(2), 154–167.

Grell, F. (2013). Frühkindliche Bildung in historischer Perspektive. In Stamm, Margrit and Edelmann, Doris (Ed.), *Handbuch frühkindliche Bildungsforschung* (pp. 147–164). Wiesbaden: Springer.

Guo, Y., Piasta, S. B., Justice, L. M. & Kaderavek, J. N. (2010). Relations among preschool teachers' self-efficacy, classroom quality, and children's language and literacy gains. *Teaching and Teacher Education*, *26*(4), 1094–1103. doi:10.1016/j.tate.2009.11.005

Hardy, I., Jonen, A., Möller, K. & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of― floating and sinking." *Journal of Educational Psychology*, *98*(2), 307. doi:10.1037/0022-0663.98.2.307

Hardy, I., Kleickmann, T., Koerber, S., Mayer, D., Möller, K., Pollmeier, J., … Sodian, B. (2010). Die Modellierung naturwissenschaftlicher Kompetenz im Grundschulalter. Projekt Science-P. In Klieme, Eckhard and Leutner, Detlev and Kenk, Martina (Ed.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (pp. 115–125).

Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In Schweizer, Karl (Ed.), *Leistung und Leistungsdiagnostik* (pp. 127–143). Berlin Heidelberg: Springer. doi:10.1007/3-540-33020-8

Hasselhorn, M. & Schneider, W. (2016). Förderung schulrelevanter Kompetenzen in Vorschule und Schule. In Hasselhorn, Marcus and Schneider, Wolfgang (Ed.), *Förderprogramme für Vor-und Grundschule* (Vol. 14, pp. 1–11). Göttingen: Hogrefe Verlag. doi:10.1026/02772-000

Hauser, B., Vogt, F., Stebler, R. & Rechsteiner, K. (2014 . Förderung früher mathematischer Kompetenzen. *Frühe Bildung*, *3*(3), 139–145. doi:10.1026/2191-9186/a000144

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling. Retrieved from http://www.afhayes.com/public/process2012.pdf

Hecker, J. & Tansaway, T. (2008). Das Haus der kleinen Forscher. *Spektrum Der Wissenschaft*, (12), 106.

Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B. & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, *317*(5843), 1360–1366.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, *16*(3), 235–266. doi:1040-726X/04/0900-0235/0

Holodynski, M. (2005). Am Anfang war der Ausdruck. Meilensteine und Mechanismen der Emotionsentwicklung. *Psychologie in Erziehung Und Unterricht*, *52*(4), 229–249.

Inagaki, K. & Hatano, G. (1993 . Young Children‘s Understanding of the Mind-Body Distinction. *Child Development*, *64*(5), 1534–1549. doi:10.2307/1131551

Inagaki, K. & Hatano, G. (1996 . Young children‘s recognition of commonalities between animals and plants. *Child Development*, *67*(6), 2823–2840. doi:10.1111/j.1467-8624.1996.tb01890.x

Inagaki, K. & Hatano, G. (2004 . Vitalistic causality in young children‘s naive biology. *Trends in Cognitive Sciences*, *8*(8), 356–362. doi:10.1016/j.tics.2004.06.004

Izard, C. E. & Ackerman, B. P. (2000). Motivational, organizational, and regulatory functions of discrete emotions. In Lewis, Michael and Haviland-Jones, Jeanette M. (Ed.), *Handbook of emotions (2nd edition)* (pp. 253–264). New York, NY: Guilford Press.

Jirout, J. & Klahr, D. (2012 . Children‘s scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review*, *32*(2), 125–160. doi:10.1016/j.dr.2012.04.002

Johnson, S. C. & Solomon, G. E. (1997). Why dogs have puppies and cats have kittens: The role of birth in young children‘s understanding of biological origins. *Child Development*, *68*(3), 404–419. doi:10.2307/1131668

Johnston, J. S. (2009). What does the skill of observation look like in young children? *International Journal of Science Education*, *31*(18), 2511–2525. doi:10.1080/09500690802644637

Jonen, A., Möller, K. & Hardy, I. (2003). Lernen als Veränderung von Konzepten - am Beispiel einer Untersuchung zum naturwissenschaftlichen Lernen in der Grundschule. In Cech, Diethard and Schwier, Hans-Joachim (Ed.), *Lernwege und Aneignungsformen im Sachunterricht* (pp. 93–108). Bad Heilbrunn: Klinkhardt.

Kelly, K. M. (2013). *Science Journals in the Garden: Developing the Skill of Observation in Elementary Age Students*. Master Thesis, Portland State University. Retrieved from http://pdxscholar.library.pdx.edu/open_access_etds/1536/

Keltner, D., Ekman, P., Gonzaga, G. C. & Beer, J. (2003). Facial expression of emotion. In Davidson, Richard J. and Scherer, Klaus R. and Goldsmith, H. Hill (Ed.), *Handbook of affective sciences* (pp. 415–432). New York, NY: Oxford University Press.

Kind, P. M. & Osborne, J. (2016). Styles of scientific reasoning - a cultural rationale for science education? *Science Education*, *in press*.

Klahr, D. (2002). *Exploring Science*. Cambridge, MA: MIT press.

Koerber, S., Mayer, D., Osterhaus, C., Schwippert, K. & Sodian, B. (2015). The Development of Scientific Thinking in Elementary School: A Comprehensive Inventory. *Child Development*, *86*(1), 327–336. doi:10.1111/cdev.12298

Koerber, S., Sodian, B., Thoermer, C. & Nett, U. (2005). Scientific reasoning in young children: Preschoolers‘ ability to evaluate covariation evidence. *Swiss Journal of Psychology*, *64*(3), 141. doi:10.1024/1421-0185.64.3.141

Kohlhauf, L. (2013). *Spielend biologisch beobachten: Messung und Förderung biologischer Beobachtungskompetenz im Vorschulalter*. Berlin: Logos-Verlag.

Kohlhauf, L., Rutke, U. & Neuhaus, B. (2011). Influence of previous knowledge, language skills and domain-specific interest on observation competency. *Journal of Science Education and Technology*, *20*(5), 667–678. doi:10.1007/s10956-011-9322-3

Kosso, P. (2011). *A summary of scientific method*. Dordrecht: Springer Science \& Business Media. doi:10.1007/978-94-007-1614-8

Krajewski, K., Renner, A., Nieding, G. & Schneider, W. (2009 . Frühe Förderung von mathematischen Kompetenzen im Vorschulalter. In Roßbach, Hans-Günther and Blossfeld, Hans-Peter (Ed.), *Frühpädagogische Förderung in Institutionen* (pp. 91–103). Wiesbaden: Springer.

Kristen, S., Thoermer, C., Hofer, T., Aschersleben, G. & Sodian, B. (2006). Skalierung von —Theory of Mind —Aufgaben. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *38*(4), 186–195. doi:10.1026/0049-8637.38.4.186

Kuger, D.-P. S. & Kluczniok, D.-P. K. (2009). Prozessqualität im Kindergarten—Konzept, Umsetzung und Befunde. In Roßbach, Hans-Günther and Blossfeld, Hans-Peter (Ed.), *Frühpädagogische Förderung in Institutionen* (pp. 159–178). Wiesbaden: Springer.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, *96*(4), 674–89. doi:10.1037/0033-295X.96.4.674

Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher*, *28*(2), 16–46.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, *9*(5), 178–181. doi:10.1111/1467-8721.00088

Kuhn, D. (2011). What is scientific thinking and how does it develop? In *The Wiley-Blackwell handbook of childhood cognitive development (2nd ed.)* (pp. 497 – 523). Wiley-Blackwell. Retrieved from http://search.ebscohost.com.emedien.ub.uni-muenchen.de/login.aspx?direct=true&db=psyh&AN=2010-22950-019&site=ehost-live

Kuhn, D. & Franklin, S. (2006). The second decade. What develops (and how). In Kuhn, Deanna and Siegler, R. S. (Ed.), *Handbook of child psychology: Vol. 2. Cognition, perception and language* (pp. 953–993). Hoboken, NJ: Wiley.

Laevers, F. (1993). Deep level learning: an exemplary application on the area of physical knowledge. *European Early Childhood Education Research Journal*, *1*(1), 53–68. doi:10.1080/13502939385207351

Laevers, F. (2000). Forward to Basics! Deep-Level-Learning and the Experiential Approach. *Early Years*, *20*(2), 20–29. doi:10.1080/0957514000200203

Laevers, F. & Heylen, L. (2003). *Involvement of children and teacher style. Insights from an International Study on Experiental Education*. Leuven: Leuven University Press.

Laevers, F., Kog, M. & Vandenbussche, E. (1997). A process-oreinted child monitoring system for young children.

Laukenmann, M., Bleicher, M., Fuss, S., Gläser-Zikuda, M., Mayring, P. & Rhöneck, C. v. (2000). Eine Untersuchung zum Einfluss emotionaler Faktoren auf das Lernen im Physikunterricht. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *6*, 139–155.

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L. & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners‘ conceptions of nature of science. *Journal of Research in Science Teaching*, *39*(6), 497–521. doi:10.1002/tea.10034

Lehrer, R. & Schauble, L. (2006). Scientific thinking and science literacy. In Renninger, K. Ann and Sigel, Irving E. and Damon, William and Lehrer, Richard (Ed.), *Handbook of child psychology (6th edition)* (Vol. Four: Child Psychology in Practice). Hoboken, NJ: Wiley.

Leuchter, M., Saalbach, H. & Hardy, I. (2010). Die Gestaltung von Aufgaben in den ersten Bildungsjahren. In Leuchter, Miriam (Ed.), *Didaktik für die ersten Bildungsjahre* (pp. 98 – 111). Seelze, Deutschland: Klett.

Leuchter, M., Saalbach, H. & Hardy, I. (2014). Designing Science Learning in the First Years of Schooling. An intervention study with sequenced learning material on the topic of ―floating and sinking." *International Journal of Science Education*, *36*(10), 1751–1771. doi:10.1080/09500693.2013.878482

Linacre, J. M. (2006). WINSTEPS Rasch measurement computer program. Retrieved from http://www.winsteps.com/

Linnenbrink, E. A. (2007). The role of affect in student learning: A multi-dimensional approach to considering the interaction of affect, motivation, and engagement. In Schutz, Paul A. and Pekrun, Reinhard (Ed.), *Emotion in education* (pp. 107–124). San Diego, CA: Elsevier.

Liu, D. W. C. (2012). Science Denial and the Science Classroom. *CBE Life Sciences Education*, *11*(2), 129–134. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3366896/

Lohmann, H. & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development*, *74*(4), 1130–1144.

Lommen, C. M. (2012). *How does the use of digital photography affect student observation skills and data collection during outdoor field studies?* Master Thesis, Montana State University. Retrieved from http://scholarworks.montana.edu/xmlui/handle/1/1747

Lück, G. (2004). Naturwissenschaften im frühen Kindesalter. In Fthenakis, Wassilios E. and Oberhuemer, Pamela (Ed.), *Frühpädagogik International* (pp. 331–343). Wiesbaden: Springer.

Lück, G. (2013 . Förderung naturwissenschaftlicher Bildung. In Stamm, Margrit and Edelmann, Doris (Ed.), *Handbuch frühkindliche Bildungsforschung* (pp. 557–572). Wiesbaden: Springer.

Lück, G. (2015). *Handbuch der naturwissenschaftlichen Bildung: Theorie und Praxis für die Arbeit in Kindertageseinrichtungen*. Freiburg, Deutschland: Herder.

Lück, G. & Demuth, R. (1998). Naturwissenschaften im frühen Kindesalter. *ChemKon*, *5*(2), 71–78. doi:10.1002/ckon.19980050204

Mähler, C. & Ahrens, S. (2003). Naive Biologie im kindlichen Denken. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *35*(3), 153–162. doi:10.1026//0049-8637.35.3.153

Matson, G. & Cline, T. (2012). The impact of specific language impairment on performance in science and suggested implications for pedagogy. *Child Language Teaching and Therapy*, *28*(1), 25–37.

Mayer, D., Sodian, B., Koerber, S. & Schwippert, K. (2014). Scientific reasoning in elementary school children: Assessment and relations with cognitive abilities. *Learning and Instruction*, *29*, 43–55. doi:10.1016/j.learninstruc.2013.07.005

Mercer, N., Dawes, L., Wegerif, R. & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Research Journal*, *30*(3), 359–377.

Michalik, K. (2010). Didaktische Konzepte für die naturwissenschaftliche Grundbildung von Kindern im Elementarbereich. In Fischer, Hans-Joachim and Gansen, Peter and Michalik, Kerstin (Ed.), *Sachunterricht und frühe Bildung* (pp. 93–107). Bad Heilbrunn: Klinkhardt.

Mogel, H. (2008 . Das Kinderspiel: Motor der Persönlichkeitsentwicklung und Lebensbewältigung. In *Psychologie des Kinderspiels: Von den frühesten Spielen bis zum Computerspiel* (pp. 153–211). Heidelberg: Springer. doi:10.1007/978-3-540-46644-4_5

Möller, K. & Steffensky, M. (2010). Naturwissenschaftliches Lernen im Unterricht mit 4- bis 8-jährigen Kindern. In Leuchter, Miriam (Ed.), *Didaktik für die ersten Bildungsjahre* (pp. 162 –178). Seelze, Deutschland: Klett.

Monteira, S. F. & Jiménez-Aleixandre, M. P. (2015). The practice of using evidence in kindergarten: The role of purposeful observation. *Journal of Research in Science Teaching*, *53*(8), 1232–1258. doi:10.1002/tea.21259

Moors, A., Ellsworth, P. C., Scherer, K. R. & Frijda, N. H. (2013). Appraisal theories of emotion: State of the art and future development. *Emotion Review*, *5*(2), 119–124. doi:10.1177/1754073912468165

Murray, N., Sujan, H., Hirt, E. R. & Sujan, M. (1990). The influence of mood on categorization: A cognitive flexibility interpretation. *Journal of Personality and Social Psychology*, *59*(3), 411. doi:10.1037/0022-3514.59.3.411

Möller, K. (2002 . Anspruchsvolles Lernen in der Grundschule - am Beispiel naturwissenschaftlich-technischer Inhalte. *Pädagogische Rundschau*, *56*(4), 411–435.

Möller, K., Hardy, I., Jonen, A., Kleickmann, T. & Blumberg, E. (2006 . Naturwissenschaften in der Primarstufe Zur Förderung konzeptuellen Verständnisses durch Unterricht. In *Untersuchungen zur Bildungsqualität von Schule* (pp. 161–193). Münster, Deutschland: Waxmann.

Niklas, F. (2015). Die familiäre Lernumwelt und ihre Bedeutung für die kindliche Kompetenzentwicklung. *Psychologie in Erziehung Und Unterricht*, *62*(2), 106–120. doi:10.2378/peu2015.art11d

Niklas, F. & Schneider, W. (2013). Home literacy environment and the beginning of reading and spelling. *Contemporary Educational Psychology*, *38*(1), 40–50. doi:10.1016/j.cedpsych.2012.10.001

Nores, M. & Barnett, W. S. (2013). Early Child Development Programs and Research in the United States. In Stamm, Margrit and Edelmann, Doris (Ed.), *Handbuch frühkindliche Bildungsforschung* (pp. 223–237). Wiesbaden: Springer.

Norris, S. P. (1984). Defining observational competence. *Science Education*, *68*(2), 129–142. doi:10.1002/sce.3730680206

Oerter, R. (2012). Lernen en passant: Wie und warum Kinder spielend lernen. *Diskurs Kindheits-Und Jugendforschung*, *7*(4). Retrieved from http://nbn-resolving.de/urn:nbn:de:0168-ssoar-390184

Oguz, A. & Yurumezoglu, K. (2007). *The Primacy of Observation in Inquiry-Based Science Teaching*. Paper presented at the International Workshop: Science Education in School. Retrieved from http://eric.ed.gov/?id=ED498653

Opfer, J. E. & Gelman, S. A. (2010). Development of the animate-inanimate distinction. In Goswami, Usha (Ed.), *The Wiley-Blackwell Handbook of Childhood Cognitive Development (2nd edition)* (pp. 213–238). Chichester, United Kingdom: Wiley.

Osterhaus, C., Koerber, S. & Sodian, B. (2015 . Children‘s Understanding of Experimental Contrast and Experimental Control: An Inventory for Primary School. *Frontline Learning Research*, *3*(4), 56–94. doi:10.14786/flr.v3i4.220

Osterhaus, C., Koerber, S. & Sodian, B. (2016). Scientific Thinking in Elementary School: Children‘s Social Cognition and Their. *Manuscript Submitted for Publication*.

Pascal, C., Bertram, T., Mould, C. & Hall, R. (1998). Exploring the relationship between process and outcome in young children‘s learning: stage one of a longitudinal study. *International Journal of Educational Research*, *29*(1), 51–67. doi:10.1016/S0278-4343(98)00013-5

Pauen, S. (1997). Überlebt der Animismus? Kritische Evaluation einer Hypothese zum präkausalen Denken. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *29*, 97–118.

Pauen, S. (2009). *Vom Klein-Sein zum Einstein. Abschlussbericht der wissenschaftlichen Begleitung*. Ludwigshafen: Heinrich Pesch Haus, Kath. Akademie Rhein-Neckar.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341. doi:10.1007/s10648-006-9029-9

Pekrun, R. & Bühner, M. (2014). Self-report measures of academic emotions. In Pekrun, Reinhard and Linnenbrink-Garcia, Lisa (Ed.), *International Handbook of Emotions in Education* (pp. 561–579). New York, NY: Routledge.

Pekrun, R., Hall, N. C., Goetz, T. & Perry, R. P. (2014). Boredom and academic achievement: Testing a model of reciprocal causation. *Journal of Educational Psychology*, *106*(3), 696. doi:10.1037/a0036006

Pekrun, R. & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In Christenson, Sandra L. and Reschly, Amy L. and Wylie, Cathy (Ed.), *Handbook of research on student engagement* (pp. 259–282). New York, NY: Springer.

Pekrun, R. & Perry, R. P. (2014). Control-value theory of achievement emotions. In Pekrun, Reinhard and Linnenbrink-Garcia, Lisa (Ed.), *International Handbook of Emotions in Education* (pp. 120–141). New York, NY: Routledge.

Pekrun, R. & Stephens, E. J. (2010). Achievement emotions: A control-value approach. *Social and Personality Psychology Compass*, *4*(4), 238–255. doi:10.1111/j.1751-9004.2010.00259.x

Penner, D. E. & Klahr, D. (1996). The Interaction of Domain-Specific Knowledge and Domain-General Discovery Strategies: A Study with Sinking Objects. *Child Development*, *67*(6), 2709–2727. doi:10.1111/j.1467-8624.1996.tb01884.x

Perner, J. (1991). *Understanding the representational mind.* Cambridge, MA: The MIT Press.

Piaget, J. (1978). *Das Weltbild des Kindes*. Stuttgart, Deutschland: Klett-Cotta.

Piaget, J. (1983). Piaget's theory. In Mussen, P (Ed. , *Handbook of child psychology. 4th ed.* (Vol. 1). New York: Wiley.

Piekny, J., Grube, D. & Maehler, C. (2013a). The Development of Experimentation and Evidence Evaluation Skills at Preschool Age. *International Journal of Science Education*, *2*(36), 1–21. doi:10.1080/09500693.2013.776192

Piekny, J., Grube, D. & Maehler, C. (2013b . The relation between preschool children's false-belief understanding and domain-general experimentation skills. *Metacognition and Learning*, *8*(2), 103–119. doi:10.1007/s11409-013-9097-4

Piekny, J. & Maehler, C. (2013). Scientific reasoning in early and middle childhood: The development of domain-general evidence evaluation, experimentation, and hypothesis generation skills. *British Journal of Developmental Psychology*, *31*(2), 153–179. doi:10.1111/j.2044-835X.2012.02082.x

Reich, K. (2001). Konstruktivistische Ansätze in den Sozial-und Kulturwissenschaften. In Hug, T. (Ed.), *Wie kommt die Wissenschaft zu ihrem Wissen?* (Vol. 4). Baltmannsweiler: Schneider.

Renkl, A., Mandl, H. & Gruber, H. (1996). Inert knowledge: Analyses and remedies. *Educational Psychologist*, *31*(2), 115–121. doi:10.1207/s15326985ep3102_3

Richards, D. D. & Siegler, R. S. (1986 . Children's understandings of the attributes of life. *Journal of Experimental Child Psychology*, *42*(1), 1–22. doi:10.1016/0022-0965(86)90013-5

Rogoff, B., Paradise, R., Arauz, R. M., Correa-Chávez, M. & Angelillo, C. (2003). Firsthand learning through intent participation. *Annual Review of Psychology*, *54*(1), 175–203. doi:10.1146/annurev.psych.54.101601.145118

Rosenfelder, D. (2006). *Lernwerkstatt Natur – Frühe Wege zur Natur(Wissenschaft) und Technik*. Retrieved from https://www.hf.uni-koeln.de/data/eso/File/Schaefer/Konzeption_LernwerkstattNatur.pdf

Roßbach, H.-G., Kluczniok, K. & Kuger, D.-P. S. (2009). Auswirkungen eines Kindergartenbesuchs auf den kognitiv-leistungsbezogenen Entwicklungsstand von Kindern. In Roßbach, Hans-Günther and Blossfeld, Hans-Peter (Ed.), *Frühpädagogische Förderung in Institutionen* (pp. 139–158). Wiesbaden: Springer.

Ruffman, T., Perner, J., Olson, D. R. & Doherty, M. (1993). Reflecting on Scientific Thinking: Children's Understanding of the Hypothesis-Evidence Relation. *Child Development*, *64*(6), 1617–1636. doi:10.1111/j.1467-8624.1993.tb04203.x

Rutherford, F. J. (1997). *Sputnik and science education*. Symposium ―Reflecting on Sputnik: Linking the Past, Present, and Future of Educational Reform". Washington, DC. Retrieved from http://www.nas.edu/sputnik/ruther1.htm

Ryan, R. M. & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist* , *55*(1), 68–78. doi:10.1037/0003-066X.55.1.68

Sammons, P., Anders, Y., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B. & Barreau, S. (2009 . Children‘s cognitive attainment and progress in English primary schools during Key Stage 2: Investigating the potential continuing influences of pre-school education. In Roßbach, Hans-Günther and Blossfeld, Hans-Peter (Ed.), *Frühpädagogische Förderung in Institutionen* (pp. 179–198). Wiesbaden: Springer.

Schäfer, G. E. (2001). *Prozesse frühkindlicher Bildung*. Retrieved from https://www.hf.uni-koeln.de/data/eso/File/Schaefer/Prozesse_Fruehkindlicher_Bildung.pdf

Schäfer, G. E. (2009). *Natur als Werkstatt*. Kiliansroda, Deutschland: Verlag Das Netz.

Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, *32*(1), 102–119. doi:10.1037/0012-1649.32.1.102

Scherres, C. (2013). Unterrichtskonzepte zur Differenzierung. In *Niveauangemessenes Arbeiten in selbstdifferenzierenden Lernumgebungen* (pp. 21–33). Wiesbaden: Springer. doi:10.1007/978-3-658-02083-5_2

Schmitt, S. A., McClelland, M. M., Tominey, S. L. & Acock, A. C. (2015). Strengthening school readiness for Head Start children: Evaluation of a self-regulation intervention. *Early Childhood Research Quarterly*, *30*, 20–31. doi:10.1016/j.ecresq.2014.08.001

Schraw, G., Crippen, K. J. & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, *36*(1-2), 111–139. doi:10.1007/s11165-005-3917-8

Schüßler, I. (2004). *Lernwirkungen neuer Lernformen*. Arbeitsgemeinschaft Betriebliche Weiterbildungsforschung e.V. Retrieved from http://www.abwf.de/content/main/publik/materialien/materialien55.pdf

Sha, L., Schunn, C., Bathgate, M. & Ben-Eliyahu, A. (2016). Families support their children‘s success in science learning by influencing interest and self-efficacy. *Journal of Research in Science Teaching*, *53*(3), 450–472. doi:10.1002/tea.21251

Sinatra, G. M. & Chinn, C. A. (2012). Thinking and reasoning in science: Promoting epistemic conceptual change. In Harris, Karen R. and Graham, Steve and Urdan, Tim and Bus, Adriana G. and Major, Sonya and Swanson, H. Lee (Ed.), *APA educational psychology handbook, Vol 3: Application to learning and teaching* (pp. 257–282). Washington, DC: American Psychological Association. doi:10.1037/13275-011

Sinatra, G. M. & Taasoobshirazi, G. (2011). Intentional conceptual change: The self-regulation of science learning. In Zimmermann, B., Shunk, D. (Ed.), *Handbook of self-regulation of learning and performance* (pp. 203–216). Routledge.

Snow, C. (1983). Literacy and language: Relationships during the preschool years. *Harvard Educational Review*, *53*(2), 165–189. doi:10.17763/haer.53.2.t6177w39817w2861

Sodian, B., Zaitchik, D. & Carey, S. (1991 . Young children's differentiation of hypothetical beliefs from evidence. *Child Development*, *62*(4), 753–766. doi:10.1111/j.1467-8624.1991.tb01567.x

Spelke, E. S., Phillips, A. & Woodward, A. L. (1995 . Infants' knowledge of object motion and human action. In Sperber, Dan and Premack, David and Premack, Ann James (Ed.), *Causal cognition: A multidisciplinary debate* (pp. 44–78). New York, NY: Clarendon Press/Oxford University Press.

Staatsinstitut für Frühpädagogik München (IFP). (2012). *Der Bayerische Bildungs-und Erziehungsplan für Kinder in Tageseinrichtungen bis zur Einschulung* (5th ed.). Berlin: Cornelsen. Retrieved from http://www. ifp. bayern. de/imperia/md/content/stmas/ifp/bildungsplan. pdf

Stanat, P., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., … Weiß, M. (2002). *PISA 2000: Overview of the Study. Design, Method and Results*. Berlin: Max Planck Institute for Human Development. Retrieved from https://www.mpib-berlin.mpg.de/Pisa/PISA-2000_Overview.pdf

Statistisches Bundesamt. (2016). Ländermonitor Frühkindliche Bildung. Forschungsverbund DJI/TU Dortmund. Retrieved 8, 2016, from http://www.laendermonitor.de/uebersicht-grafiken/indikator-4a-bildungsbeteiligung-von-kindern-in-kindertagesbetreuung/indikator/4/indcat/15/indsubcat/40/index.nc.html

Steffensky, M., Lankes, E.-M., Carstensen, C. H. & Nölke, C. (2012 . Alltagssituationen und Experimente: Was sind geeignete naturwissenschaftliche Lerngelegenheiten für Kindergartenkinder? *Zeitschrift Für Erziehungswissenschaft*, *15*(1), 37–54. doi:10.1007/s11618-012-0262-3

Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J. A., Wagner, R. K., Williams, W. M., … Grigorenko, E. L. (2000). *Practical intelligence in everyday life*. Cambridge, United Kingdom: Cambridge University Press.

Stiftung Haus Der Kleinen Forscher. (2016). *Stiftung Haus der kleinen Forscher - Themen*. Retrieved from http://www.haus-der-kleinen-forscher.de/de/praxisanregungen/experimente-themen/

Strauss, A. & Corbin, J. (1994). Grounded theory methodology. In Denzin, Norman K. and Lincoln, Yvonna S. (Ed.), *Handbook of qualitative research* (pp. 273–285). Thousand Oaks, CA: Sage Publications.

Stutz, M. (2013). Frühes Lesen und Rechnen und ihre Auswirkungen auf die spätere Schullaufbahn: Zur Genese sozialer Ungleichheit im Kontext des Bildungssystems. In Stamm, Margrit and Edelmann, Doris (Ed.), *Handbuch frühkindliche Bildungsforschung* (pp. 695–711). Wiesbaden: Springer.

Tokarczyk, S. (2015). *Using Technology and Phenology to Increase Students' Observation Skills*. Master Thesis, Hamline University, Saint Paul, Minnesota. Retrieved from http://digitalcommons.hamline.edu/hse_all/162/

Tominey, S. L. & McClelland, M. M. (2011). Red light, purple light: Findings from a randomized trial using circle time games to improve behavioral self-regulation in preschool. *Early Education \& Development*, *22*(3), 489–519. doi:10.1080/10409289.2011.574258

Tomkins, S. S. (1962). *Affect imagery consciousness: The positive affects* (Vol. 1). New York, NY: Springer.

Tomkins, S. & Tunnicliffe, S. D. (2007 . Nature tables: stimulating children's. *Journal of Biological Education*, *41*(4), 150–155.

Tunnicliffe, S. D. & Reiss, M. J. (1999). Opportunities for sex education and personal and social education (PSE) through science lessons: the comments of primary pupils when observing meal worms and brine shrimps. *International Journal of Science Education*, *21*(9), 1007–1020. doi:10.1080/095006999290282

Van der Graaf, J., Segers, E. & Verhoeven, L. (2015). Scientific reasoning abilities in kindergarten: dynamic assessment of the control of variables strategy. *Instructional Science*, *43*(3), 381–400. doi:10.1007/s11251-015-9344-y

Van der Graaf, J., Segers, E. & Verhoeven, L. (2016). Scientific reasoning in kindergarten: Cognitive factors in experimentation and evidence evaluation. *Learning and Individual Differences*, *49*, 190–200. doi:10.1016/j.lindif.2016.06.006

Van der Sluis, S., de Jong, P. F. & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*, *35*(5), 427–449.

Van Els, T. & Van Hest, E. (1992). *The Dutch national action programme on foreign languages: Recommendations and policy reactions*. Zoetermeer, The Netherlands: Ministry of Education and Science.

Veenman, M. V. J., Hout-Wolters, B. H. A. M. & Afflerbach, P. (2006). Metacognition and learning: conceptual and methodological considerations. *Metacognition and Learning*, *1*(1), 3–14. doi:10.1007/s11409-006-6893-0

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental process*. Cambridge, MA: Harvard University Press.

Walberg, H. J. & Tsai, S.-L. (1983). Matthew effects in education. *American Educational Research Journal*, *20*(3), 359–373. doi:10.3102/00028312020003359

Wasik, B. A., Bond, M. A. & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology*, *98*(1), 63. doi:10.1037/0022-0663.98.1.63

Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In Rychen, Dominique Simone and Hersh Salganik, Laura (Ed.), *Defining and selecting key competencies* (pp. 45–65). Seattle: Hogrefe & Huber.

Wellman, H. M., Cross, D. & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, *72*(3), 655–684. doi:10.1037/0022-0663.98.1.63

Wellnitz, N., Fischer, H. E., Kauertz, A., Mayer, J., Neumann, I., Pant, H. A., … Walpuski, M. (2012). Evaluation der Bildungsstandards: eine fächerübergreifende Testkonzeption für den Kompetenzbereich Erkenntnisgewinnung. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *18*, 261–291.

Wellnitz, N. & Mayer, J. (2011). *Modelling and assessing scientific methods*. Annual Meeting of the National Association of Research in Science Teaching (NARST). Orlando, Florida. Retrieved from https://www.uni-kassel.de/fb10/fileadmin/datas/fb10/biologie/didaktik/Publikationslisten/Artikelanh%C3%A4nge/Wellnitz_NARST2011.pdf

Wellnitz, N. & Mayer, J. (2012). Beobachten, Vergleichen und Experimentieren: Wege der Erkenntnisgewinnung. In Harms, Ute and Bogner, Franz X. (Ed.), *Lehr- und Lernforschung in der Biologiedidaktik* (Vol. 5, pp. 63–79). Innsbruck: Studienverlag.

Wertfein, M., Wirts, C. & Wildgruber, A. (2015). Bedingungsfaktoren für gelingende Interaktionen zwischen Erzieherinnen und Kindern. *IFP-Projektbericht*, *27*. Retrieved from http://www.ifp.bayern.de/imperia/md/content/stmas/ifp/projektbericht_bike_nr_27.pdf

Wilhelm, P. & Beishuizen, J. (2003). Content effects in self-directed inductive learning. *Learning and Instruction*, *13*(4), 381–402. doi:10.1016/S0959-4752(02)00013-0

Wirts, C., Wertfein, M. & Wildgruber, A. (2016). *Interaktionsqualität in der Kita – Qualitätsdimensionen und Situationen machen den Unterschied*. 50. Kongress der Deutschen Gesellschaft für Psychologie, Berlin, Germany.

Wolf, K. M., Schroeders, U. & Kriegbaum, K. (2016). Metaanalyse zur Wirksamkeit einer Förderung der phonologischen Bewusstheit in der deutschen Sprache. *Zeitschrift Für Pädagogische Psychologie*, *30*(1), 9–33. doi:10.1024/1010-0652/a000165

Woltman, H., Feldstain, A., MacKay, J. C. & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 52–69. doi:10.20982/tqmp.08.1.p052

Zaitchik, D., Iqbal, Y. & Carey, S. (2014). The effect of executive function on biological reasoning in young children: an individual differences study. *Child Development*, *85*(1), 160–175. doi:10.1111/cdev.12145

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review*, *20*(1), 99–149. doi:10.1006/drev.1999.0497

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, *27*(2), 172–223.

List of Tables

List of Figures

Appendix

Appendix A – Communication

- Preschool Information Letter
- Parent Consent Form

Appendix B – Intervention

- Photos of Materials

Appendix C – Tests

- Observation Situation
- Biology Understanding Questionnaire
- Mouse-House-Task Picture Cards and Text
- Cake Task Picture Cards and Text
- Prior Knowledge Question Sheet
- Theory of Mind Question Sheet
- Executive Functions Instructions

Appendix D – Coding Schemes

- Observation Coding Scheme
- Well-Being and Involvement Coding Scheme
- Biology Understanding Coding Scheme
- Training Evaluation Form

Statement of Scientific Integrity

Spielend Biologisches Beobachten Lernen

Konzept:

Im Kindergarten ist im Bereich des biologischen wissenschaftlichen Denkens die Fähigkeit zu beobachten von besonderer Bedeutung. Durch Spiele im Kindergarten und in der Natur soll das Beobachten der Kinder geschärft werden. Eine gute Beobachtungsgabe ist für die Kinder in allen Lebensbereichen von hohem Wert – nicht nur im wissenschaftlichen Kontext, sondern auch im Alltag, beim Beginn der Schule und beim Spielen.

In unserem Projekt wollen wir herausfinden, ob sich das wissenschaftliche Beobachten von Kindern mit einem spielerisch angelegten Training schulen lässt. Hierfür suchen wir Kindergärten, die ein Interesse daran haben, dass Vorschüler an unserem Training teilnehmen. Hierzu kommt eine gut geschulte Trainerin aus unserem Team zu Ihnen ins Haus. Wir würden gerne zur Datenerhebung (siehe unten) am Anfang und am Ende mit zwei Gruppen von Vorschülern kleine Spiele und Interviews durchführen, um mehr darüber herauszufinden, wie Beobachten und wissenschaftliches Denken im Vorschulalter funktioniert. Eine Gruppe Vorschüler nimmt anschließend am Training (siehe unten) teil, die zweite Gruppe Vorschüler dient als Vergleichsgruppe und nimmt nicht am Training teil. Gerne stellen wir aber den Erziehern nach Abschluss unseres Trainings Materialien zur Verfügung, sodass sie auf Wunsch die Spiele und Aktivitäten selbst mit diesen Kindern durchführen können.

Training:

Für das Training sind 12-14 Trainingstermine von je etwa 90 Minuten vorgesehen. Die Trainerin wird mit Ihnen einen guten Termin vereinbaren, an dem sie regelmäßig einmal wöchentlich von etwa Mitte Oktober bis etwa Mitte Januar vorbeikommt und mit den Kindern die von uns vorbereiteten Spiele und Aktivitäten macht. Ein/e Erzieher/in aus Ihrem Team kann die ganze Zeit zugegen sein und die Spiele beobachten und auch gerne mitmachen. Die Kinder erhalten zum Ende eine Urkunde über die Teilnahme an dem Training.

Beispiele für Spiele und Aktivitäten: Quartettspiel mit Vogelarten, Bilder von Tieren malen, Tiere live beobachten und nachahmen, Ameisenfarm, Blumen pflanzen, …

Datenerhebung:

Wir wollen herausfinden, ob die Beobachtungsfähigkeit der Kinder geschult werden konnte und auch, mit welchen anderen Kompetenzen der Kinder die Beobachtungsfähigkeit zusammenhängt. Hierzu würden wir vor und am Ende der Trainingsphase mit den Kindern verschiedene Spiele durchführen, die wir auf Video aufzeichnen, um verschiedene Fähigkeiten zu messen. Die Spiele und

Tests sind altersgerecht konzipiert und machen den Kindern Spaß. Sollte ein Kind einmal keinen Spaß haben, kann es die Testung auf eigenen Wunsch sofort abbrechen. Die Videos werden streng vertraulich behandelt und sind zu jedem Zeitpunkt nur den Mitarbeitern des Projekts zugänglich.

Für die Vorerhebung Ende September planen wir drei Untersuchungstage ein, an denen jedes Kind jeweils etwa eine halbe Stunde mit einem unserer gut geschulten Testleiter Spiele spielt. Bei der Nacherhebung im Januar/Februar werden es nur noch zwei Testtage sein. Die Termine werden rechtzeitig mit Ihnen und Ihrem Team abgesprochen und an Ihren Zeitplan angepasst.

Außerdem würden wir auch Sie, die Sie als Erzieher die Kinder schon sehr gut kennen, bitten, einen kurzen Fragebogen zur Einschätzung der Kinder auszufüllen. Diesen werden wir Ihnen rechtzeitig geben, sodass genügend Zeit bleibt, diesen in Ruhe auszufüllen.

Die Eltern werden von uns selbstverständlich auch ausführlich über das Projekt informiert und um ihre Zustimmung zur Teilnahme Ihrer Kinder gebeten. Auch für die Eltern wird es einen kurzen Fragebogen geben.

Kontakt:

Wenn Sie Nachfragen zu dem Projekt haben, dürfen Sie sich gerne jederzeit an mich wenden:

Janina Klemm
Leopoldstraße 44
Room: 609

80802 München

Tel: 089/2180-72561
Mobil: [redacted]

janina.klemm@psy.lmu.de

Wir würden uns sehr freuen, wenn Ihre Einrichtung an unserem Projekt teilnehmen möchte!

Herzliche Grüße

Janina Klemm

LMU | LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN | DIDAKTIK DER BIOLOGIE | DIDAKTIK DER BIOLOGIE LMU

LMU München · Didaktik der Biologie · FB Biologie
Winzererstr. 45II · 80797 München

Prof. Dr. Birgit J. Neuhaus

Telefon +49 (0)89 2180-6489
Sek. +49 (0)89 2180-6490
Telefax +49 (0)89 2180-6491

birgit.neuhaus@lrz.uni-muenchen.de

http://www.zi.biologie.uni-muenchen.de/institute/idb/

An die Erziehungsberechtigten

des Kindergartens XXX

München, Datum

Sehr geehrte Eltern, sehr geehrte Erziehungsberechtigte,

der Lehrstuhl für Entwicklungspsychologie der LMU München führt derzeit ein Forschungsprojekt durch, in dem die Entwicklung des naturwissenschaftlichen Denkens und Verstehens sowie der dazu notwendigen Grundlagen untersucht wird. Dieses Projekt wird durch das Elitenetzwerk Bayern gefördert.

Im Rahmen des Projektes planen wir derzeit eine Studie über die Fähigkeit von Vorschulkindern, in biologischen Kontexten zu beobachten. Aus früheren Studien wissen wir bereits, dass Vorschulkinder zu ersten wissenschaftlichen Beobachtungen in der Lage sind. Nun wollen wir sehen, ob sich diese Fähigkeit durch ein spielerisches Training im Kindergarten und in der Natur fördern lässt.

Eine gute Beobachtungsgabe ist für die Kinder in allen Lebensbereichen von hohem Wert – nicht nur im wissenschaftlichen Kontext, sondern auch im Alltag, beim Beginn der Schule und beim Spielen. Hierzu kommt eine gut geschulte Trainerin aus unserem Team in die Kindertageseinrichtung und führt im Zeitraum von Anfang Oktober bis Ende Februar wöchentlich eine Spielstunde zum Beobachten durch.

Wir würden gerne zur Datenerhebung am Anfang und am Ende mit den Vorschülern kleine Spiele und Interviews durchführen, um mehr darüber herauszufinden, wie Beobachten und wissenschaftliches Denken im Vorschulalter funktioniert. Eine Gruppe Vorschüler nimmt anschließend am Training (siehe unten) teil, die zweite Gruppe Vorschüler dient als Vergleichsgruppe und nimmt nicht am Training teil. Gerne stellen wir aber den Erziehern nach Abschluss unseres Trainings Materialien zur Verfügung, sodass sie auf Wunsch die Spiele und Aktivitäten selbst mit diesen Kindern durchführen können.

Die erste Erhebung soll zu Schuljahresbeginn im September/Oktober stattfinden. In drei etwa 30-minütigen Interview- und Spielsequenzen werden spielerische Aufgaben zum Beobachten, zum wissenschaftlichen Denken, zum Vorwissen und zu den Sprachfähigkeiten der Kinder erhoben. Bei der zweiten Erhebung im März 2015 werden ähnliche Aufgaben verwendet, allerdings wird es sich hierbei um weniger Spiele und nur noch um zwei Sequenzen handeln. Unsere Erfahrungen mit ähnlichen Studien zeigen, dass die Kinder sehr viel Freude bei der Bearbeitung der geplanten Aufgaben haben, da die Aufgaben spielerisch und altersgemäß angelegt sind. Die Studie wird von erfahrenen, gut ausgebildeten Mitarbeiterinnen der Arbeitsgruppe durchgeführt.

Selbstverständlich führen wir unsere Interviews nur im Einverständnis mit Ihnen und Ihrem Kind durch. Ihr Sohn bzw. Ihre Tochter hat somit jederzeit die Möglichkeit, die Teilnahme abzubrechen. Sie selbst können Ihre Einverständniserklärung jederzeit ohne Angabe von Gründen widerrufen.

Für die Auswertung unserer Untersuchungen werden wir die Interviews mit Ihrem Kind auf Video aufzeichnen. Die erhobenen Daten werden streng vertraulich behandelt und vor der Auswertung so aufbereitet, dass ein Rückschluss auf das jeweilige Kind nicht möglich ist. Alle Daten sind jederzeit nur Mitarbeiterinnen und Mitarbeitern des Projekts zugänglich und werden für keinen anderen Zweck verwendet. Dritte haben keinen Zugriff auf die erhobenen Daten. Wir sind verpflichtet, die Videos zehn Jahre gesichert aufzubewahren, danach werden sie unwiderruflich gelöscht. Sollten Sie nach Aufzeichnung der Videos Ihre Meinung über die Teilnahme Ihres Kindes ändern, können Sie aber jederzeit die Löschung der Daten verlangen.

Wir möchten Sie hiermit darum bitten, einer Teilnahme Ihres Kindes an der Studie zuzustimmen. Bitte füllen Sie dazu die angefügte Erklärung aus und geben Sie sie im Kindergarten ab. Sollten Sie noch Fragen zu unserer Untersuchung haben, so stehen wir Ihnen gerne zur Verfügung. Ihre Ansprechpartnerin ist Dipl.-Psych. Janina Klemm (Kontaktdaten: janina.klemm@psy.lmu.de, Tel. 089/2180-72561).

Wir würden uns über die Teilnahme Ihres Kindes am Projekt sehr freuen!

Mit freundlichen Grüßen

Janina Klem

Dipl.-Psych. Janina Klemm

Einverständniserklärung eines/einer Erziehungsberechtigten

Mein Sohn/meine Tochter

____

O darf

O darf nicht

an der von der Ludwig-Maximilians-Universität durchgeführten Studie teilnehmen.

München, den __________ ______________________________

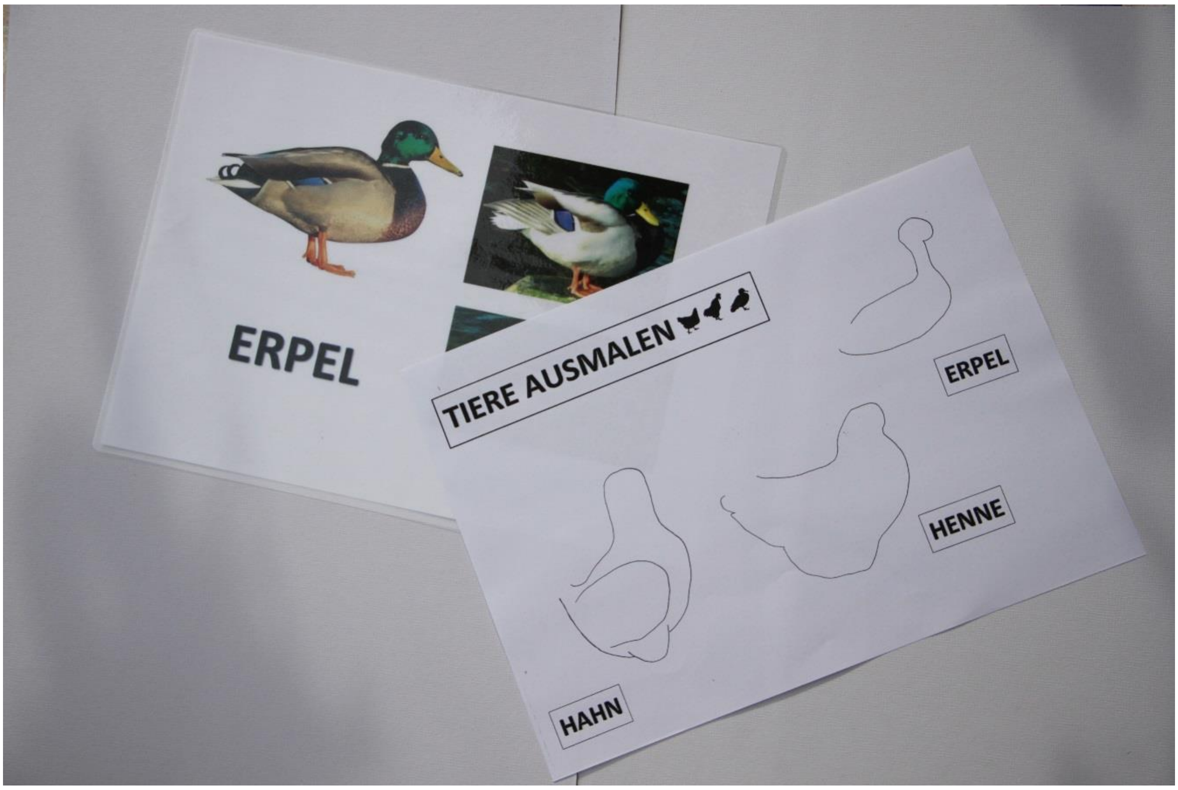
Unterschrift eines/einer Erziehungsberechtigten

Mimikwürfel

Geruchsmemory

Blumenbilder

Vogelmemory



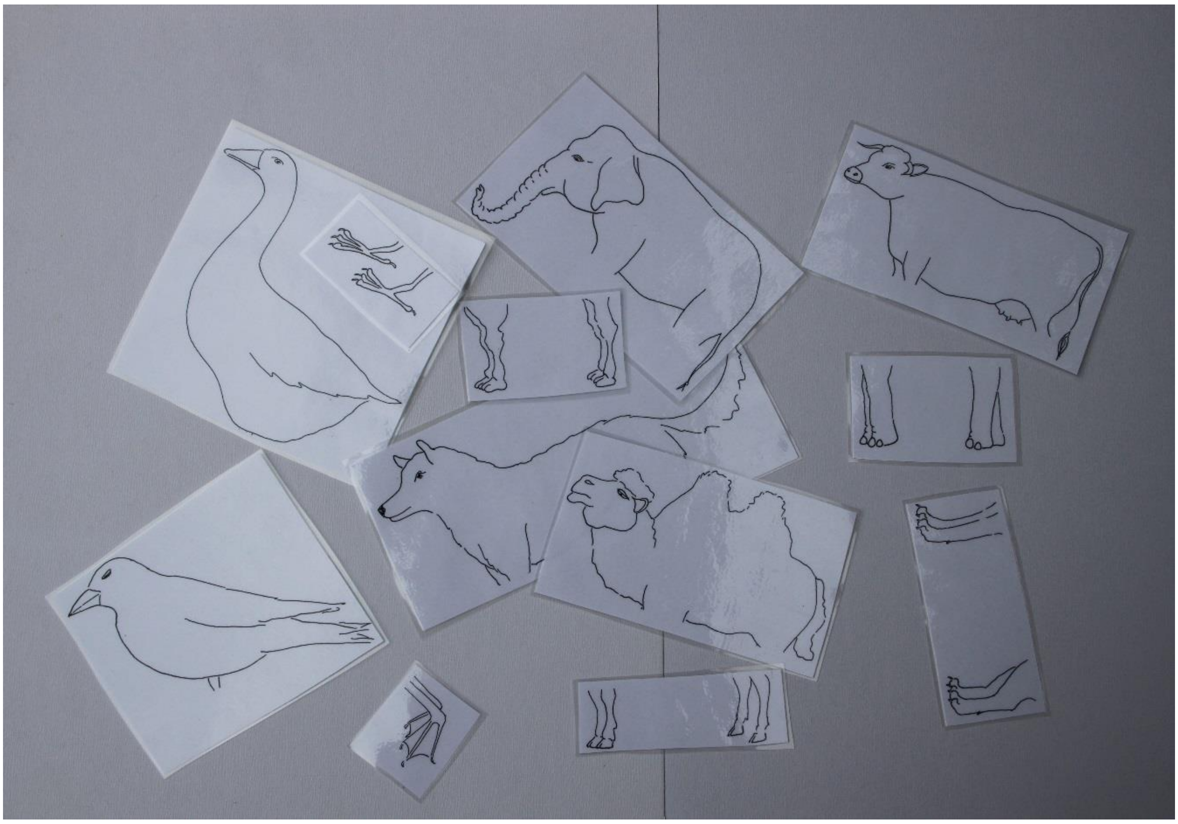
Ergänzungsbilder Vögel


Vogelquartett

Ausschnittbilder

Welches Tier krabbelt hier?

Detailbilder



Ergänzungsbilder Tiere

Spurensuche

Urkunde für Kinder

KINDERFORSCHER

...immer auf der Suche nach dem Unbekannten

Wir überreichen

diese Auszeichnung für ihren Beitrag zur Erforschung des wissenschaftlichen Denkens in den frühen Lebensjahren.

Er/sie hat aktiv am Kurs „Spielend Biologisch Beobachten Lernen“ teilgenommen und erste Beobachtungen zum „Beobachten wie Biologen“ gemeistert.

München, 23.02.2015
Ort, Datum

Janina Klemm
Unterschrift der Versuchsleiterin

Observation Situation

EINFÜHRUNG:

Hallo, ich bin __________, das hier ist der Emil. Wie heißt du?

Emil: Oh, hallo *NAME*! Wollen wir zusammen spielen, Komm mal mit, die ______ und ich haben ganz spannende Sachen mitgebracht! Wir wollen jetzt mal Biologen sein!

Ja, genau, Emil. Weißt du, was ein Biologe ist? Das ist jemand der sich mit Tieren beschäftigt. Wollen wir das jetzt auch mal machen?

EINFÜHRUNG TESTINSTRUMENTE (Lineal, Lupe, Waage, Thermometer)

BEOBACHTUNG TIERE (Fische, Schnecken, Maus):

Schau mal, weißt du was das ist? Und der Emil hält sich jetzt mal die Augen zu. Kannst du ihm *TIER* beschreiben? Kannst du beschreiben was du beobachtest?

Super. So und jetzt sind wir also Biologen und beobachten *TIER*. Wie gehst du da vor, was machst du?

Falls keine selbstständige Fragestellung:

Hast du eine Idee, was du über *TIER* herausfinden willst?

Falls keine Antwort:

Ah der Emil hat eine Idee:

EMIL: Ohja, wie wäre es mit/ ich wollte schon immer wissen….

Falls keine selbstständige Hypothese:

Mhm, und was glaubst du denn was die Antwort sein könnte?

Falls keine selbstständige Testung:

Okay, und hast du eine Idee wie wir das rausfinden können?

Falls keine selbstständige Idee:

Hat der Emil eine Idee?

EMIL: Ja, wir könnten…

Falls keine selbstständige Beschreibung:

Was kannst du beobachten?

Falls keine selbstständige Interpretation:

Was hast du jetzt herausgefunden?

Stimmt das, was wir uns am Anfang gedacht haben?

Super, vielen Dank dass du mitgemacht hast.

Mögliche Forschungsfragen

FISCH

Wie geht der nicht unter?

Welche Flosse benutzt er zum Antreiben/ sich bewegen

Wie atmet der Fisch? Muss er auftauchen?

SCHNECKE

Wie bewegt sie sich fort?

Wie schnell ist sie?

Was frisst sie?

Wie viele Fühler hat sie?

Gewicht Haus und Körper

Spiralenform vom Haus (kreise?)

Welche Schnecke ist schwerer?

MAUS

Wann hat sie die Augen zu? Beim Graben?

Wo ist die Maus mehr, im Tunnel oder oben?

Was macht die Maus mit dem Ast?

Schwanzlänge (so lang wie Körper?)

Schnurhaare

Biology Understanding Questionaire

Name Kind: ________________ Testleiter:__________________ Nummer:________

Q1: Was bedeutet es, am Leben zu sein?

Q2: Kannst Du einige Dinge nennen, die am Leben sind, die Lebewesen sind?

Q3: Kannst Du einige Dinge nennen, die nicht lebendig / am Leben sind?

Q4: Urteile: Ja / Nein-Entscheidungen: Ist ein x am Leben?

Objekt		Ja	nein
a	Berg		
b	Glocke		
c	Sonne		
d	**Wind**		
e	Tisch		
f	Fliege		
g	Auto		
h	Feuer		
i	**Katze**		
j	Bleistift		
k	Vogel		
l	**Baum**		
m	Schlange		
n	Fahrrad		
o	Uhr		
s	Blume		
q	Flugzeug		
r	**Lampe**		
s	Wolke		
t	Regen		

Q5: Begründungen

a Du hast gesagt, dass eine Katze lebendig / nicht lebendig ist. Woher weißt du das?

b Du hast gesagt, dass ein Baum lebendig / nicht lebendig ist. Woher weißt du das?

c Du hast gesagt, dass der Wind lebendig / nicht lebendig ist. Woher weißt du das?

d Du hast gesagt, dass eine Lampe lebendig / nicht lebendig ist. Woher weißt du das?

Cake Task Picture Cards and Text

Das sind Lars und Tom. Sie sind Geschwister. Lars und Tom haben eine Maus im Keller. Die Maus kommt nur nachts raus. Lars und Tom wissen, dass es nur eine Maus ist. Sie wissen nicht, ob die Maus groß oder klein. Lars denkt, die Maus ist klein. Tom denkt, die Maus ist groß.

Lars und Tom wollen die Maus mit Käse füttern. Sie nehmen zwei kleine Häuser, die keine Fallen sind. Jedes Haus hat einen Eingang. Die Eingänge sind unterschiedlich groß. Das eine Haus hat einen großen Eingang. Das andere Haus hat einen kleinen Eingang. Wenn die Maus durch den Eingang passt, kann sie den Käse essen und wieder verschwinden.

Kontrollfragen:
Passt eine große Maus durch den großen Eingang?
Passt eine kleine Maus durch den großen Eingang?
Passt eine große Maus durch den kleinen Eingang?
Passt eine kleine Maus durch den kleinen Eingang?

Effekt:
Lars und Tom wollen den Käse für die Maus in eines der Häuser legen. Denk daran, sie wissen nicht, ob die Maus groß oder klein ist. Welches Haus sollten sie nehmen, damit die Maus den Käse essen kann, egal ob sie groß oder klein ist?

Das mit dem kleinen Eingang
Das Haus mit dem großen Eingang

Warum?

Find out:
Wie können sie herausfinden, ob die Maus groß oder klein ist? ((Offene Frage))

Am nächsten Tag wollen Lars und Tom herausfinden, ob die Maus nun groß oder klein ist. Sie legen Käse in eines der Häuser. Wenn der Käse aus dem Haus verschwunden ist, hat die Maus durch die Öffnung gepasst und den Käse gegessen.

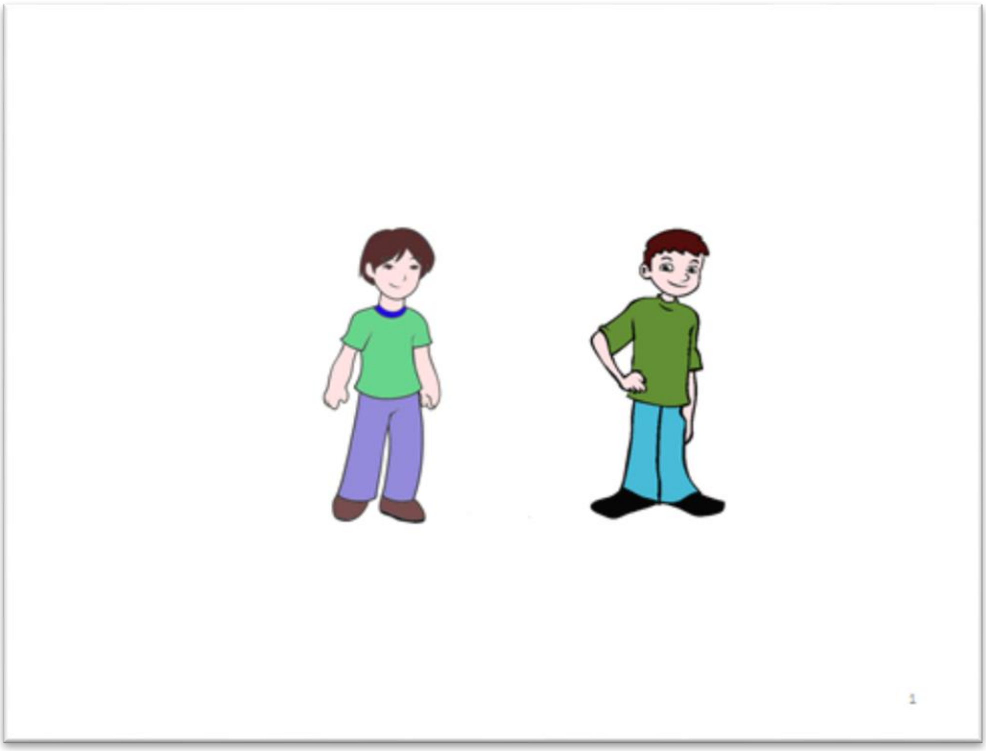
In welches von den beiden Häusern sollten sie den Käse tun, wenn sie herausfinden wollen, ob die Maus groß oder klein ist? In das

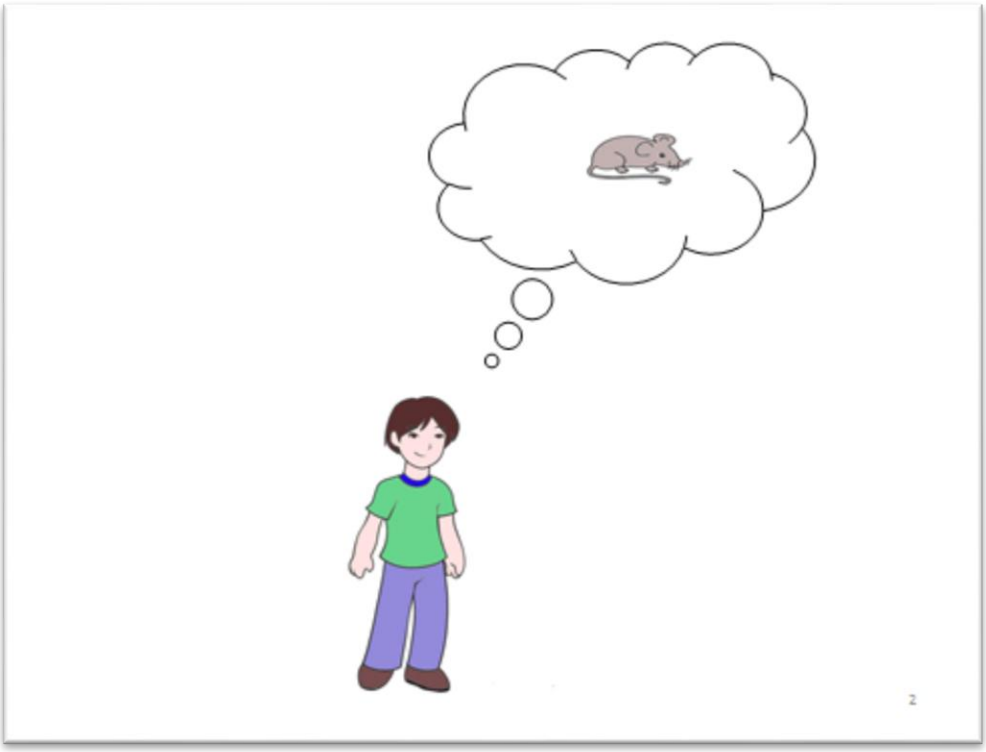
Haus mit dem kleinen Eingang oder
Haus mit dem großen Eingang

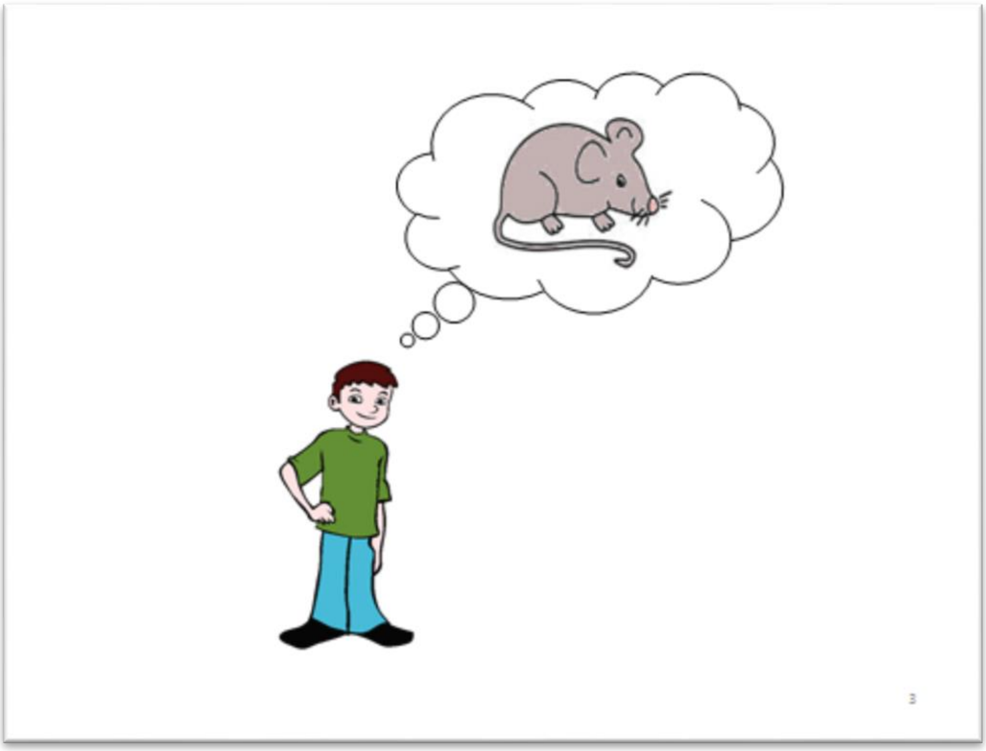
Warum?

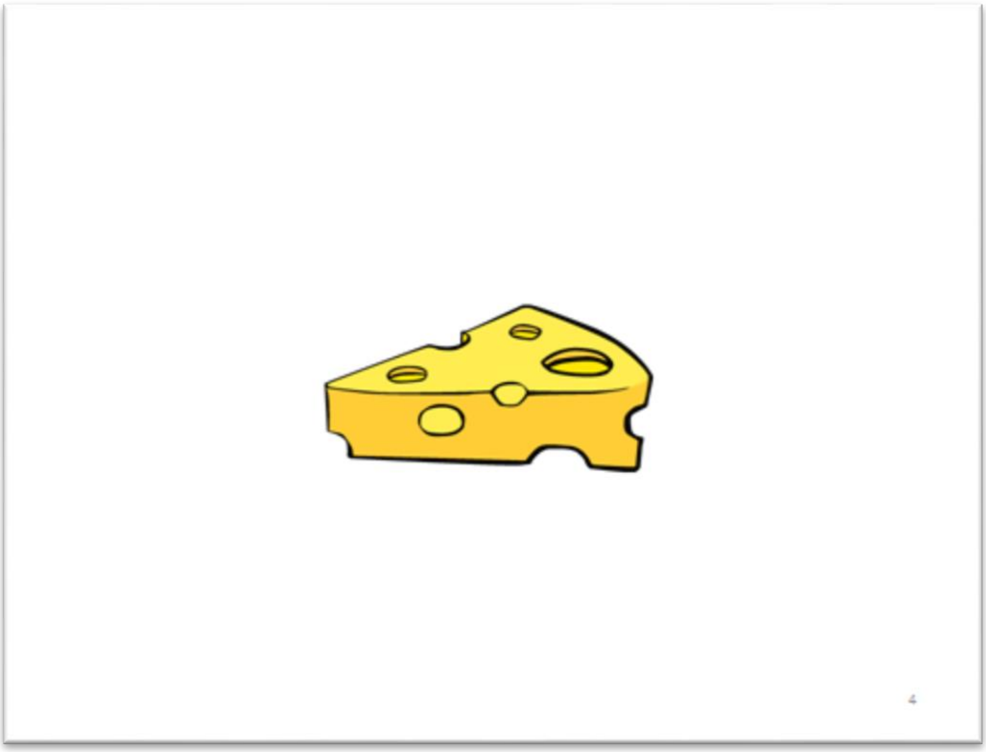
Posthoc:
Lars und Tom nehmen das Haus mit dem großen Eingang und legen den Käse hinein. Am nächsten Morgen ist der Käse weg.

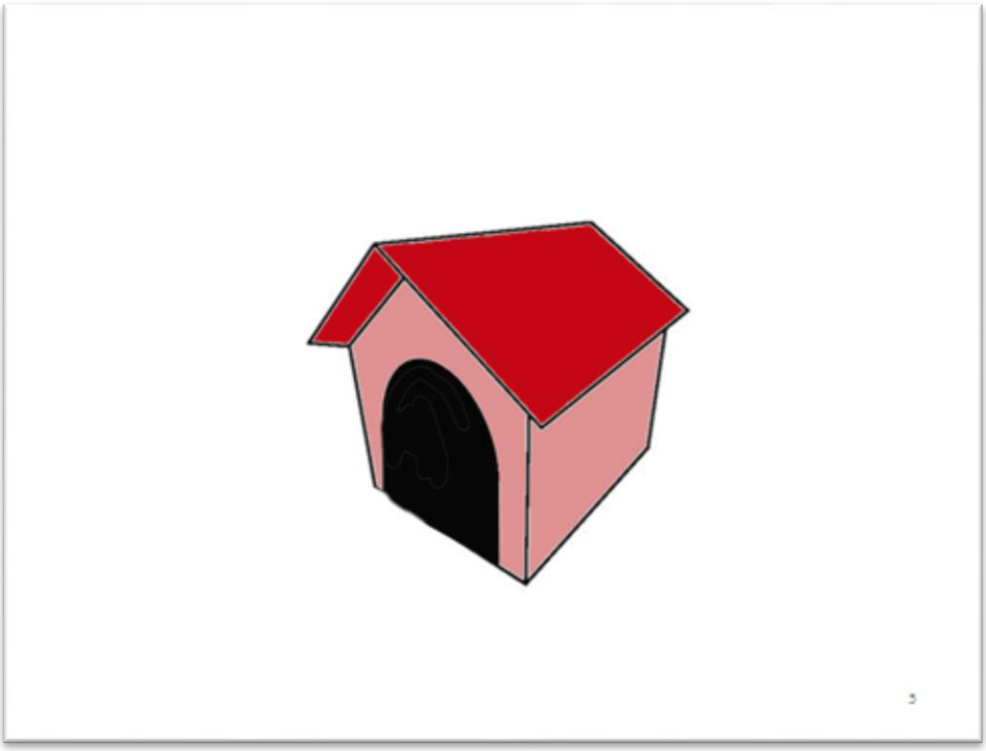
Wissen Tom und Lars nun, ob die Maus groß oder klein ist?
Warum? Warum nicht?

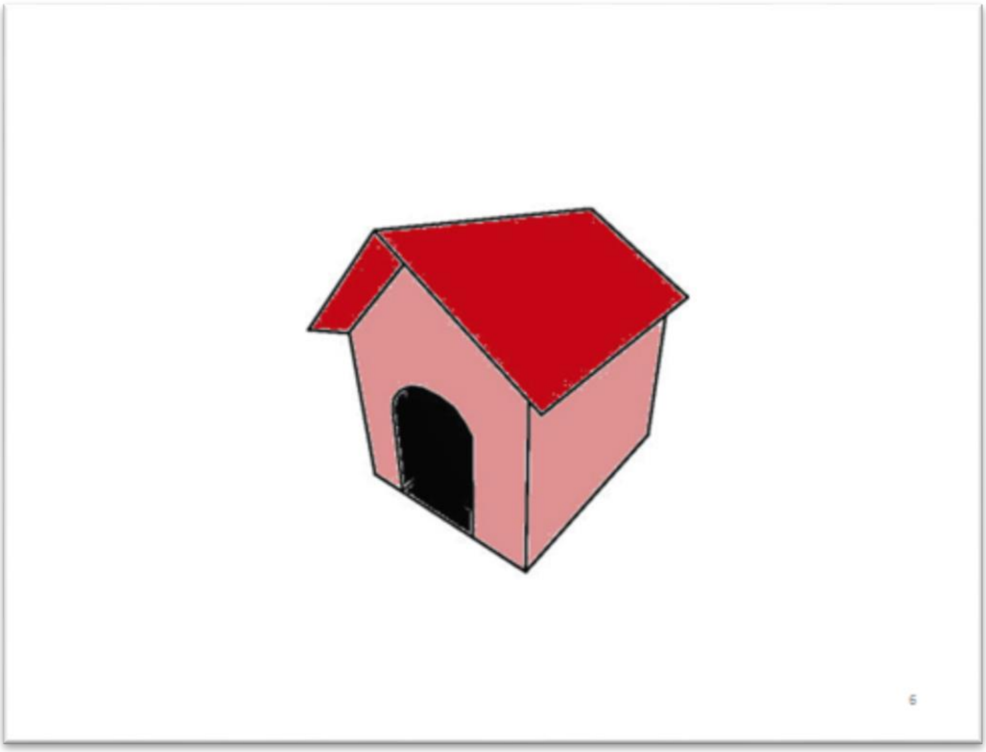

1


2


3


4


5


6

Cake Task Picture Cards and Text

In der Geschichte geht es um eine Familie mit der Mama und den drei Kindern Mia, Lisa und Anna. Die Mama backt einen Kuchen. Sie hat ein Rezept, das sie immer verwendet. Eines Tages geht sie in den Supermarkt und kauft zwei neue Zutaten: KATMUS und GANORE. Sie backt den Kuchen und tut KATMUS und GANORE hinein. Die Familie isst den Kuchen und alle drei Kinder finden den Kuchen total lecker. Sie sagen: „Oh Mama, der neue Kuchen schmeckt so gut! Er schmeckt besser, als die Kuchen, die Du vorher gebacken hast!“ Was macht den Kuchen so lecker? Die Mutter sagt: „ Ich habe zwei neue Zutaten für den Kuchen verwendet aber ich weiß nicht welche von beiden den Kuchen so gut macht“.

Effekt:
Die Mutter möchte den leckeren Kuchen für eine Geburtstagsfeier nochmal backen.
Wie schafft sie es den leckeren Kuchen nochmal genauso gut zu backen?
Warum?

Die drei Kinder der Familie hatten folgende Ideen.
Mia sagt: „ Backen wir den Kuchen mit KATMUS und GANORE zusammen“.
Lisa sagt: „ Backen wir den Kuchen nur mit KATMUS“.
Anna sagt: „ Backen wir den Kuchen in einer viereckigen Form“

Was glaubst Du, wer hat Recht?
Warum?

Find Out:
Die Zutaten sind teuer und die Mutter möchte wissen, welche der beiden Zutaten den Kuchen lecker macht.
Wie kann die Mutter herausfinden welche der beiden Zutaten den Kuchen lecker macht?
Warum?

Die drei Kinder hatten folgende Ideen.
Mia: „Lasst uns zwei Kuchen backen: einen nur mit KATMUS und einen nur mit GANORE“.
Lisa: „Lasst uns zwei Kuchen backen: einen mit KATMUS und GANORE und einen ohne KATMUS und GANORE“.
Anna: „ Lasst uns zwei Kuchen backen: einen in einer eckigen Form und einen in einer runden Form“.

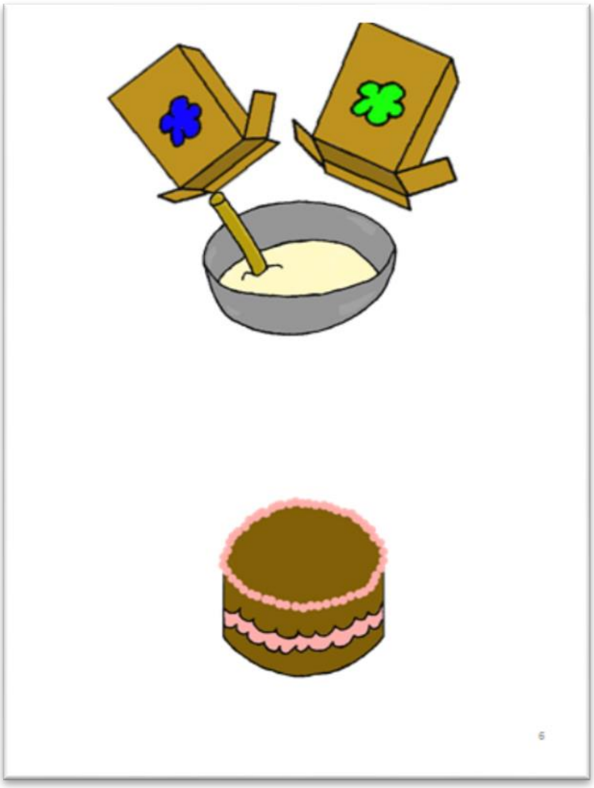
Was glaubst Du, wer hat Recht?
Warum?

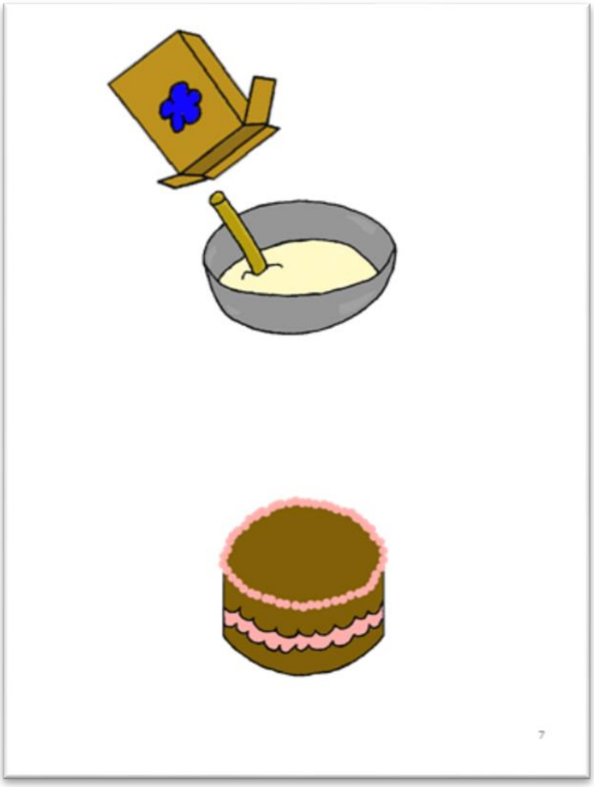
Posthoc:
Die Familie entscheidet sich zwei Kuchen zu backen. Für den einen verwendet sie KATMUS und GANORE, für den zweiten Kuchen verwendet sich keine der beiden. Glaubst Du, sie finden so heraus welche Zutat den Kuchen lecker macht?
Warum?

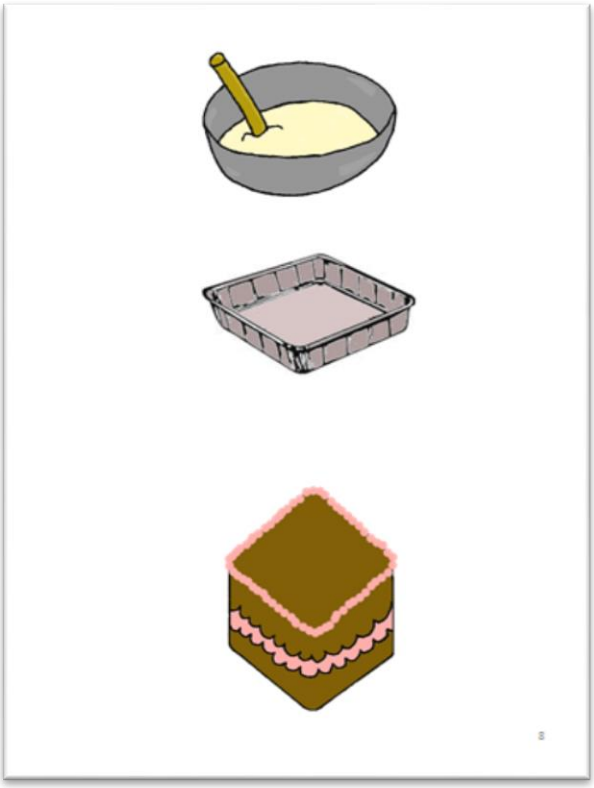

M
L
A
1

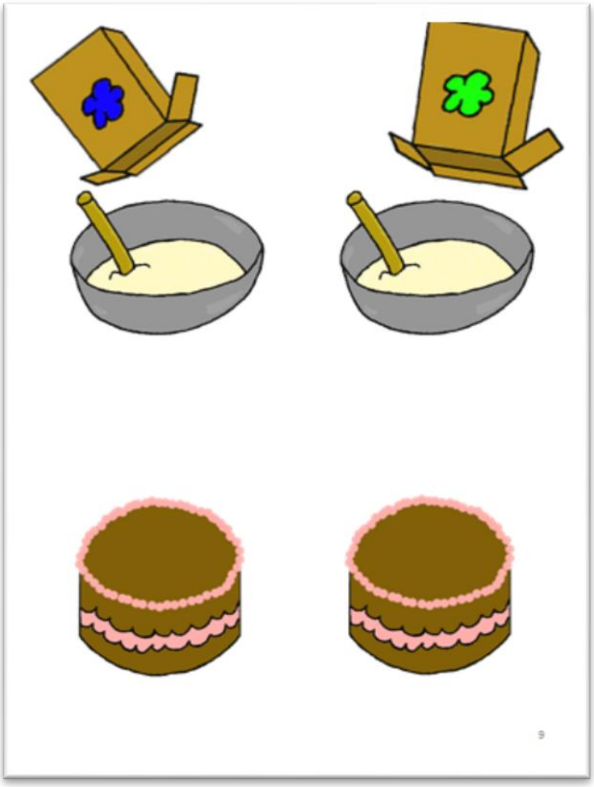

2

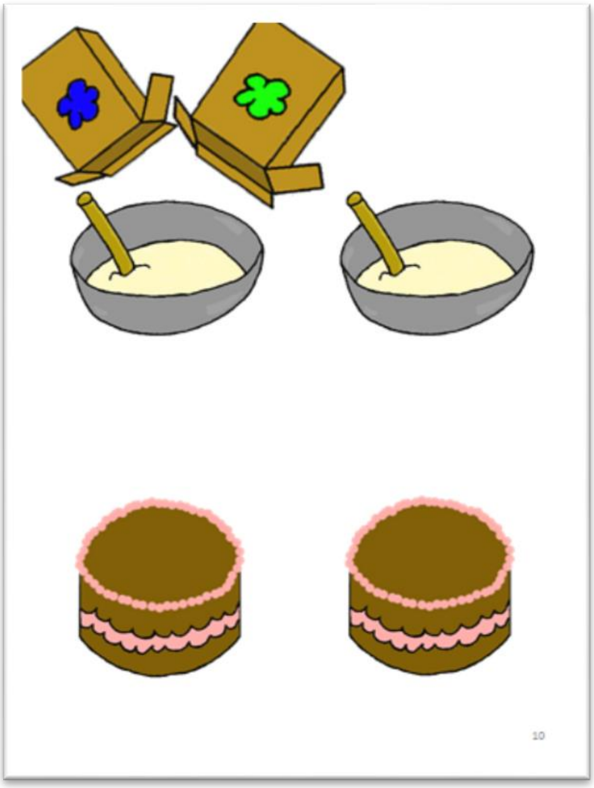

M
L
A
5

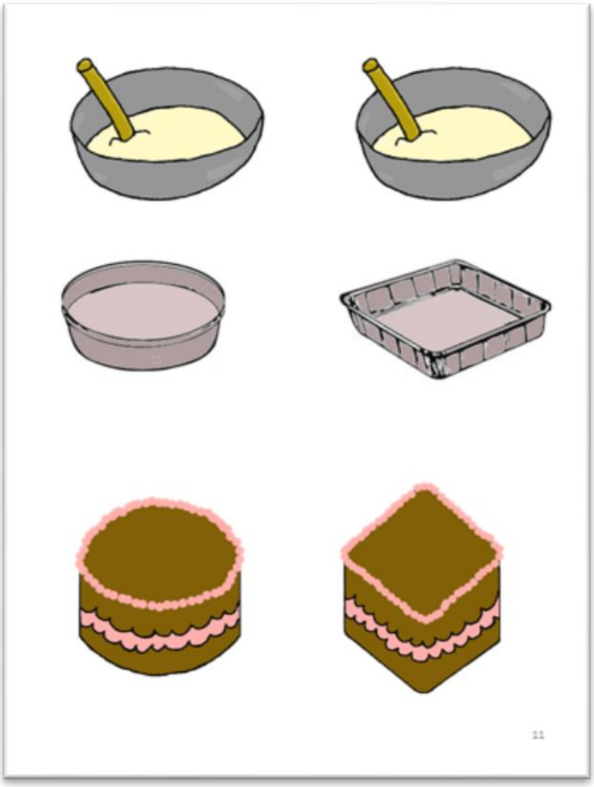

6


7


8







Prior Knowledge Question Sheet

<u>Name</u>: <u>Testleiter</u>: <u>Nummer</u>:

Beantworte die Fragen kurz! Meist ist ein einziges Wort ausreichend…

Weinbergschnecke

1. Wie viele Fühler hat die Weinbergschnecke?

2. Wo genau sitzen die Augen der Weinbergschnecke?

3. Mit welchem Körperteil atmet die Weinbergschnecke?

4. Nenne das Material, aus dem das Gehäuse von Weinbergschnecken besteht!

5. Womit zerkleinern Weinbergschnecken ihre Nahrung?

Maus

1. Wie nennt man die vergrößerten Schneidezähne bei Mäusen?

2. Wie viele dieser vergrößerten Schneidezähne haben Mäuse?

3. Von was ernähren sich neugeborene Mäuse?

4. Wie viele Zehen haben Mäuse an den Hinterfüßen (pro Fuß)?

5. Nenne die Sinnesorgane, die der Maus dabei helfen, im Dunkeln nirgends anzustoßen!

6. Nenne eine Funktion des Mäuseschwanzes!

Fische

1. Nenne das Organ, das Fischen hilft, mit anderen Fischen im Schwarm nicht zusammenzustoßen?

2. Womit atmen die meisten Fische?

3. Nenne den Grund dafür, dass Fische ihre Augen nie schließen!

4. Nenne eine Aufgabe der Schwimmblase von Fischen!

5. Wie viele Flossen haben die meisten Fische?

6. Wie nennt man die hinterste Flosse bei Fischen?

7. Wie nennt man die Knochenblättchen, die den Körper der meisten Fische bedecken?

Theory of Mind Question Sheet

<u>Psy-Testung – Theory of Mind Aufgaben</u>

Name: __________________ Testleiter: __________________ Nummer: ________

Aufgabe	**Anmerkungen**
Smarties	Was **glaubt** Lukas, was in der Schachtel ist? (Smarties oder ein Schwein?) ☐ Smarties ☐ Schwein Kontrollfrage: Hat Lukas schon mal in diese Schachtel **geschaut**? ☐ ja ☐ nein
Emotionen	Wie **fühlt sich Tim wirklich**? ☐ glücklich ☐ traurig ☐ zwischendrin Welches Gesicht wird **Tim versuchen zu machen**? ☐ glücklich ☐ traurig ☐ zwischendrin

Executive Functions Instructions

HERZEN

Das ist das Spiel mit den HERZEN!

Wenn du ein Herz siehst, drücke die Taste auf der GLEICHEN Seite wie das Herz.

Also: Wenn das Herz auf dieser Seite ist, drücke DIESE Taste! Und wenn das Herz auf dieser Seite ist, drücke DIESE Taste!

Jetzt bist du an der Reihe es auszuprobieren! Das ist nur ein Übungsversuch du musst dich also nicht beeilen, die Herzen warten auf dich!

DENK DRAN: Beim Spiel mit den Herzen, drücke die Taste auf der GLEICHEN Seite wie das Herz.

ÜBUNGSVERSUCH *→ jede Leistung loben!* (Gut, Prima, sehr gut, genau, super, toll, ja, richtig, …)

Danach auf die Frage fertig: y für ja (wenn Kind verstanden hat und schnell genug ist), n für nein (wenn Kind noch mehr Übung braucht)

Wenn noch mehr Übung:

Gut gemacht! Im richtigen Spiel sind die Bilder noch ein bisschen schneller. Lass uns noch ein paar schnellere probieren
ODER
Lass uns noch ein paar Runden mehr spielen!

Frage Fertig? → Ja

SEHR GUT GEMACHT! Jetzt ist es Zeit, das Spiel mit den Herzen richtig zu spielen! Versuche so schnell zu sein wie du kannst, aber mach langsam genug, sodass du die richtige Taste drücken kannst! Pass gut auf, jetzt geht's los!

Das hast du TOLL gemacht!

BLUMEN

Jetzt spielen wir das Spiel mit den Blumen!

Wenn du eine Blume siehst, drücke die Taste auf der ANDEREN Seite als die Blume!

Also: Wenn die Blume auf dieser Seite ist, drückst du DIESE Taste! Und wenn die Blume auf dieser Seite ist, drückst du DIESE Taste!

DENK DRAN Beim Spiel mit den Blumen, drücke die Taste auf der ANDEREN Seite als die Blume.

ÜBUNGSVERSUCH *→ jede Leistung loben!* (Gut, Prima, sehr gut, genau, super, toll, ja, richtig, …)

Danach auf die Frage fertig: y für ja (wenn Kind verstanden hat und schnell genug ist), n für nein (wenn Kind noch mehr Übung braucht)

Wenn noch mehr Übung:

Gut gemacht! Im richtigen Spiel sind die Bilder noch ein bisschen schneller.
Lass uns noch ein paar schnellere probieren
ODER
Lass uns noch ein paar Runden mehr spielen!

Frage Fertig? → Ja

Fantastisch! Jetzt ist es Zeit, das Spiel mit den Blumen zu spielen! Versuche so schnell zu sein wie du kannst, aber mach langsam genug, sodass du die richtige Taste drücken kannst! Denk dran: Blumen bedeuten ANDERE Seite. Pass gut auf! Jetzt geht's los!

KOMBINIERT

Prima gemacht! Jetzt ist es Zeit, das Spiel mit den Blumen und den Herzen ZUSAMMEN zu spielen! Diesmal gibt es keine Proberunde, Versuche also wieder so schnell zu sein wie du kannst, aber mach langsam genug, sodass du die richtige Taste drücken kannst! Denk dran.. Herz bedeutet GLEICHE Seite! Blume bedeutet ANDERE Seite! Pass gut auf, jetzt geht's los!

Observation Coding Scheme

FISCHE **Testzeitpunkt:_____ID-Nr:______ Auswertung von:___________**

Unspezifische Details

Dimensionen ☐ Unspezifische Details ☐ Spezifische Details ☐

Frage

- ☐ Spontane Frage ______________________
- ☐ Frage auf Nachfrage ______________________
- ☐ Frage mit viel Hilfe Emil:______________________
- ☐ Keine Frage ______________________
- ☐ Emil

Hypothese

- ☐ Spontane Hypothese ______________________
- ☐ Hypothese auf Nachfrage ______________________
- ☐ Hypothese mit viel Hilfe ______________________
- ☐ Keine Hypothese ______________________
- ☐ Keine Hypothese abgefragt ______________________

Testung

- ☐ (zu größten Teilen) Selbstständig
- ☐ Idee selbstständig, Hilfe bei der Umsetzung
- ☐ Idee vom Testleiter, weitestgehend selbstständige Umsetzung
- ☐ Idee und Umsetzung durch Testleiter
- ☐ Keine Beobachtung, nur Hypothese

Spezifische Details:

Interpretation

Zusammenfassung der Ergebnisse ☐ von selbst ☐ mit Aufforderung ☐ keine/falsch

Rückbezug auf Hypothese ☐ von selbst ☐ mit Aufforderung richtig ☐ keine/falsch ☐ keine Aufforderung

Trennung von Interpretation und Beobachtung ☐ vorhanden ☐ nicht vorhanden

Schnecke **ID-Nr:_____ Auswertung**

von:____________

Unspezifische Details

Dimensionen [ ] Unspezifische Details [ ] Spezifische Details [ ]

Frage

- ☐ Spontane Frage ____
- ☐ Frage auf Nachfrage ____
- ☐ Frage mit viel Hilfe Emil:____
- ☐ Keine Frage ____
- ☐ Emil

Hypothese

- ☐ Spontane Hypothese ____
- ☐ Hypothese auf Nachfrage ____
- ☐ Hypothese mit viel Hilfe ____
- ☐ Keine Hypothese ____
- ☐ Keine Hypothese abgefragt ____

Testung

- ☐ (zu größten Teilen) Selbstständig
- ☐ Idee selbstständig, Hilfe bei der Umsetzung
- ☐ Idee vom Testleiter, weitestgehend selbstständige Umsetzung
- ☐ Idee und Umsetzung durch Testleiter
- ☐ Keine Beobachtung, nur Hypothese

Spezifische Details:

Interpretation

Zusammenfassung der Ergebnisse ☐ von selbst ☐ mit Aufforderung ☐ keine/falsch

Rückbezug auf Hypothese ☐ von selbst ☐ mit Aufforderung richtig ☐ keine/falsch ☐ keine Aufforderung

Trennung von Interpretation und Beobachtung ☐ vorhanden ☐ nicht vorhanden

Maus **ID-Nr:_____ Auswertung von:___________**

Unspezifische Details

Dimensionen ☐ Unspezifische Details ☐ Spezifische Details ☐

Frage

- ☐ Spontane Frage ______________________________
- ☐ Frage auf Nachfrage ______________________________
- ☐ Frage mit viel Hilfe Emil:______________________________
- ☐ Keine Frage ______________________________
- ☐ Emil

Hypothese

- ☐ Spontane Hypothese ______________________________
- ☐ Hypothese auf Nachfrage ______________________________
- ☐ Hypothese mit viel Hilfe ______________________________
- ☐ Keine Hypothese ______________________________
- ☐ Keine Hypothese abgefragt ______________________________

Testung

- ☐ (zu größten Teilen) Selbstständig
- ☐ Idee selbstständig, Hilfe bei der Umsetzung
- ☐ Idee vom Testleiter, weitestgehend selbstständige Umsetzung
- ☐ Idee und Umsetzung durch Testleiter
- ☐ Keine Beobachtung, nur Hypothese

Spezifische Details:

Interpretation

Zusammenfassung der Ergebnisse ☐ von selbst ☐ mit Aufforderung ☐ keine/falsch

Rückbezug auf Hypothese ☐ von selbst ☐ mit Aufforderung richtig ☐ keine/falsch ☐ keine Aufforderung

Trennung von Interpretation und Beobachtung ☐ vorhanden ☐ nicht vorhanden

Well-Being and Involvement Coding Scheme

Tier: ___________________ **Nummer:**___________ **Ausgewertet von:** ___________________

Welche Signale sind für Wohlbefinden und Engagiertheit zu beobachten	Nicht vorhanden	Niedrig	Mittel	Hoch	Begründung
Emotionales Wohlbefinden					
Offenheit					
Flexibilität					
Selbstvertrauen					
Durchsetzungsvermögen					
Vitalität					
Entspannung, innere Ruhe					
Genießen können					
Im Einklang mit sich selbst					
Engagiertheit					
Gezielte Aufmerksamkeit					
Energie					
Vielschichtige Kreativität					
Gesichtsausdruck, Körperhaltung					
Ausdauer					
Genauigkeit					
Reaktionsbereitschaft					
Verbale Äußerungen					
Zufriedenheit					

Engagiertheitsskala:__________

Konkrete Beispiele:

Emotionales Wohlbefinden	
Offenheit	Erzählt was aus Privateleben, z.B.: das hab ich mit meinem Vater schon mal gemacht, meine Oma hat auch Fische… Blickkontakt Auch wenn schüchtern kann Kind offen sein -> reagieren auf VL
Flexibilität	Geht sofort auf Aufgabenstellung ein; verwendet die kennengelernten Utensilien; niedrig bei sprachlichen Problemen, wenn einfach nein gesagt wird weil das Kind nichts versteht oder nicht antworten kann
Selbstvertrauen	Nimmt sich einfach Gegenstände, verhält sich sehr natürlich
Durchsetzungsvermögen	Bringt eigene Ideen und will diese weiterführen; will beispielsweise unbedingt eine Schnecke aus dem Käfig raus haben; weiß, was es will
Vitalität	Bewegt sich viel, dreht sich zu Vl oder zu Tier, aber nicht hampelig, strahlt Vitalität aus
Entspannung, innere Ruhe	Niedrig bei Übersprunghandlungen oder wenn es nervös wirkt, hoch wenn es gelassen wirkt und nicht überspannt ist, nicht entspannt bei Ticks
Genießen können	Lacht, lächelt, Ausrufe wie „ Oh ist das süß“; stilles genießen, zufriedener Gesichtsausdruck
Im Einklang mit sich selbst	Entspannte Körperhaltung, entspannter Gesichtsausdruck

Engagiertheit	
Gezielte Aufmerksamkeit	Wendet sich der VL zu wenn es angesprochen wird, wendet sich dem Tier mit ganzem Körper zu wenn es beobachtet, schweift nicht mit den Augen ab (beim Nachdenken Augen schweifen lassen ist eher zeichen für Konzentration als für Ablenkung)
Energie	Mit viel Freude bei der Sache, erzählt, beschreibt mit Händen und Worten, Energie auf die Sache bezogen
Vielschichtige Kreativität	Fragt nach bringt ständig neue Ideen, Anregungen, nicht vorhanden, wenn Kind nur Anweisungen der Vl befolgt
Gesichtsausdruck, Körperhaltung	Hoch, wenn es sich dem jeweiligen Objekt (Emil, VL, Tier) direkt zuwendet aber entspannte Haltung hat, und entspannten oder gespannten Gesichtsausdruck hat, niedrig bei geringer Körperspannung gelangweilter Haltung und Gesichtsausdruck oder komplett abwesendem Gesichtsausdruck
Ausdauer	Hoch wenn es nicht ablenkbar ist und voll konzentriert den ganzen Zeitraum über mitarbeitet
Genauigkeit	Wie genau geht es mit Hilfswerkzeugen um, wie genau beschreibt es Situationen oder das Tier etc, hier zählt auch, wenn Kind beim beschreiben das Terrarium von den Schnecken beispielsweise genau beschreibt, „ da

	liegen Äpfel drin…“
Reaktionsbereitschaft	Hoch wenn es auf alle Denk- und Handlungsanstöße eingeht, Anregungen aufgreift
Verbale Äußerungen	Redet flüssig, zusammenhängende Sätze, etc
Zufriedenheit	Ist fasziniert, lächelt, positive Ausrufe „ Oh wie süß“, genießt Gespräche etc, lächelt, oder stille Zufriedenheit

Signale für emotionales Wohlbefinden:

Offenheit:

- Offen und empfänglich für seine Umgebung
- Gesichtsausdruck sind offen und direkt
- Empfängt und erwidert Blicke, Berührungen, Ermunterung, Trost oder angebotene Hilfe (Versuchsleitung)
- Kann positive und negative Gefühle ausdrücken
- Ist bereit, neue, fremde Situationen oder Menschen kennenzulernen

Flexibilität

- Findet sich schnell in seiner Umgebung zurecht (Nebenraum im Kindergarten)
- Kann sich auf neue oder ungewohnte Situationen einstellen und angepasst handeln (Ist ungewohnte Situation) also eher kann sich auf die neue, ungewohnte Situation einstellen
- Bleibt in Problem oder Frustration nicht stecken, sondern zeigt Bereitschaft, Alternativen anzunehmen oder Kompromisse zu schließen

Selbstvertrauen und Selbstwertgefühl

- Strahlt gehörige Portion Selbstbewusstsein aus
- Traut sich einiges zu, man kann es sehen und hören
- Zeigt ein beachtliches Selbstwertgefühl (ich bin da und wert, dass man mich wahrnimmt)
- Traut sich an Neuerungen heran
- Nimmt Herausforderungen an
- Es will ausprobieren, auch mit dem Risiko, es nicht zu schaffen
- Misserfolge belasten es nicht dauerhaft (schwer zu Bewerten weil nur kurze Videosequenz)
- Sucht Herausforderungen, die seinem Können und seinem Niveau entsprechen

Durchsetzungsvermögen

- Beansprucht Beachtung von seiner Umgebung (Versuchsleiterin)
- Möchte mit einbezogen und angenommen werden (eher in Gruppensituation zu bewerten)
- Vertritt eigene Wünsche, Bedürfnisse und Anliegen
- Lässt sich nicht schnell beeinflussen
- Geht nicht ohne weiteres auf Befehle oder Vorstellungen anderer Kinder ein, wenn dies den eigenen Bedürfnissen und Interessen widerspricht (eher in Gruppensituation, es werden keine Richtigen Befehle erteilt)

Appendix D – Coding Schemes

Vitalität

- Voller Leben und Energie, Lebensfreude und Kraft
- Vitalität im Gesichtsausdruck und Körperhaltung
- Augen strahlen (oft)
- Aufrechte Haltung, energischer Eindruck

Entspannung und innere Ruhe

- Gesichtsausdruck ohne ungewöhnliche Bewegungen oder Grimassen
- Entspannung in Körperhaltung und –bewegung
- Muskeln nicht ständig angespannt oder verkrampft
- Bewegungen geschmeidig und gleichmäßig
- Normales Sprechtempo und Stimmvolumen
- Anspannung nur zeitweilig und an eine bestimmte Aktivität gebunden (ist ja zeitweilig und eine bestimmte Aktivität…
- Nach aufregendem Spiel oder spannungsvoller Aktivität wieder schnell und völlig entspannt

Genießen können

- Kann genießen auch als „ stiller Genießer“
- Zeigt authentische, echte Freude, keine neurotische, ungesunde Form von Vergnügen
- Zeigt begeisterung
- Strahlt Zufriedenheit aus
- Spontanes singen, lacht oft

Im Einklang mit sich selbst

- Im Einklang mit eigenen Bedürfnissen, Wünschen, Gefühlen, Gedanken
- Scheint für sich selbst zu wissen, was es braucht
- Durchlebt Erfahrung intensiv
- Kennt zeitweise unangenehme, negative Gefühle, lässt sie zu und bewältigt sie (schwer in der kurzen Zeit zu beurteilen)
- Lebt in Frieden mit sich selbst rein Subjektiv wie der Beobachter sich in das Kind versetzt
- Fühlt sich verbunden mit anderen, mit der Natur,…

Stufen der Engagiertheit

Stufe 1: keine Aktivität

- Nicht – akitv
- Mit nichts beschäftigt, teilnahmslos, abwesend

Stufe 2: häufig unterbrochene Aktivität

- Momente echter Akitivität
- Nimmt weniger als die Hälfte der Zeit ein
- Lange Unterbrechungen

Oder:

- Mehr oder weniger dauerhaft aktiv
- Komplexität der Tätigkeit entspricht <u>nicht</u> Möglichkeiten und Fähigkeiten
- Eher mechanische Handlung, gewisse geistige Abwesenheit möglich

Stufe 3: mehr oder weniger andauernde Aktivität

- Mehr oder weniger beständige Beschäftigung
- Keine deutlichen Anzeichen von Engagiertheit
- Scheinen interessenlos, gleichgültig, kaum Eifer
- Handeln absichtlich aber nicht ganz bei der Sache
- <u>Auch</u> relativ intensive Aktivität die von Phasen der Nicht – Aktivität unterbrochen wird

Stufe 4: Aktivität in intensiven Momenten

- Kann vom äußerlichen Ablauf ähnlich zu Stufe 3 sein
- Mindestens die Hälfte der Zeit Elemente der aktiven Engagiertheit
- Wichtig / bedeutsam für Kind

Oder:

- Große Konzentration (vgl 5)
- Aktionen sehr motiviert
- Entsprechend bestimmten Zweck
- Aber innerhalb der Tätigkeit eher Routine, beinhalten keinen hohen Einsatz oder Leistung

Stufe 5: anhaltende intensive Aktivität

- Größtmögliche Engagiertheit
- Vertieft, gefesselt von Tätigkeit
- Augen nahezu ununterbrochen auf Material gerichtet
- Reize aus Umgebung können es nicht erreichen
- Erfordert geistige Anstrengung aber auf natürliche, selbstverständliche Art
- Gewisse Spannung wird deutlich
- Signale Konzentration, Energie, Ausdauer, Komplexität müssen deutlich wahrnehmbar sein

Biology Understanding Coding Scheme

Q1: Was bedeutet es, am Leben zu sein?

3 Punkte: Biologische Prozesse (Wachstum, Tod, Lebenszyklus)
2 Punkte: Biologisch relevantes Verhalten (Atmung, Essen)
2 Punkte: bedeutet, dass man stirbt
1 Punkt: Bewegung, Tätigkeit oder "nicht tot"
0 Punkte: weiß nicht, fehlende Reaktion

Q2: Kannst Du einige Dinge nennen, die am Leben sind, die Lebewesen sind?

2 Punkte: Menschen, Pflanzen, Tiere, entweder allgemein oder spezielle Beispiele
1 Punkt: Menschen und / oder Tiere, aber nicht Pflanzen
0 Punkte: weiß nicht, fehlende Reaktion
* 1 Punkt für jeden unbelebten Objekt subtrahiert

Q3: Kannst Du einige Dinge nennen, die nicht lebendig / am Leben sind?

2 Punkte: Jedes unbelebte Objekt
1 Punkt: Tote Menschen, Tiere oder Pflanzen
0 Punkte: Dinosaurier oder imaginären Einheiten

Ja / Nein-Entscheidungen: Ist ein x am Leben?

4 Punkte: Ja zu allen Tiere und Pflanzen; Nein zu Naturphänomenen und Dingen
3 Punkte: Ja zu allen Tieren, nein zu einer / beiden Pflanzen, Naturphänomenen und Dingen
2 Punkte: Ja zu allen Tieren und einem oder mehreren Naturphänomenen; nein zu Dingen
1 Punkt: Ja zu allen Tieren und einem oder mehreren Dingen.

Begründungen:
Du hast gesagt, dass ein x lebendig / nicht lebendig ist. Woher weißt du das?

2 Punkte:
Für das Tier, z.B. eine biologische Antwort "Eine Katze lebt, weil sie atmet".
Für Dinge, von Menschen verursachte Herkunft, beispielsweise "Ein Flugzeug ist nicht am Leben, weil die Menschen es geschaffen haben."

1 Punkt: Autonome **Bewegung**, beispielsweise "Eine Katze lebt, weil sie Dinge tut, von selbst", "ein Flugzeug ist nicht am Leben, weil es sich nicht von selbst bewegen kann."

0 Punkte: falsche Antworten, Bewegung, Aktivität, Körperteile , wie zB "Eine Katze lebt, weil sie sich bewegt und Füße hat."

-1 Punkt: Existenz, z.B. "Eine Katze lebt, weil ich es sehe."

Für jede der Begründungsfragen wurden Punkte enstprechend dieses Schemas zugeschrieben, so lange die jeweiligen Begründungen nicht zuvor schon erwähnt wurden; z.B. die Antwort "ein Tisch ist nicht mehr am Leben, weil er nicht atmet" würde keine Punkte geben, wenn sie auf eine Antwort, dass "eine Katze lebt, weil sie atmet", folgt; würde aber 2 Punkte geben, wenn sie auf eine Antwort folgen würden, dass "eine Katze lebt, weil sie wächst."

Die Punkte wurden addiert, um den Gesamt-Animismus-Score zu errechnen.

Training Evaluation Form

Trainingsstunde Nr. und Datum:

Thema:

Das lief gut:
Das hat nicht so gut geklappt:
Das habe ich verändert:

Evaluation Kinder

1= trifft gar nicht zu, 2= trifft kaum zu, 3=teils-teils, 4=trifft eher zu, 5=trifft sehr zu

Name Kind

… hat aktiv am Programm teilgenommen	1	2	3	4	5
… hatte Schwierigkeiten den Inhalten zu folgen	1	2	3	4	5
… hat heute Beobachtungsverhalten gezeigt	1	2	3	4	5

Name Kind

… hat aktiv am Programm teilgenommen	1	2	3	4	5
… hatte Schwierigkeiten den Inhalten zu folgen	1	2	3	4	5
… hat heute Beobachtungsverhalten gezeigt	1	2	3	4	5

Name Kind

… hat aktiv am Programm teilgenommen	1	2	3	4	5
… hatte Schwierigkeiten den Inhalten zu folgen	1	2	3	4	5
… hat heute Beobachtungsverhalten gezeigt	1	2	3	4	5

Name Kind

… hat aktiv am Programm teilgenommen	1	2	3	4	5
… hatte Schwierigkeiten den Inhalten zu folgen	1	2	3	4	5
… hat heute Beobachtungsverhalten gezeigt	1	2	3	4	5

Name Kind

… hat aktiv am Programm teilgenommen	1	2	3	4	5
… hatte Schwierigkeiten den Inhalten zu folgen	1	2	3	4	5
… hat heute Beobachtungsverhalten gezeigt	1	2	3	4	5

…

Eidesstattliche Versicherung
Statement of Scientific Integrity

Klemm, Janina

Name, Vorname
Last name, first name

Ich versichere, dass ich die an der Fakultät für Psychologie und Pädagogik der Ludwig-Maximilians-Universität München zur Dissertation eingereichte Arbeit mit dem Titel:
I assert that the thesis I submitted to the Faculty of Psychology and Pedagogy of the Ludwig-Maximilian-Universität München under the title:

Biological Observation Competency in Preschool – the Relation to Scientific Reasoning and Opportunities for Intervention

selbst verfasst, alle Teile eigenständig formuliert und keine fremden Textteile übernommen habe, die nicht als solche gekennzeichnet sind. Kein Abschnitt der Doktorarbeit wurde von einer anderen Person formuliert, und bei der Abfassung wurden keine anderen als die in der Abhandlung aufgeführten Hilfsmittel benutzt.
is written by myself, I have formulated all parts independently and I have not taken any texts components of others without indicating them. No formulation has been made by someone else and I have not used any sources other than indicated in the thesis.

Ich erkläre, das ich habe an keiner anderen Stelle einen Antrag auf Zulassung zur Promotion gestellt oder bereits einen Doktortitel auf der Grundlage des vorgelegten Studienabschlusses erworben und mich auch nicht einer Doktorprüfung erfolglos unterzogen.
I assert I have not applied anywhere else for a doctoral degree nor have I obtained a doctor title on the basis of my present studies or failed a doctoral examination.

München, 10.11.2016

Ort, Datum
Place, Date

Janina Klemm

Unterschrift Doktorandin/Doktorand
Signature of the doctoral candidate