UNIVERSITÄT DUISBURG ESSEN

Improving the Quality of Instruction for Egyptian Science Classes: Developing and Evaluating a Teaching Program for Middle School Students

Dissertation

von

Mohamed Ali Ahmed Shahat

aus Aswan, Ägypten

eingereicht zur Erlangung
des Doktorgrades der Naturphilosophie (Dr. phil. nat.)
(Didaktik der Physik)

an der Fakultät für Physik
der Universität Duisburg-Essen, Deutschland

Essen, 2013

Erstgutachter: Prof. Dr. Hans E. Fischer
Zweitgutachter: Prof. i.R. Dr. Dr. h.c. Reinders Duit

Tag der mündlichen Prüfung: 01.02.2013

Acknowledgements

It is my great pleasure to thank the many people who supported me on this study over the last four years. Unfortunately, I am unable to name everyone here, so to those of you not specifically mentioned below: Thank you!

First of all, I would like to express my deep gratitude to my supervisor, Professor Hans Fischer. I deeply enjoyed working as your PhD-student and am grateful that you made this contribution between Germany and Egypt possible. I am also very thankful for your professional guidance, constructive comments, intellectual discussions and support. Your kindness will stay in my memory forever.

My enormous gratitude also goes to ‘my’ post-doc Annika Ohle, who always supported me while she was at Essen and continued to when she left to Dortmund. Thank you, Annika, for your helpful comments and valuable feedback.

I would like to thank all my NWU colleagues who so freely gave their time and were happy to share discussions, especially my officemates, Felix Schoppmeier, and Bettina Kreiter. Special thanks also to Jon Higgins, whose the eagle eye proved invaluable at the proofreading stage.

It is also important that I recognize and thank the financial support of ‘my country’ the Arab Republic of Egypt to this project. I also thank all my professors, and colleagues at the Faculty of Education, Aswan University for their kind support as well as all the students who participated in my study and each school which facilitated this participation. Their generosity in allowing me access to their daily schooling made this research possible.

My family has of course always been the most important aspect of my life. Special thanks go to my wife, my son Ahmed, and my daughter Yara, for allowing me the time to complete my study in Germany, and especially for inspiring me and supporting me during those critical moments and times of doubt. I would never have finished this study without you in my life.

And last but certainly not least, I am deeply grateful to my parents and my sister for their unwavering support over the years.

Table of Contents

1. Introduction

In TIMSS 2003 and 2007, the average science score of Egyptian eighth grade students was 421 and 408, which was significantly less than the TIMSS 2003 and 2007 scale means of 474 and 500 (Martin, Mullis, Gonzalez & Chrostowski, 2004; Martin, Mullis & Foy, 2008). As a result, Egypt ranked 41st out of the 59 participating countries in TIMSS 2007 (Martin et al., 2008). The TIMSS Study results additionally showed that low student achievement scores was not only a problem in Egypt, but in many other countries as well. So in the years since then, educators and researchers have debated which school variables influence student outcomes. As policymakers become more involved in school reform, this question has taken on new importance since many initiatives rely on presumed relationships between various education-related factors and learning outcomes (Darling-Hammond, 2000). Some studies on the question of what influences student outcomes have suggested that the low student achievement scores are assumed to be a result of student effort, social context and the role of teachers in school (see Darling-Hammond, 2000; Ronald, 2009, Roderick & Engel, 2001). Another recent study assumes that teachers' instruction methods affect student performance (Good & Brophy, 2008).

In addition, the Ministry of Education in Egypt found that in secondary schools, students have been enrolling in the liberal arts increasingly more than they have been in the sciences (UNESCO, 2008), which can be taken as an indicator of lower interest in sciences. Newburghl (2008) states that the loss of science students is an alarming trend and is an extremely relevant issue for developing countries where significant changes are taking place in science programs offered to young citizens, such as in Egypt.

Currently, educators and policymakers envision the future of education in Egypt as one in which schools provide high-quality education for every learner. They assert that the learning environment should be based on active learning to enable learners to acquire self-learning, scientific thinking, critical thinking, and problem solving skills, which is related to increase student achievement and interest (Ministry of Education-Egypt, MOE, 2010). The MOE, therefore, demands reforms in science teaching and the development of new teaching approaches which motivate younger students to study science (MOE, 2010). With this in mind, this study aims to develop and evaluate a teaching program based on a scientific inquiry approach, as demanded by the Ministry of Education. Therefore, a model is first needed that describes what students need to do during a problem solving process. In the presented study, such a model has been derived from a psychological (Klahr & Dunbar, 1988) and a practical (Oser & Baeriswyl, 2001) model for problem development. This model has been evaluated by expert ratings and was later on used for planning this study's teaching program. The general idea behind increasing quality of instruction in the presented study is that the teacher offers learning opportunities that are used by their students in individual ways.

The lesson plans of the teaching program were constructed based on the established model for problem solving. The program itself consists of five double lessons, providing content knowledge on the topic *density and buoyancy* and introducing the method of problem solving. This topic was chosen because it is part of the Egyptian curriculum and because students tend to have difficulties learning about this topic. In addition, the topic provided a broad basis for problem solving tasks with many already well-established experiments. Six instruments were applied to assess students' problem solving abilities, experimental strategy knowledge, science achievement, perceptions on the quality of science lessons, motivation towards science and, finally, teachers' quality of instruction. All instruments had to be translated into Arabic from German or English; the translation process and the quality control were closely screened.

In order to verify the learning effects of this program, a classical control group design was used. Altogether four classes of seventh grade students participated in this study; two classes were taught as intervention-group that learned with the developed program, while two classes as control-group were taught according to the common, regular teaching methods. To ensure the comparability and quality of teaching in both groups, a second science teacher filled in observer sheets (Clausen, 2002). Students in both the control and intervention groups were similar in age, gender, academic pre-knowledge, background, and cognitive abilities. The analyses confirmed that there were no significant differences between students in the both groups (CG & IG) regarding the pervious control variables. The hypothesis, which assumed that there are no significant differences between the control and intervention group regarding the teachers' quality of instruction, was also tested. The analyses confirmed, based on a standardized questionnaire (Clausen, 2002), that there were no significant differences between both groups (CG & IG) regarding the teachers' quality of instruction.

Students' problem solving abilities, experimental strategies knowledge, achievement, perceptions on the quality of science lessons and motivation towards science were assessed in a pre-post-design. Analyses concluded that there are differences between both groups (CG & IG) regarding the increase of students' problem solving abilities, experimental strategy knowledge, perceptions on the quality of science lessons, achievement and motivation towards science. Students in the intervention group performed significantly better with an overall small and medium effect size.

In summary, the teaching program and instruments were partly self-developed and partly adapted to the school conditions in Egypt from newly performed studies on strategy knowledge and problem solving in Germany, and PISA and TIMSS tests. From September to October 2011, an intervention study was conducted in a quasi-experimental design with pre- and post-testing (e.g. Shadish, Memphis, Dood & Evanston, 2002). The sample included n=147 seventh grade students in a general middle school in the city of Aswan, Egypt during the 2011-2012 school year. The intervention group consisted of 74 students (50% males), while the control group contained 73 students (49.5% males). Both groups showed no significant differences in age, cognitive abilities, background, academic pre-knowledge and teachers' quality of instruction. Two treatments were used with the control and intervention group: the newly developed teaching program was used to teach the intervention group while the control group was taught according to regular teaching methods. Both groups were taught by the same teacher (the author, who is a qualified science teacher). To ensure the comparability and quality of teaching in both groups, a second science teacher filled in observer sheets (Clausen, 2002). The study was conducted during a six week-long unit at a middle school in Egypt. Intervention and control group students were compared to determine differences in learning outcomes and motivation. The data were appropriately analyzed according to the Rasch model and classical test theory. The results revealed that students in the intervention group performed significantly better with a small and medium effect size ($.1<r<.5$ & $.1<d<.8$) in the aforementioned dependent variables. Regarding the small sample size, the generalizability of results is limited. Nevertheless, the presented study provides evidence that this teaching program can be effective and encourages further research to confirm the presented results on a larger sample.

2. Theoretical background

The Egyptian Ministry of Education demands science teaching reforms and the implementation of new teaching approaches which focus on scientific inquiry and problem solving processes. Further, motivating more students to pursue science studies in earlier stages is needed. Thus, the aim of the presented study is to develop and evaluate a science teaching program based on a model for problem solving on seventh grade students' problem solving abilities, experimental strategy knowledge, achievement, perceptions on the quality of science lessons and motivation towards science in Egyptian classes.

In this chapter, the context of this study will be clarified. The current situation of science teaching and learning, and the new vision of education in Egypt are presented. The elements of the Egypt educational system relevant for this study are also described. Following this, the model for quality of instruction, which was used to develop the teaching program, is described. After that, the description of problem solving and strategy knowledge will be provided. At the end of this chapter, the research questions are gleaned from the theoretical background and the according hypotheses will be stated.

2.1 Situation of science teaching and learning in Egypt

In TIMSS 2003 and 2007, the average science score of Egyptian eighth grade students was 421 and 408, which is significantly less than the TIMSS 2003 and 2007 scale means of 474 and 500 (Martin et al., 2004; Martin et al., 2008). As a result, Egypt ranked 41st out of the 59 participating countries in TIMSS 2007 (Martin et al., 2008). The TIMSS Study results additionally showed that low student achievement scores is not only a problem in Egypt, but in many other countries as well. In this way, educators and researchers have debated for many years which school variables influence student outcomes. As policy-makers become more involved in school reform, this question takes on new importance since their many initiatives rely on presumed relationships between various education-related factors and learning outcomes (Darling-Hammond, 2000). Some studies have suggested that the low student achievement scores are assumed to be a result of student effort, social context and role of teachers in school (e.g. expectations, efficacy and support, and structure of activities) (see Darling-Hammond, 2000; Roderick & Engel, 2001; Ronald, 2009). One study assumes that teacher's instruction methods affect in student outcomes (Good & Brophy, 2008).

In addition, the Ministry of Education in Egypt found that in secondary schools, students have been enrolling in the liberal arts increasingly more than they have been in the sciences (UNESCO, 2008). For example, Newburgh (2008) states that the loss of science students is an alarming trend, and is an extremely relevant issue for developing countries where significant changes are taking place in science programs offered to young citizens, such as in Egypt. Therefore, a reform of science teaching, and finding new teaching approaches, which motivate students in earlier stages towards science is demanded (Ministry of Education, MOE, 2010).

Traditional patterns of science education, which build a strong theoretical tradition with less emphasis on laboratory and practical experiences, are changing rapidly. New programs and systems with more emphasis on research are being introduced. The content and teaching methods incorporated in these programs are based on the most up-to-date theories about which science is most worth knowing. The primary goal of the new generation science programs is to reflect the national interest in having a scientifically literate population (Hassan, 1997).

Egypt's teaching vision

"The amount of scientific knowledge has increased exponentially over the past several decades" (Bloom, 2006, p. 4). During this period, "there has been a rethinking by philosophers of science about the nature of scientific inquiry (i.e., the practices and process of the growth of scientific knowledge)" (Duschl & Hamilton, 2011, p.78). Most "curricular programs are moving away from isolated skills to the integration of several different aspects of inquiry skills such as hypothesis generation and evaluation" (Jeong & Songer, 2008, p.195). Instruction used to focus either on content knowledge or process skills, but in recent years the development of content knowledge and inquiry skills are increasingly addressed simultaneously (e.g. Chang, 2011; Kim & Hannafin, 2011; Tang, Coffey, Elby & Levin, 2009; Yoon, 2009) by emphasizing student questioning, observing, inferring, classifying, interpreting, analyzing, predicting, investigating, and problem solving (DeBoer, 2006; Lederman & Lederman, 2012, Quintana et al., 2004).

In Egypt, the common method of science education can be described as one of explanation and lecturing: In science class, the teacher presents examples of concepts, writes the rules, generalizations and characters related to these concepts, gives students time to answer some questions individually to evaluate them, and finally he or she draws conclusions on the blackboard. The teacher may also use some concrete materials to explain specific content features of the lesson. One can say that, typically, the teacher acts and speaks much more than his or her students (Hashimoto, Pillay & Hudson, 2008; UNESCO, 2008). Since Egyptian society needs new generations to be capable of dealing with the requirements of a *new knowledge*, the Ministry demands that the following skills, knowledge and values be improved: advanced knowledge, self-learning, citizenship values, dialogue and acceptance of others, enlightened morals, religious values and social cohesion, problem solving, creativity and scientific research. In order to foster these skills, the Ministry is following several trends and pursuing methods that are considered the basis of improving education and its outcomes. These trends and methods are the learner-centered approach, the encouragement of learning through exploration, linking education and learning to real-life contexts, continual meditation and reflection in educational processes and outcomes, and continuous and comprehensive evaluation of the learner's performance (see Ministry of Education, MOE, 2008, 2010; UNESCO, 2008). Therefore, since new teaching approaches need to be developed and evaluated in Egypt, instruments for evaluating these teaching methods are needed. Thus, students' achievement, problem solving abilities, experimental strategy knowledge, perceptions on the quality of science lessons, and motivation towards science have to be fostered and evaluated. Designing programs and instruments for these processes are needed to measure their implementation.

Since the study presented here focuses on seventh grade students in middle school science classes, it is important to know how the school system in Egypt is structured.

2.1 The educational system in Egypt

The Egyptian educational system is the largest in the Middle East North Africa (MENA) region and is one of the largest in the world. It comprises 17 million students, 821,000 teachers and 40,000 schools. The system employs about 1.6 million teachers and administrators, which is the largest group of civil servants in Egypt (UNESCO, 2008). The system is centralized in the Ministry of Education (MOE), the Ministry of Higher Education (MOHE) and Al-Azhar religious education system. The MOHE is responsible for universities, including the faculties of education and institutions of technical and professional training. The Azhar education system is supervised by the Supreme Council of Al-Azhar Institution and is independent from the MOE. It follows the same direction of the general education. Pre-tertiary education involves three stages, namely, primary, preparatory, and secondary stages (general, technical and vocational). Until this moment, early childhood education (ECE) is not considered as part of the official educational system in Egypt (UNESCO, 2008).

Referring to Figure 1 below, the educational phases in Egypt (BouJaoude, Wiles, Asghar & Alters, 2011; Mansour, 2010; OECD, 2010; UNESCO, 2008) can be described as five phases, and they are: Basic Education, General Secondary Education, Technical Secondary Education, University and Higher Education, and Al-Azharite Education. Relevant elements of the Egypt educational system for this study are described below.

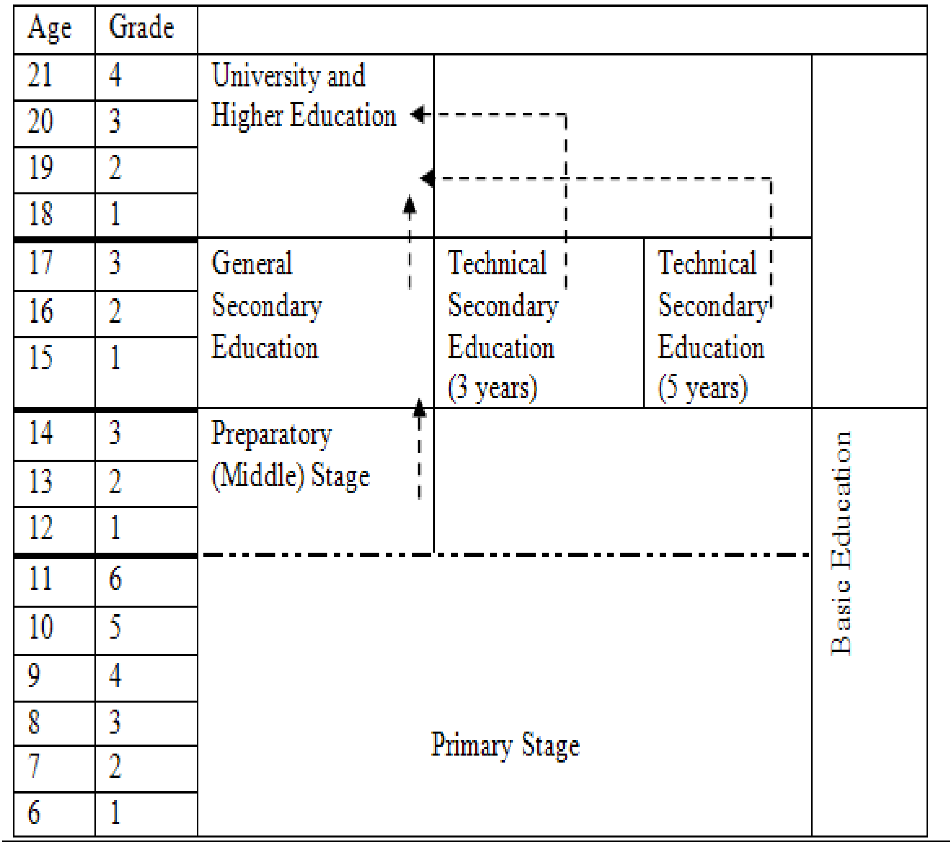


Figure1. Structure of Educational System in Egypt (UNESCO, 2008, p. 75)

Basic Education (Primary and Preparatory (middle)) consists of six years of primary school starting at the age of six and a subsequent three-year period at preparatory school. Students learn physics, chemistry and biology in their science courses. There are three types of schools in addition: schools for boys, schools for girls and co-ed schools. At the end of preparatory school – after Grade 9 – students are divided into one of two tracks: general secondary schools or technical secondary schools.

General Secondary Education comprises a three-year stage that starts from Grade 10 and prepares students for work and further education. Students learn physics, chemistry and biology as separate courses. There are three types of schools in addition: schools for boys, schools for girls and co-ed schools. Graduates of this track normally join higher education institutes in a highly competitive process based mainly on their results on the secondary school graduation examination.

Technical Secondary Education (industrial, agricultural, commercial and other school types) includes two tracks. The first provides technical education in a three-year technical secondary school. The second provides more advanced technical education in an integrated five-year school; the first three years are similar to those of the former type and the final two years prepare graduates for jobs as technicians. Students learn physics, chemistry and biology as separate courses. Graduates of both tracks may access higher education depending on their results in the final exam. However, their transition rates are low in comparison to graduates of general secondary education.

Al-Azharite Education follows the same direction of the general education with regard to hours of study for each school subject. However, Al-Azhar offers religious instruction as part of the curriculum. In primary education, 62% of students attend public schools, 29% private schools and 9% religious Al-Azhar schools. Public education plays a greater role in general secondary education accounting for 92% of all enrolments. The average class size in Egypt is 39.65 students (MOE, 2008), which had to be taken into account for planning the learning program. The primary language of instruction in Egyptian schools is Arabic.

The current study was conducted on a sample of seventh grade middle (preparatory) school students. The sample is described in Chapter 3.2. Since new teaching approaches need to be developed and evaluated in Egypt, instruments for evaluating these teaching methods are needed, and those are the two aims of the presented study. In order to increase students' motivation and achievement, changes in teaching and learning processes are needed, which are implemented in the teaching program by using a model for quality of instruction (Fischer et al., 2005; Ohle, 2010).

2.2 A Model for quality of instruction

According to Jimerson, Burns and Van Der Heyden (2007) fostering the success of students is the primary goal of educational systems. And identifying individual needs by providing appropriate interventions is important for enhancing student outcomes. In the presented study, in order to develop the proposed teaching program, a model for quality of instruction is needed. The following literatures give an overview of the different ways quality of instruction variables have been operationalized.

Carroll's (1963) model of school learning stated that students' degree of learning is depended on the student's time and needs on something in order to learn it. He suggested some characteristics for quality of isntruction such as clarity of the learning goals, adequate presentation of the learning material, and, a planned series of learning steps.

Bloom (1976 cf. Neumann, Kauertz & Fischer, 2012) focused on the importance of students' prerequisites, specially their cognitive abilities, for the learning process. He identified a set of characteristics that could influence the learning process, such as: cues and feedback, reinforcement and participation. He has mentioned that, the overall influences of quality of instruction on student achievement are considered to be only moderate, however, students' cognitive abilities are considered to have the highest influence.

Creemers (1994) and Creemers & Kyriakides (2008) described quality of instruction based on Carroll (1963) by distinguishing between three components of the quality of instruction: school level (e.g. quality/educational, quality/organizational, time opportunity), classroom level (e.g. curriculum, groping procedures, and teacher behaviors), and students level (e.g. time on task, opportunities used, motivation, attitudes, social background, achievement).

Another model that has evolved from Carroll's (1963) model of school learning is the model of educational productivity proposed by Walberg (1981). A major new feature in Walberg's (1981) model was the provision of the learning environment and its influence on students' learning time. Altogether, Walberg (1981) identifies nine factors that influence affective, behavioral and cognitive learning: ability or prior achievement, age and development, motivation or self-concept, quantity of instruction or time engaged in learning, quality of instruction, home environment, classroom environment, peer group environment, and the mass media.

Dunkin and Biddle (1974, cf., Neumann et al., 2012) have developed a model of classroom learning. The model focuses on four classes of factors:

teacher characteristics, context variables, process variables, and product variables.

Walberg et al. (1981) have described that there are some factors significant, and they are: socio-economic status, motivation, quality of instruction, class social-psychological environment and home conditions. While other factors as race and gender were controlled, however, the analysis showed that the social- psychological environment had the only unequivocal effect on science learning.

An all-embracing review of process-product research was written by Jere Brophy and Thomas Good (Brophy & Good 1986, cf., Neumann et al., 2012), describing two dimensions of characteristics as characteristics related to quantity and pacing of instruction and qualitative characteristics.

They found that there is a positive impact of some quantitative characteristics on students' instructional outcomes like the amount of opportunities to learn, the content covered, role definition, expectations, allocation, classroom management, student engaged time, consistent success, academic learning time, and active teaching.

Quality of instruction research underwent a major revival due in a large part to the so-called TIMSS Video Study (Stigler, Gonzales, Kawanaka, Knoll & Serrano, 1999). Based on an opportunity-to-learn model of instructional quality, analysis of video lessons was based on three dimensions of teaching quality: classroom management, cognitive activation and student-centered orientation (Lipowsky et al., 2005).

In more recent work, another model describing also the processes in the classroom is Helmke's model for quality of instruction (Helmke, 2003). The general idea behind this mediation model (see Figure 2) is that the teacher offers learning opportunities that are used by the students in individual ways. The model includes teachers' personality on the one side, embracing professional knowledge, motivation, educational background as well as values and aims.

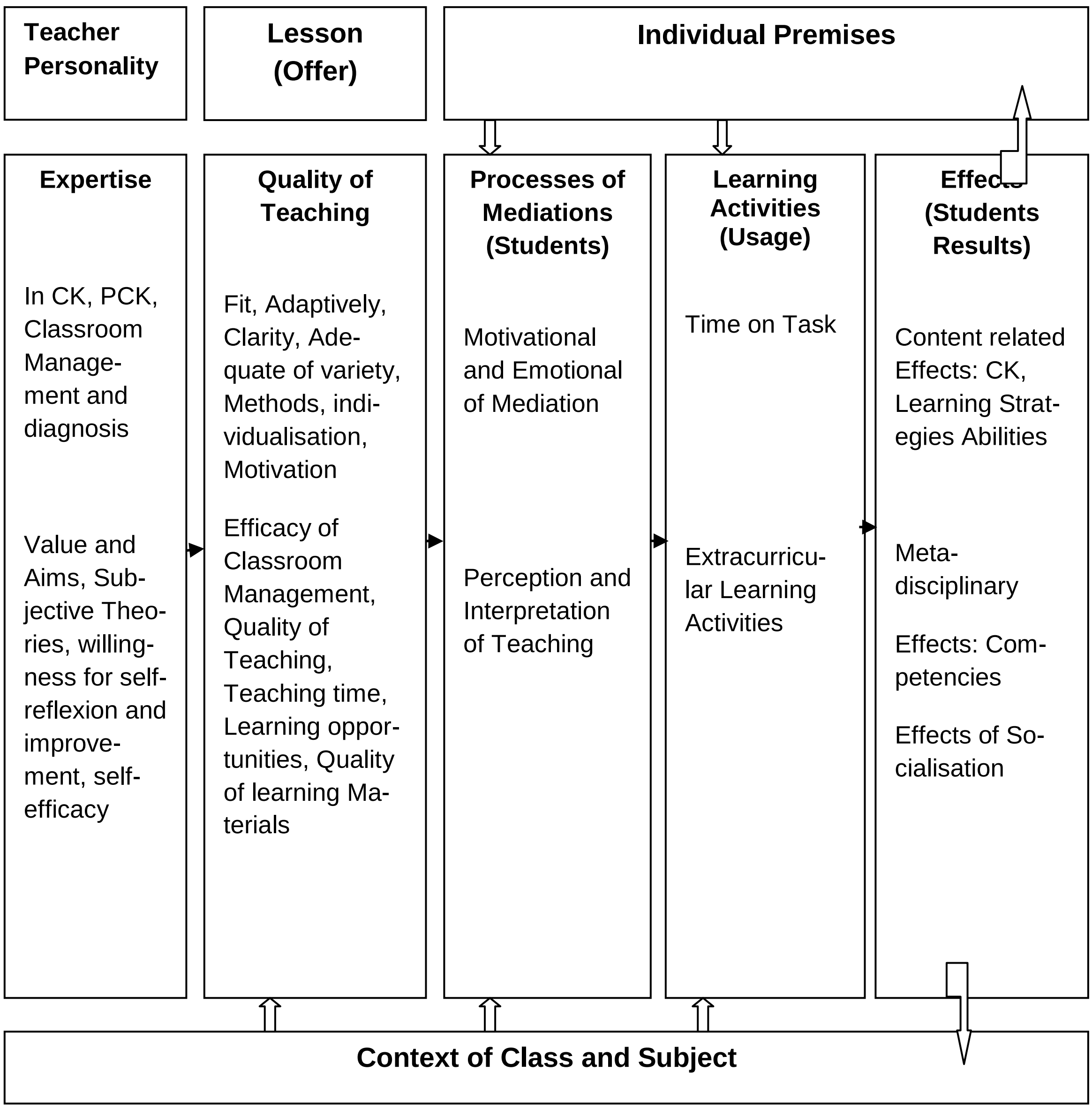


Figure 2. Model for quality of instruction (Helmke, 2003 cf., Ohle, 2010)

As this shows, teachers offer learning opportunities. These learning opportunities have to be appropriate to students' pre-conditions (e.g. abilities and motivational prerequisites), which are therefore also included in these models. In the classroom students, use the teachers' learning offers, which have to fulfill quality criteria like clarity, efficacy of classroom management or quality of learning materials. The way students use the learning offers is mediated by students' motivation and their perception of learning offers. Though it was not possible to videotape the classes during the intervention study's period in Egypt, the quality of instruction in this study was assessed by measuring the following:

First, students' perceptions on the quality of science lessons using an adapted instrument (Olson, Martin & Mullis, 2008; Reyer, 2004).

Second, the teachers' quality of instruction was also evaluated using a standardized instrument (Clausen, 2002).

The Clausen questionnaire measures the following aspects: rule clarity, discipline, repeat practice, challenge practice, diagnosis of competence in social aspects, positive attitude of students, diagnosis of competence in achievement, individual reference of attitude, classroom management, volatility, motivation, clarity and structure of instruction, interaction rate, pacing, amount of wasted time in instruction, instruction's individualization, time on task, aggression towards teacher, aggression towards classmates, scientific production of the students, error culture, structuring aids and individual learning support. Accordingly, this study considers several variables that might impact student outcomes in order to evaluate the quality of instruction for teacher and students. Variables assessed in the presented study are teacher's quality of instruction, students' perceptions on the quality of science teaching and individual learning premises, as well as students' outcomes. Class size, as well as regional, cultural, and diversity aspects also influence the teaching processes (e.g. Blatchford et al., 2003; Rice, 1999; Ohle, 2010) – all of which need to be taken into account for learning program and interpreting results.

Thus, the aims of this study are to develop a model for problem solving and to design instruments for evaluating Egyptian science classes, which are taught according to a new, more student-centered teaching approach. The instruments need to assess students' content knowledge on the taught topic, problem solving- and experimental strategy knowledge, and perceptions on the quality of science lessons as well as students' motivation. In closing, Figure 3 illustrates which variables in the current study are assessed.

Lesson			**Individual learning premises**
Teacher action	⇔ Student action		Understanding of the topic (pre) Motivation (pre) Cognitive abilities (pre) Gender (pre) Age (pre) Social background (pre) Problem solving (pre) Experimental strategy knowledge (pre) Perception (pre)
Learning opportunities	Perception of learning opportunities	⇐	
Quality of instruction (during the intervention time)	Perception of implementation of students' interest and understanding		

⇩

Outcomes
Understanding of the topic (post) Motivation (post) Problem solving (post) Experimental strategy knowledge (post) Perception (post)

Figure 3. Variables assessed in the current study

2.3 Describing problem solving and experimental strategy knowledge

Describing students' problem solving abilities

The national science education standards in many countries (Bernholt, Neumann & Nentwig, 2012) recommend that teachers focus on inquiry, and predominantly on real phenomena in classrooms, outdoors, or in laboratory settings, where students are given investigations or guided toward designing investigations/experiments that are demanding but within their capabilities. Although in general, it is true that inquiry requires new ways of thinking, developing these habits of mind, especially through hands-on experiments, is not yet standard practice in many science classrooms. In part, this is because teachers lack guidance or training in designing investigations in ways that facilitate students' practicing and learning to inquire and think critically about evidence and their discoveries (Sutman, Schmuckler & Woodfieldand, 2008). "In this sense, scientific inquiry is viewed as a teaching approach used to communicate scientific knowledge to students (or allow students to conduct their own knowledge) as opposed to an educational outcome that students are expected to achieve" (Lederman & Lederman, 2012, p. 339). For example, according to Zimmerman (2000) scientific inquiry is a problem-solving activity that uses the same information processing mechanisms that have been identified in other problem-solving contexts. Based on this background and to enable students to engage in a scientific inquiry process in science classes, students need a model-guided instruction. Therefore, inquiry learning should be structured in a way that students are provided with information related to a certain problem situation and encouraged to plan, conduct, and evaluate their own investigation (Berg, Bergendahl, Lundberg & Tibell, 2003). Scientific inquiry and problem solving processes have been investigated in various studies and, as a consequence, are operationalized in different ways (see Figure 4). The following table gives an overview of operationalizations of scientific inquiry and problem solving processes.

Table1. Overview of some models of scientific inquiry and problem solving

	Isaksen & Treffinger (1985)	**Klahr & Dunbar (1988)**	Kneeland (1999)	White, Shimoda & Frederiksen (1999)	**Oser & Baeriswyl (2001)**	Klos et al. (2008)	Dogru (2008)
Steps	-Mess Finding -Data Finding -Problem-Finding -Idea-Finding -Solution Finding -Acceptance Finding	**Formulate hypotheses based on pieror knowledge** **-Conduct experiments** **-Evaluate the results**	Awareness of the problem -Gathering of relevant facts -Definition of the problem -Development of solution options, -Selection of the best solution -Implementation of the solution	**-** Question -Hypothesize -Investigate -Analyze -Model -Evaluate	**-Understand the problem (problem presentation, the discovery of a problem, reformulation of the problem task)** **-Develop hypotheses about possible ways to find a solution** **-Test the hypotheses** **-Evaluate and apply the solutions**	Finding an idea/hypothesis - Planning and conducting an experiment -Drawing conclusions from experimental evidence	Understanding the problem, -Gathering information about problem solving, Solution and interpretation of the information about the problem -Determining ways of solution -Determining the best effective solution -Preparing report and its evaluation

Because the future vision of education in Egypt, which focuses in particular on the need for improving self-learning, scientific inquiry and problem solving, and because there is no current attempt being made to use the aforementioned psychological and practical teaching models of problem solving in teaching environments, this study aims to develop a model on problem solving that deals with the demanded aspects described above and combines Klahr and Dunbar’s (1988) model of scientific discovery as a dual search (SDDS) with Oser and Baeriswyl's (2001) practical teaching theory into one synthesized model. The two models were combined as illustrated in Figure 2 for two reasons. (1) Students and teachers are not used to solving such problems and following such steps. (2) Before activating pre-knowledge (Klahr & Dunbar), students need to identify and formulate the problem (Oser & Baeriswyl). In addition, before formulating hypotheses (Klahr & Dunbar), students need to define and represent the problem (Oser & Baeriswyl). Furthermore, before performing the solving process (Klahr & Dunbar), students need to explore a possible way of solving the problem (Oser & Baeriswyl), and before evaluating the results (Klahr & Dunbar), students need to fix data and calculate (Oser & Baeriswyl). The combined model of problem solving was validated (see Chapter 4.3) with an expert-rating. This model, as shown in Figure 4, has also been used to develop a student test, assessing their knowledge on problem solving and strategy knowledge.

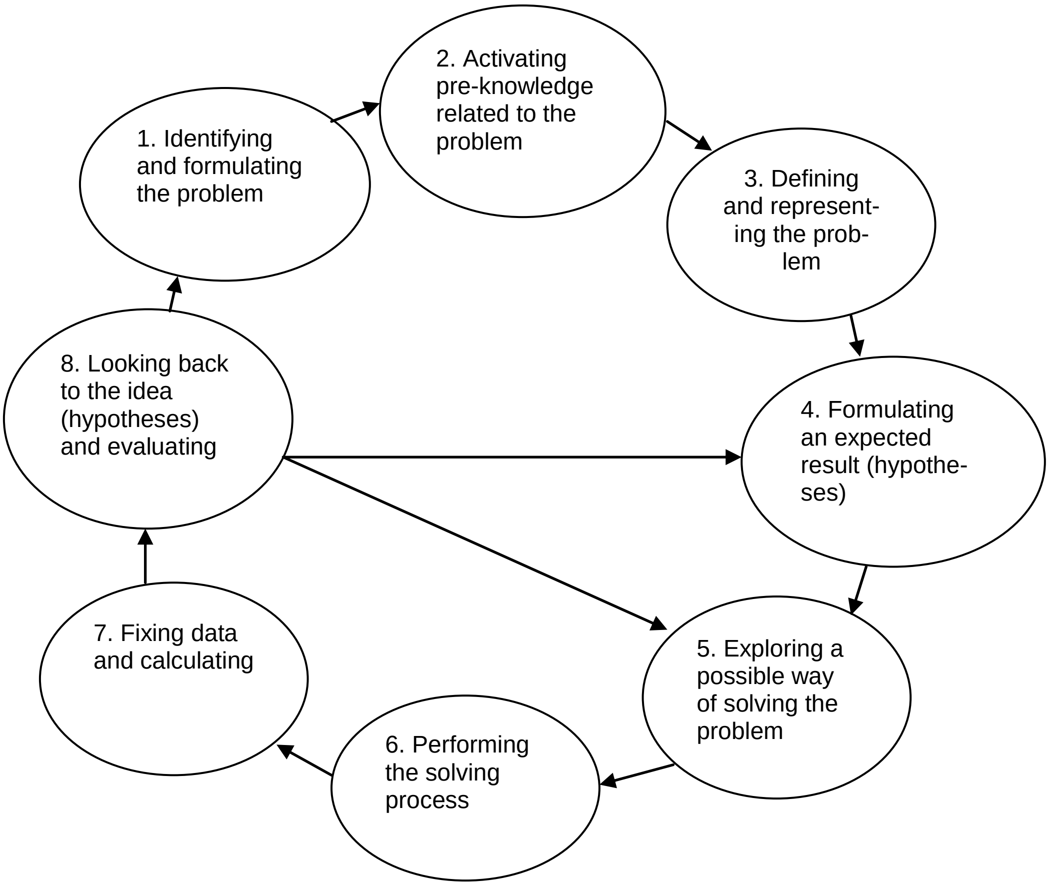


Figure 4. Model of problem solving to be used within the study

Here, a definition for problem solving is still needed. Before defining problem solving, it is important to know what *problem*, in the context of education, means. Posamentier and Krulik (2009) defined a problem as a situation that confronts the learner, that requires resolution, and for which the path to the answer is not immediately known. However, Robertson (2001) assumes that a problem arises when a living creature has a goal but does not know how this goal is to be reached. This study assumed that in a problem-solving situation, the experimentee ideally and typically knows the solution, but he or she does not know how to reach it (Oser & Baeriswyl, 2001).

Several researchers and much literature have defined problem solving. One good example would be the work of VanGundy (2005), which stated that if one accepts a problem as a gap between the state and desired state, then problem solving can be defined as the process of making something into what one wants it to be. That is, when you solve a problem, you transform *what is* into *what should be*. This means you have to figure out how to do something differently. You have to change the status quo. Eventually it may provide a solution which can then be analyzed retrospectively (Wragg, 2005). Solutions often elude the rational of an a priori approach. Onetcenter (2008) defined problem solving skills as developed capacities used to solve novel, ill-defined problems in real-world settings. Problem solving skills are then said to involve identifying complex problems and reviewing related information to develop and evaluate options and implement solutions.

According to Klahr and Dunbar (1988), Oser and Baeriswyl (2001) and PISA 2003 (OECD/PISA, 2003b), problem solving can be defined as an individual's capacity to use cognitive processes (e.g. identify and formulate a problem, define and represent a problem, explore various avenues of solving a problem, look back to the idea [hypotheses] and evaluate) in cross-disciplinary situations where the solution path is not immediately obvious and where the content areas or curricular areas that might be applicable are within a single subject area of science.

Describing students' experimental strategy knowledge

Imparting scientific inquiry-related skills to students is a primary goal of education systems in industrialized countries. For example, in the American national curriculum standards for science, the goals of inquiry skill development for grades 5–8 are the following (National Research Council, 1996):

- To identify questions that can be answered through scientific investigations;
- To design and conduct a scientific investigation;
- To use appropriate tools and techniques to gather, analyze, and interpret data;

- To develop descriptions, explanations, predictions, and models using evidence;
- To think critically and logically to make the relationships between evidence and explanations

Although the importance of teaching students to design, execute, and evaluate controlled experiments is emphasized in national standards, the methods of teaching these concepts and procedures are not well specified (Toth, Klahr & Chen, 2000). Thus, if the process of experimentation is meant to be a learning goal of science education, there is the need to identify if the goal has been achieved (Emden & Sumfleth, 2012). That means that identifying the definition of experimental strategy knowledge also is needed. This study also aims to describe students' experimental strategy knowledge because strategy knowledge and problem solving are both used when students conduct experiments.

According to Toth, Klahr and Chen (2000), scientific experimentation is the ability to use the control of variables strategy (CVS). Procedurally, CVS refers to the method used when experimenting with a phenomenon that has many variables. To avoid confounded experiments, all variables but those under consideration must be controlled. Conceptually, CVS is a necessary strategy in designing unconfounded experiments and in determining whether a given experiment is controlled.

However, Weinstein and Mayer (1986) have defined strategy knowledge as behavior and thoughts in which a learner engages and which are intended to influence the learner's learning process. Thus, every learner selects, acquires, organizes, or integrates new knowledge to reach a certain conceptual goal.

Schauble, Glaser, Ranghavan and Reiner (1992) classified three main types of knowledge about experimentation strategies, and they are strategies specific to generating evidence, interpreting evidence, and managing the memory requirements of the task. Recognizing that both structures of knowledge and strategies of experimentation are fundamental to scientific reasoning (Schauble, Glaser, Ranghavan & Reiner, 1991), several researchers have explored the relations between knowledge and strategy in scientific discovery and using experimental methodologies. Some of this work (e.g. Chen & Klahr, 1999; Klahr & Dunbar, 1988; Kuhn, 2005; Schauble et al., 1991, 1992) has contributed to more detailed and elaborated descriptions of the strategic component, particularly with new findings concerning how strategic processes interact with subjects' hypotheses. In addition, those studies showed also how a direct instruction on strategy knowledge led to significant improvement in students’ ability to design simple and unconfounded experiments (e.g. Chen & Klahr, 1999).

Considering the aforementioned summary, and keeping in mind Klahr and Dunbar's (1988) model of scientific discovery as dual search (SDDS), and the definitions of Wirth and Leutner (2006), Thillmann (2008) and Thillmann, Künsting, Wirth & Leutner (2009) for experimental strategy knowledge, this study further defines experimental strategy knowledge. Experimental strategy knowledge is defined in this study as knowledge about strategies to systematically identify new information (e.g. finding out relationships between variables by using the control-of-variables strategy) and strategies to systematically integrate new information into one’s own knowledge base (e.g. storing relationships between variables in one’s long term memory by elaborating on them).

2.4 Fostering students’ motivation towards science

To be motivated means to be moved to do something. A person who feels no inspiration to act is thus characterized as unmotivated, whereas someone who is activated toward an end is considered motivated. People have different levels and kinds of motivation. This means that they vary in the level and in the orientation of a certain motivation. Orientation of motivation is related to the underlying attitudes and goals (Ryan & Deci, 2000).

A student could be motivated to learn new skills because he or she understands their potential utility or value or because learning the skills will yield a good grade and the privileges a good grade affords (Ryan & Deci, 2000). Deci and Ryan’s theory of motivation of 1985 focuses on three basic needs: competence, self-determination, and social affiliation. The awareness of an imbalance in any of these needs or amongst these needs creates goals or activities and releases the energy for the related behavior to regain equilibrium (Fischer & Horstendahl, 1997). Deci and Ryan (1985) have distinguished between two types of motivation: (1) Intrinsic motivation refers to doing something because it is inherently interesting or enjoyable. (2) Extrinsic motivation refers to doing something because it leads to a separable outcome.

According to Schiefele, Krapp and Winteler (1992) there are three broad classes of factors that are considered to be especially relevant to a successful prognosis of academic success: (1) general cognitive factors, (2) general motivational factors, and (3) specific preferences for particular subject areas. The latter group is commonly referred to as *interests*. "The concept Interest has often been associated with intrinsically motivated behaviors because students seem to adopt these behaviors out of interest" (Deci, 1992, p.45).

The interest of students very often contradicts the interest of the learning environment. The students feel forced to carry out certain activities that are in contrast to his or her actual interest (Fischer & Horstendahl, 1997). This attempt can describe the situation in Egypt. The Ministry of Education found that secondary school students enroll in the liberal arts increasingly more than in the sciences (UNESCO, 2008), which can be taken as an indicator of lower interest in sciences. Newburghl (2008) states that the loss of science students is an alarming trend and is seen as an extremely relevant issue for developing countries.

According to Gilbert (2002), the teachers' activities in the classroom have an influence on the students' behavior and motivation for learning. For instance, a student may start a lesson with low interest in science in general, but the teacher's learning offers and methods, the structure of the lessons and interactions with students can have an effect on student motivation for learning (Helmke, 2003). This means we cannot separate strategies of learning from motivation to learn. Accordingly, the teacher has to find strategies to foster students' motivation for learning, which may at the end of a longer period also influence the students' interest in science (Gilbert, 2002).

Student-oriented instruction is seen as one factor to motivate students in earlier stages of schooling towards the study of science (Keziah, 2010). So, the use of problem solving as instructional strategy is seen as one way to foster students' achievement and motivation which have been found to positively impact students' outcomes (e.g. Faessler, Hinterberger, Dahinden & Wyss, 2006; Gök & Sılay, 2010; Keil, Haney & Zoffel, 2009; Mason & Singh, 2010; Sekuk, Caliskan & Erol, 2008). With this in mind, this study aims to develop and evaluate a teaching program based on an established model of problem solving to foster students' motivation towards science and to increase students cognitive outcomes of related lessons (OECD, 2009). The direction of the relation between students' motivation and knowledge or competence in science is not a matter of this study.

2.5 Summary of the theoretical background

The main goal of the presented chapter was to outline the framework of the study. This study focuses on science teaching in the first year of middle school, which is Grade 7 in the Egypt school system. The future vision of education in Egypt focuses on fostering scientific inquiry and problem solving skills. Egyptian students poorer science achievement in TIMSS 2003 and 2007 made clear that students' motivation needs to be fostered and examined in the earlier grades (Ministry of Education (MOE), 2010). It is assumed that this is directly related to instruction quality. To begin to deal with this, the presented study therefore focuses on developing and evaluating a teaching program using a model for quality of instruction. Moreover, it aims to develop or adapt six instruments to assess seventh grade Egyptian students' outcomes and evaluate teachers' quality of instruction. This study considers several variables that might influence students' outcomes. These include teachers' quality of instruction, students' perceptions on the quality of science lessons and their individual learning premises (e.g. understanding of the topic, motivation, cognitive abilities, gender, age, background, problem solving, experimental strategy knowledge), and students' outcomes (e.g. problem solving, experimental strategy knowledge, achievement, perceptions on the quality of science teaching and motivation towards science).

2.6 Aims and research questions

The main goals of this study, based on the theoretical background, are to use a model for quality of instruction to develop and evaluate a teaching program based on a model for problem solving. Moreover, the study should show how the inquiry/discovery process not only enlivens lessons based on or driven by laboratory investigations, but also builds deeper understandings of science content. This instructional approach offers a framework for structuring lessons so that students can and will become more deeply engaged and take greater interest, by hands-on activities, in their learning (Holstermann, Grube & Bögeholz, 2010; Sutman et al., 2008), and that teachers have a guideline for planning and performing a lesson accordingly. The related effect must be assessed in a future study. Moreover, the study will develop or adapt six instruments to assess seventh grade Egyptian students' problem solving abilities, experimental strategy knowledge, achievement, perceptions on the quality of science teaching and motivation towards science and evaluate teachers' teaching processes.

To develop the teaching program, the first goal is to develop lesson plans for the topic density and buoyancy, as well as phases of teaching. According to Van de Walle, Karp and Bay-Williams (2010), content and task decisions are often overlooked when lessons are planned without considering the content expectations and the needs of students—yet this is the most important part of the planning process. This study focuses here on designing activities based on a model for problem solving (Flick & Lederman, 2006) of the topic density and buoyancy for seventh grade students that accomplish the goals outlined in appendix 8.6 for the five double lessons. As Bentley and Watts (2005) showed that before beginning problem solving, we should ask the following questions: a) does one first deliberately teaches all the facts, concepts and skills so that the students will have all the relevant information at their fingertips? b) Or does the teacher use the motivating power of problem-solving as a means of allowing the students to decide and satisfy their own knowledge needs? Furthermore, learning to solve problems requires practice in solving problems (Jonassen, 2011), not just learning about problem solving.

Thus, the teaching program, in this study, contains three teaching phases, and they are:

(1) learning about density and buoyancy.

(2) Learning about problem solving and strategy knowledge.

(3) Training on problem solving and experimental strategies.

These lesson plans were constructed for the topic of density and buoyancy. This topic was chosen because it is part of the Egyptian curriculum and because students tend to have difficulties learning about this topic. In addition, the topic provided a broad basis for problem solving tasks with many already well-established experiments. The lesson plans planned according to conditions in Egyptian schools, cooperative learning principles (Slavin, 2011), the proposed model of problem solving and the bases of lesson plan construction from previous literature (see D'Amico & Gallaway, 2010; Feasey, 2007; Frey & Fisher, 2011; Newton & Newton, 2008; Shaw, 2008, Vargas, 2009). In order to assess students' problem solving abilities, experimental strategy knowledge, achievement, perceptions on the quality of science teaching and motivation towards science, the first goal is to develop or adapt six instruments accordingly to the conditions in Egypt. To use those items in an intervention study in Egypt, the reliability and validity of instruments have to be verified on an Egyptian sample. The control variables of age, gender, academic pre-knowledge, background and cognitive abilities are also assessed. Based on the aims of this research, the following questions are developed:

1. Does the teaching program that is based on the problem solving model (within the topic of density and buoyancy) have an impact on students' achievement in terms of content knowledge, problem solving abilities and experimental strategy knowledge?
2. Is there an influence of the teaching program on students' motivation?
3. Are there effects of the teaching program on students' perceptions about science lessons?

2.7 Hypotheses

The primary goal of instructional design is to improve the quality of learning and instruction. Instructional designers have focused on a number of areas of critical concern and developed a variety of techniques to achieve this goal (Seidel, Perencevich & Kett, 2007). While it is debatable whether any new program or practice is ever exactly replicated by users in different settings (Anderson, 2010), the idea behind this study is that the quality of instruction could be substantially improved if teacher offers learning opportunities. These learning opportunities have to be adequate to students' premises (e.g. abilities and motivational prerequisites). Moreover, teachers' teaching process can be changed, using the proposed teaching program, in order to foster students' motivational, problem solving abilities, experimental strategy knowledge, perceptions on the quality of science lessons and content-related achievement.

Several researchers have found a positive relation between using problem solving as independent variable and students' outcomes (see Afamasaga-Fuata'i, 2009; Gök & Sılay, 2010; Goldberg, Otero, Robinson, 2010; Keil et al., 2009; Mason & Singh, 2010; Sekuk et al., 2008; Valiotis, 2008). The current study developed a teaching program in order to foster students' problem solving abilities, experimental strategy knowledge, achievement and motivation towards science. The quality of instruction was assessed by measuring students' perceptions on the quality of science lessons and evaluating teachers' quality of instruction. Hence, the hypotheses for this study are as follows:

<u>Hypotheses to Research Question 1</u>

H1.1. Relative to the control group, the intervention group would exhibit larger gains in science achievement.

H1.2. Relative to the control group, the intervention group would exhibit larger gains in problem solving abilities.

H1.3. Relative to the control group, the intervention group would exhibit larger gains in experimental strategy knowledge.

Hypothesis to Research Question 2

H2. There are differences between the control group and intervention group regarding the increase of motivation towards science. Students in the intervention group will perform significantly better.

Hypotheses to Research Question 3

H3.1. There are no significant differences between the control and intervention group regarding the teachers' quality of instruction.
H3.2. Relative to the control group, the intervention group would exhibit more positive perceptions on the quality of science lessons.

Hypotheses H1.1 to H.2 and H3.2 will be assessed using t-test and Mann–Whitney U-test for independent samples (Field, 2009), comparing the learning gains of each instrument between the control and intervention group. The hypotheses will be confirmed if differences between the two groups are significant on an Alpha $< .05$ level with an overall small and medium effect size ($.1<r<.5$ & $.1<d<.8$) based on this study's small sample size (Jackson, 2009).

In order to prove the last hypothesis, H3.1, an independent-sample t-test will be used, and the mean of teacher quality of instruction between the control and intervention group will be analyzed. The result of this test has in past studies proven that there was no significant difference between the means of these two samples (Field, 2009).

In the next chapter, the design of this study will be illustrated. This design includes the following sections: the study's schedule, information on sample and data collection, the treatment, which was implemented with intervention and control groups, and finally, the differences between intervention and control groups.

3. Design

In this chapter the design of this study will be explained. This design includes the following sections: the study’s planning and organization, information on sample and data collection, and finally, the treatment, which was implemented with intervention and control groups.

3.1 Planning and organization

The study was conducted in Aswan, Egypt, in the autumn of 2011. The study started in February 2009 and had a period of four years. The following figure shows the chronology of the teaching program design, instrument development, pilot and main studies, and data analysis:

<table>
<tr><td>2009</td><td>2010</td><td>2011</td><td>2012</td></tr>
<tr><td>February - September
Program design & in-struments development</td><td></td><td></td><td></td></tr>
<tr><td colspan="2">October 2009 - March 2010
Translation and editing processes</td><td></td><td></td></tr>
<tr><td></td><td>October – November
Pilot study for paper-pencil-instruments</td><td></td><td></td></tr>
<tr><td></td><td colspan="2">December 2010 – March 2011
Data analysis of pilot study</td><td></td></tr>
<tr><td></td><td></td><td>September – October
Intervention study</td><td></td></tr>
<tr><td></td><td></td><td colspan="2">November 2011 - February 2012
Data analysis</td></tr>
</table>

Figure 5. Schedule of the study

3.2 Sample and data collection

From September to October 2011, an intervention study was conducted on a sample of seventh grade classes. The sample was selected from two schools (two classes of boys and two classes of girls). The sample included 147 students in a general middle school in the city of Aswan, Egypt during the 2011-2012 school year. The control group contained two classes, one with 36 boys and the other with 37, both of whose students had a mean age of 12.35 years, while the intervention group consisted of two classes: one with 37 boys and the other with 37 girls; with a mean age of 12.43 years.

This study was planned as a quasi-experimental study (e.g. Babbie, 2011; Shadish et al., 2002) with a classic pre-post design for evaluation of the intervention. Two treatments were used with the control and intervention groups: the newly developed teaching program, as described in Chapter 1, p.5, was used to teach the intervention group while the control group was taught according to regular teaching methods, as described in Chapter 1, p.5. Both groups were taught by the same teacher (the author, who is a qualified science-teacher) and to ensure the comparability of teaching in both groups, an observer sheet was used.

The study was conducted during a six week-long unit at a middle school in Egypt. The lesson plans of the teaching program were constructed based on a model for problem solving (see Figure 4, p.21), which combined established problem solving models (Klahr & Dunbar, 1988; Oser & Baeriswyl, 2001) according to the requirements in Egyptian science classes. The program itself consisted of five double lessons (one per week) which provided content knowledge on the topic *density and buoyancy* and introduced the method of problem solving. This topic was chosen because it is part of the Egyptian curriculum and because students tend to have difficulties learning about this topic. In addition, the topic provided a broad basis for problem solving tasks with many already well-established experiments.

Six instruments were developed/adapted assessing students' problem solving abilities, experimental strategy knowledge, achievement, perceptions on the quality of science teaching, motivation towards science, and finally teachers' quality of instruction. The assessment methodology used for these instruments was criterion-referenced (Peers, 2006). To test the quality of instruments, a pilot study was conducted. The instruments were tested with Egyptian science students (n=44, 7^{th} grade). A quantitative research method was used in this study. The data from the six instruments were analyzed according to the Rasch model and classical methods. (Engelhard, 2011; Meyer, 2010) and final instruments were compiled. Intervention and control group students were compared to determine possible significant differences in learning outcomes or abilities. Results of the quality of instruments in the pilot study and the learning effects of this program on students' dependent variables in the main study will be reported in Chapter 5.

3.3 Treatment in the intervention and control groups

In this section, a description of the lesson plans for the teaching program based on the model of problem solving, which was used with the intervention group, will be provided. It will be followed by a description of the treatment within the control group.

3.3.1 Treatment in the intervention group

Students' learning of scientific inquiry and problem solving can be enhanced by providing them with an explicit model of problem solving. This approach is particularly effective for lower-achieving students (White, Frederiksen & Collins, 2009). This study offers a framework for structuring lessons so that students can and will become more deeply engaged and take greater interest in their learning through the application of experimental activities (Holstermann et al., 2010). In addition, teacher changes away from memorizing the facts to guide students to use and transfer their knowledge in new situations. Moreover, teachers have a guideline for planning and performing a lesson accordingly. Figure 6 illustrates the structure of the intervention-lesson plans.

No. of lessons	Learning goals	Time in min.
1	Achieve content knowledge in the topic density and buoyancy	90
2		90
3	Learn about problem solving and strategy knowledge	90
4	Training on problem solving and experimental strategies.	90
5		90

Figure 6. Intervention - Lesson plans

The following figures (7, 8, 9) describe how the three programs' phases (for more details see Appendix 8.6) were structured and sequenced based on aforementioned aspects (see Chapter 2.6, p. 28). They show how content knowledge and problem solving were taught and trained during this study.

Title: Density and buoyancy		Grade: 7
Curriculum: Science	Concepts and Competencies: Density	Duration: 90 minutes (One session)
Description: Students will determine whether various objects sink or float in water. Whether an object sinks or floats in a liquid depends mainly on Archimedes´ Principle and density.		

Figure 7. Example – One lesson about Archimedes' principles and buoyancy

Title: Learning about problem solving and strategy knowledge		Grade: 7
Curriculum Area: Science	Concepts and Competencies: Problem solving	Duration: 90 minutes (One session)
Description: This lesson is an explanation of the stages of scientific inquiry and problem solving, and how they can be implemented in a scientific situation.		

Figure 8. Example – One lesson on learning about problem solving

Title: Problem solving activity		Grade Range: 7
Integrated Curriculum Area: Science	Concepts and Competencies: buoyancy, density and sinking or floating	Duration: 90 Minute (One session)
Description: This lesson is an example of how we can implement the practice of scientific inquiry and problem solving in the context of swimming and sinking.		

Figure 9. Example – One lesson on training of problem solving

3.3.2 Treatment in the control group

The following procedures were implemented in the control group during the intervention study:

- The topic of density and buoyancy was presented via regular lecture format as a teacher-centered presentation.
- All experiments and problem solving processes were done by the teacher (the author) while students only took notes.
- The control group received the model of problem solving using a regular lecture format.
- The same amount of time was used with the control group (450 min.) to learn the content and to solve problems.
- Students worked individually.

Based on the above explanation of the treatment, which was implemented with both intervention and control groups, the differences between the treatments of the intervention and control groups can be summarized as follows:

Table 2. The differences between the treatments of the intervention and control groups

Categories	Control group (CG)	Intervention group (IG)
Instruction approach	Teacher-centered approach	Student-centered approach
Teaching process		
Teacher role	- explains and does all experiments - solves problems	- monitors student performance and provides extra help to students who have difficulty or query - organize the discussion for students
Students role	take notes and monitor the teacher´s work	do scientific inquiry and solve problems
Work form	individually	In groups
No. of lessons	Five double lessons	
Time	450 minutes	
Solving the problems	Using the model of problem solving	
Materials and resources	The same materials and resources in both groups	

3.4 Summary of the design

The schedule of this study was first explained, followed by sample and data collection. The two main treatments for the intervention and control groups, including the differences between the treatments of the control and intervention groups were described in closing. This study has been planned as a quasi-experimental study with a classic pre-post design for evaluation of the intervention. Both groups were taught by the same teacher (the author, who is a qualified science-teacher). The teaching program, which was implemented with the intervention group, contained three teaching phases. These included: learning about density and buoyancy, learning about problem solving, and training on problem solving abilities.

4. Methodology

Described in this chapter is the methodology of the presented study. This methodology includes the following sections: the development of the instruments, translation and editing processes, validation of instruments, and overview of planned statistical analysis. In the presented study, six instruments are essential: three tests assessing students' experimental strategy knowledge, problem solving abilities and achievement, two questionnaires assessing students' perceptions on the quality of science lessons and motivation towards science, as well as one questionnaire assessing teachers' quality of instruction.

4.1 Instruments

The instruments used in this study were in part adapted from previous studies and in part self-developed. A special challenge was the items' original languages – English and German – and their translation into Arabic. The following table gives an overview of instruments and their language:

Table 3. Overview of instruments

Instrument	Original Language
Science achievement test (SAT)	English
Problem Solving test (PST)	English
Experimental strategy knowledge test (ESKT)	English
Motivation questionnaire (MQ)	German
Perceptions questionnaire (PQ)	English
Teachers' quality of instruction questionnaire (QI)	German

In order to use those items in an intervention study in Egypt, the validity and reliability of instruments had to be validated by experts and on an Egyptian sample (see Chapter 4.3 & 5). Described and depicted in this section are the development of the paper-pencil instruments and how the quality of instruments was tested.

4.1.1 The Science Achievement Test (SAT)

The aim of this test is to assess students' content knowledge on the topic *density and buoyancy*. This instrument contains n=30 multiple-choice items containing an item-stem, one attractor and three distractors. Items were developed in order to cover different cognitive activities and to estimate the difficulty of the items a posteriori based on the six categories of Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001). These categories include remembering, understanding, applying, analyzing, evaluating, and creating. Some of the items were self-developed and some of them were adapted based on the PISA and TIMSS tests. The items' difficulties should increase from understanding to creating. Table 4 illustrates how the 30 items are distributed over the categories of Bloom. Difficulty is expected to increase from remembering to evaluating.

Table 4. Science test items according to Bloom's Revised Taxonomy

Category	Items	Test items percentage
Remembering	11	36.66 %
Understanding	6	20 %
Applying	4	13.33%
Analyzing	5	16,66%
Evaluating	4	13,33%
Sum.	30	100%

To solve an item successfully, a student has to tick the correct answer. The test was analyzed dichotomously; if a student ticks a distractor, zero points were given.

4.1.2 The Experimental Strategy Knowledge (ESKT) and Problem solving (PST) Tests

The aim of this test was to assess students' knowledge and abilities concerning experimenting and problem solving. The test was developed according to newly performed studies in Germany (e.g. Marschner, 2011; Thillmann, 2008; Wirth & Leutner, 2006), and the science subject of density and buoyancy. To solve an item in this test (see sample of tasks in Figures 10 and 11), students had to rate different courses of action (items) according to a given situation (tasks) on a five-point Likert-scale from *strongly agree* to *strongly disagree*. In each item, a problem situation was described and students' were asked to rate alternatives.

In the experimental strategy knowledge part, the tasks were related with each other to identify new information (e.g. finding out relationships between variables by using the *control-of-variables strategy*) and to integrate new information into one's own knowledge base (e.g. storing relationships between variables in one's long-term memory by elaborating on them, see Figure 10). Experimental strategy knowledge and problem solving abilities were assessed within one test booklet.

Tasks in the problem solving part of the test included identifying and formulating a problem, exploring various avenues of solving a problem, looking back at the idea (hypotheses) and evaluating the problem solving process. Tasks in experimental strategy knowledge and problem solving were all based on the topic of density and buoyancy. Students then had to rate a list of possible action alternatives on a scale from strongly agree to strongly disagree (Figure 11). Including five items for experimental strategy knowledge and four items for problem solving, this multiple-choice test was constructed based on the model of problem solving outlined above (see Figure 4, p. 21). The following items are examples from the student test:

You want to find out on which variables the property of buoyancy (whether a body sinks, floats or swims in water) is dependent.

You consider the following procedures in order to answer the question. Review the procedures with the scale from *strongly agree* to *strongly disagree*:	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
I make a drawing with all possible variables.	□	□	□	□	□

Figure 10. Example - one item of experimental strategy knowledge

Remember that the force of buoyancy is the upward force caused by fluid pressure that keeps things afloat. Something that is positively buoyant floats in water. Something that is negatively buoyant sinks in water. Things that are neutrally buoyant neither float nor sink in water. With regard to buoyancy, two important aspects of objects are mass and volume. In an experiment, it is shown that a heavy steel model of a ship floats, but a heavy solid block of steel sinks in freshwater. Which procedure can explain the observation?

You consider the following ideas in order to answer the question. Review the procedures with the scale from *strongly agree* to *strongly disagree:*	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
I measure the mass of the steel ship model and the mass of the solid block and compare them.	□	□	□	□	□

Figure 11. Example - one item of problem solving

Test items on experimental strategy knowledge and problem solving required a different type of answering and data editing process. Students had to rate courses of actions, and those ratings had to be summed up to a total score for each student. Therefore, a group of five experts, all Ph.D. students in science education, rated the problem solving and strategy knowledge items according to the definitions as described above (see definitions of problem solving and experimental strategy knowledge on theoretical background, pp. 19-25). Having five expert ratings for each task, the median for each task was estimated and compared pair-wise with students' answers.

The medians of the experts' ratings were then calculated (Field, 2009). In order to measure the students' scores, the results were scored as *0* (no agreement between student and experts) or *2* (full agreement between student and experts). This degree of agreement is based on pair-wise comparisons (Aczbl & Saaty, 1983). For example, if the experts say that in Task 1, alternative (a) is better than alternative (b), then the student gets two points for the same order. If the student says that alternative (a) is equally good as alternative (b), a one-point is awarded, and if the student says alternative (b) is better than (a) zero points are awarded. This was done for all tasks. At the end, all points were summed up and divided by the number of comparisons. Finally, we had 62 pair-wise comparison items for the problem-solving test and 35 pair-wise comparison items for the experimental strategy knowledge test.

4.1.3 The Motivation Questionnaire (MQ)

The fourth instrument used in this study was a questionnaire assessing the students' motivation towards science instruction. This instrument consisted of 60 items in German in a 4-point and 6-point Likert-scale format and was adapted from the project of quality of instruction in physics (QUIP, university Duisburg-Essen) and PISA 2006 to the conditions in Egypt (e.g. integrated science courses, gender, middle school, period time, age) without changing the meaning of the items. Students had to rate different statements (tasks) according to their occurrence within various issues related to science, careers and <broad science>, and teaching and learning science. In this questionnaire, students have to answer by ticking a box. The example item below gives an idea of the structure of this instrument.

How much do you agree with the statements below?						
(Please tick only one box in each row)						
	SA	A	TA	TD	D	SD
I like reading about science topics.	□	□	□	□	□	□

Figure 12. Example - One item of motivation questionnaire

4.1.4 Questionnaires assessing quality of instruction

To assess quality of instruction, the analysis of videotaped lessons has become a powerful instrument. In the planning phase of the project, videotaping all intervention- and control group lessons was considered. Due to educational and political circumstances in Egypt, however, it was not possible to conduct a video study. Instead questionnaires were used to ensure the comparability of instructional quality between intervention- and control groups. For this, quality of instruction was assessed using two questionnaires: quality of teachers' instruction and students' perceptions on the quality of science teaching.

The teachers' Quality of Instruction Questionnaire (QI)

The aim of this questionnaire was to assess teachers' quality of instruction for control and intervention group. This instrument consists of 112 items in German in a 4-point Likert-scale format. It was adapted from the Clausen (2002) questionnaire. During the teaching period for each lesson, a second teacher had to evaluate instructional quality by rating different statements (items). This questionnaire has the following scales: rule clarity, discipline, repeat practice, challenge practice, diagnosis of competence in social aspects, positive attitude of students, diagnosis of competence in achievement, individual reference of attitude, classroom management, volatility, motivation, clarity and structure of instruction, interaction rate, pacing, amount of wasted time in instruction, instruction's individualization, time on task, aggression towards teacher, aggression towards classmates, scientific production of the students, error culture, structuring aids and individual learning support. Each scale was assessed with an average of five items. The example item below gives an idea of the structure of this instrument.

Scale: Rule Clarity	Strongly not correct at all	Incorrect	Mostly correct	Strongly correct
The rules that must be met are known for all students.	□	□	□	□

Figure 13. Example - One item of quality of instruction questionnaire

<u>The Students' Perceptions on the quality of Science Teaching Questionnaire (PQ)</u>

Having assessed the quality of instruction by an objective observer, the presented study also assessed students' perception of instructional quality. Previous studies indicate that both views on classroom interactions might strongly differ from each other (Clausen, 2002). The students' perception on the quality of science lessons was also assessed with a questionnaire. This instrument consisted of 24 items in English. It was adapted from TIMSS 2007 (Olson et al., 2008) and Reyer (2004) questionnaires on the topic of density and buoyancy. This questionnaire was used in a pre/post design. Students were asked to rate different statements (tasks) in a 4-point Likert-scale format according to their occurrence within science lessons before and during the intervention. Again, students had to tick one box. The example item below gives an idea of the structure of this instrument.

How often do you do these things in your science lessons? (Fill in one box \ for each line)				
	Every or almost every lesson	About half the lessons	Some lessons	Never
We work in small groups on an experiment or investigation.	□	□	□	□

Figure 14. Example - One perception questionnaire item

4.2 Translation and editing processes

After constructing the paper and pencil instruments this study faced the challenge of translating the items into Arabic based on the unique characteristics of the educational systems of each country, language and cultural differences each of which must be carefully taken into consideration to ensure the comparability of the data (Arora, Ramírez & Howie, 2006). In order to ensure a high quality of translation (see Brislin, 1970, 1980), two steps were taken:

- The four English instruments (science test, experimental strategy knowledge test, perceptions on the quality of science lessons questionnaire and problem solving test) were translated from English into Arabic by an Arabic native speaker. Two instruments (teachers' quality of instruction and motivation questionnaire) were translated from German into Arabic again by an Arabic native speaker.
- The validity of the Arabic translations was measured. The Arabic versions were translated back into English and German with other native speakers' support. The quality and accuracy of all translations were verified with the help of a native Arabic speaker who specializes in English at university, a native English speaker and native German speakers.

After the translation process was finished, various Arabic words were found that could be used for one English word. To solve this problem, three phases of comparison, with native speakers in Arabic, English and German, were done, with an acceptable inter-rater agreement of Cohen’s Kappa between 0.41-0.62 (Wirtz & Caspar, 2002), and the following steps:

- The original English (E1) and German instruments’ words (achievement test, experimental strategy knowledge test, problem solving test, perceptions on the quality of science lessons questionnaire, teachers’ quality of instruction questionnaire and motivation questionnaire) were compared with the translated English (E2) and German instruments’ words by researchers and with the help of a native English speaker and two native German speakers.
- The Arabic translated instruments’ words (achievement test, experimental strategy knowledge test and problem solving test, perceptions on the quality of science lessons questionnaire, teachers’ quality of instruction questionnaire, motivation questionnaire) were compared with the original English instruments’ words by researchers and with the help of a native Arabic speaker.
- ΔE_{12} (comparison between the original English and translated English words) and ΔEA (comparison between the English and Arabic words) were measured. ΔE_{12} was positive (+) when the two English words matched and negative when they did not. In this case, a look at ΔEA was needed and if it was also found to be negative, the Arabic word would need to be modified. If ΔEA was positive, the original English and Arabic words matched and the Arabic word could be accepted.

4.3 Validation of the instruments

In this section, the method for validation will be discussed as an important aspect about the quality of instruments.

One important aspect of a test or scale is its validity (Peers, 2006). “A more critical test of validity is called criterion validity, which concerns whether the measure (a) can accurately forecast some future behavior or (b) is meaningfully related to some other measure of behavior” (Goodwin, 2010, p.132). “The reason for gathering criterion validity evidence is that the test or measure is to serve as a “stand-in” for the measure we are really interested in” (Kaplan & Saccuzzo, 2009, p. 138).

To ensure the criterion validity of the six developed/adapted tests and questionnaires of the science achievement, problem solving, experimental strategy knowledge, motivation towards science, perceptions on the quality of science lessons and teachers' quality of instruction, an expert rating of all instruments was conducted, including 14 Ph.D. students, three post-doctoral students and five teachers, all of whom specialized in science education in Germany or Egypt. Each participant received a sample of instruments and reviewed the instruments' items according to the six categories of Bloom's Revised Taxonomy (see p. 39), problem solving and experimental strategy knowledge definitions (see pp. 19-24), PISA 2006 (OECD, 2009), Reyer (2004) and Clausen (2002) scales of motivation towards science, perceptions on the quality of science lessons and teachers' quality of instruction (see pp. 41-44), and commented on a checklist related the correctness of the items. To estimate the criterion validity of the model for problem solving (see Figure 4, p. 21), three post-doctoral students as experts in science education in Germany reviewed the proposed model according to the steps of Klahr and Dunbar (1988), and Oser and Baeriswyl (2001) and commented on a checklist related to the quality of the model. To judge the quality of the lesson plans (see Appendix 8.6), a group of eight experts was consulted. These experts included three post-doctoral students and five teachers, all of whom also specialized in science education in Germany or Egypt. Each participant also received a sample of lesson plans and commented on a checklist related to the correctness of content and the appropriateness of timing, activities, sources and materials of each lesson.

The result of the inter-rater agreement of the experts rating is acceptable with Cohen's kappa between $.58<\kappa<.87$ (Elliott & Woodward, 2007). For the science achievement test items, the agreement is between $.58<\kappa<.81$, whereas, for the problem solving and experimental strategy knowledge tests items, the agreement is between $.63<\kappa<.85$ and $.72<\kappa<.86$ respectively. For the motivation towards science instruction questionnaire items, the agreement is between $.83<\kappa<.86$. For the perceptions on the quality of science lessons questionnaire items, the agreement is between $.78<\kappa<.84$. For the teachers' quality of instruction questionnaire items, the agreement is between $.85<\kappa<.87$. For the model of problem solving the agreement is between $.79<\kappa<.87$. For the lesson plans the agreement is between $.81<\kappa<.86$. Some minor modifications were made to the lesson plans and instruments based on the comments provided. An expert review process for the test items was performed in addition.

4.4 Overview of statistical analyses

After describing the development of the instruments, the translation and editing processes and the validation of instruments, an overview of the planned statistical analyses will be given. As described in Chapter 3.2, for data collection, a quantitative research methodology was used in this study.

In the following two sections, an explanation of the applied instruments' quality criteria, which was used in the present study, will first be described. Following that will be an explanation of the procedures of inferential statistics, which were used to investigate the impact of the teaching program based on the learning gains on students' problem solving abilities, experimental strategy knowledge, achievement, perceptions on the quality of science lessons and motivation related to science.

4.4.1 Criteria of measurement quality

To ensure the quality of the instruments' items, item response and classical theories were used. In the following, an overview of these theories is given and the applied instruments' quality criteria are explained.

<u>Item response theory</u>

Item response theory IRT focuses on persons' responses on individual items. The response of an examinee is assumed to depend upon one or more factors called traits. Each item is used to measure the underlying trait. And the relationship between persons' performance on an item and the trait underlying item performance can be described by a monotonically increasing function (De Gruijter & Van der Kamp, 2008). In contrast to classical test theory, IRT consists of a class of mathematical models for which estimation procedures exist for model parameters (i.e., person and item parameters) and other statistical procedures for investigating to what extent the model at hand fits the data or persons' responses to a set of items (De Gruijter & Van der Kamp, 2008).

The simplest IRT model is often called the Rasch model (Hambleton, Swaminathan & Rogers, 1991). To investigate whether or not the predicted trait was defined by the test items, Rasch analysis (Boone & Scantlebury, 2006) was conducted on the data from the experimental strategy knowledge test, the problem solving test, the science achievement test, the perceptions questionnaire on the quality of science lessons and the motivation questionnaire towards science because it offers better results with sufficient stability and even for smaller sample sizes than do classical analyses. "Rasch analysis can handsomely handle a sample size of 25-30 to generate a sound 95% CL statistics and 50-60 for a 99% CL". (Linacre, 1994, p. 328). Infit statistics were used to assess whether a given performance or item is consistent with other performances or items (Dawson-Tunik, 2006). The acceptable range of the infit mean square statistic for each item was taken to be (MNSQ) from .5 to 1.5 for (T) from –2.0 to +2.0, (T) is only useful for salvaging non-significant MNSQ >1.5 when the sample size is small or the test length is short (Linacre, 2010). The accepted cut-off for Person Separation Indices is .7 and values above this threshold suggest higher reliability (Smith, 1998). Items for which there was evidence of potential misfit were excluded from the final analysis of data. The software of WINSTEPS® software (version 3.70.0.5) was applied for processing the data (Linacre, 2010).

Classical statistical analysis

If the instrument items showed poor Rasch-fit criteria for the partial credit model, a classical test theory was used (Wendler & Walker, 2006). Hence, to ensure the quality of items, reliability and discriminatory power of teachers' quality of instruction questionnaire items were estimated. According to the classical test theory, internal consistency of a scale measured in Cronbach's Alpha should be >.7 (Carver & Nash, 2012; Cortina, 1993) and a discriminatory power >.3 is an acceptable value

(Wendler & Walker, 2006). SPSS® software (PASW statics; version 18) was used for processing the data classically (Field, 2009).

4.4.2 Procedures of inferential statistics

To ensure the discriminant validity of the instruments, a bivariate correlation between instruments was estimated. Hypotheses concerning the impact of the teaching program on students' problem solving abilities, experimental strategy knowledge, achievement, perceptions on the quality of science teaching and motivation towards science were investigated by using the analysis of variances. In the following, the main principles of these procedures of inferential statistics are described.

Correlations

A correlation coefficient measures the degree of relationship between two or more sets of scores and can vary between -1.00 and +1.00. The stronger the relationship between the variables, the closer the coefficient will be to either -1.00 or +1.00. The weaker the relationship between the variables, the closer the coefficient will be to 0 (Jackson, 2009). The Pearson's product moment r was used for parametric tests and the Spearman's correlation coefficient r_s for non-parametric tests. The Pearson (r) was used with normally distributed data, while Spearman (r_s) was used with the non-normally distributed data (Field, 2009). These two tests indicate both the direction and the strength of a relationship between variables (Ary, Cheser & Sorensen, 2010). According to Cohen (1988, 1992 cf., Field, 2009), effect size $r \geq .10$ is indicated as a small effect, and $r \geq .30$ as a medium effect, whereas $r \geq .50$ is considered a large effect.

The method of correlating instruments is an accepted method for investigating the discriminant validity of a test (e.g. Anastasi, 1988; Borghese & Gronau, 2005). Discriminant validity occurs when scores on an instrument designed to measure a certain construct are uncorrelated with scores on other instruments that should be theoretically unrelated (Goodwin, 2010). Thus, to ensure the discriminant validity of the instruments, a bivariate correlation between instruments was estimated.

Analysis of variances

Analysis of variances through such instruments as a t-test for independent samples is usually used in situations in which there are two experimental conditions using two samples of different participants (Field, 2009). The test requires normally distributed data within each group. There is also an additional assumption called homogeneity of variance (Field, 2009). To ensure the normal distribution of data in control and intervention group, a Shapiro-Wilk test was used (Field, 2009; Peers, 2006). Compared to the Kolmogorov–Smirnov test, this test is better suited for detecting differences from normality (Field, 2009). A significant value ($\rho<.05$) indicates a deviation from normality, but this test tends to be affected by large samples in which small deviations from normality yield significant results (Field, 2009). And to ensure the homogeneity of variance between the study's control group (CG) and intervention group (IG) Levene's test was conducted (Page, Braver & MacKinnon, 2003). Therefore a significant t-test at $\rho<.05$, would increase our confidence in the assumption that the means of the two conditions are different. If, however, the t-test is non-significant (i.e. $\rho>.05$), we would not have sufficient evidence to reject the null hypothesis, which states that there is no significant difference between the means of these two samples (Field, 2009).

However, if there is no normal distribution of data within each group, special statistical procedures known as non-parametric tests should be used (e.g. Field, 2009; Spatz, 2008). Non-parametric tests work on the principle of ranking the data. The function of these tests is as follows: Finding the lowest score and giving it a rank of 1, and finding then the next highest score and giving it a rank of 2, and so on. Accordingly, the results are presented in high scores by large ranks, while low scores are represented by small ranks. This kind of tests is usually used when there is a break in the parametric assumptions (Field, 2009). Mann–Whitney U-test can be used (Field, 2009) to assess the differences between two conditions and different participants used in each condition. This test is equivalent to the parametric independent t-test (Field, 2009). Accordingly, using the t-test and U-test for independent samples is appropriate for identifying significant differences between two groups (CG

& IG), e.g. regarding students' control variables or the learning gains. In addition to the independent-groups t-test and U-test it is useful to look at the effect size, which is —"the proportion of variance in the dependent variable that is accounted for by the manipulation of the independent variable. Effect size indicates how big a role the conditions of the independent variable play in determining scores on the dependent variable" (Jackson, 2009, p. 231). Thus, by using the observed t-value (parametric test) Cohen's d was used to evaluate the effect size of the teaching program based on the learning gains on students' motivation towards science. Cohen's d interprets small effect sizes as up to .20, a medium effect size starting at .50, and a large effect size starting at .80 (Jackson, 2009). According to Rosenthal (1991, cf. Field, 2009) *z*-scores (non-parametric test) are converted into the effect size estimate, r. r was used to evaluate the impact of the teaching program based on the learning gains of students' achievement, problem solving abilities, experimental strategy knowledge and perceptions on the quality of science lessons. An effect size $r \geq .10$ is indicated as a small effect, $r \geq .30$ as a medium effect, and $r \geq .50$ is considered a large effect (Field, 2009).

4.5 Summary of the methodology

This chapter has described the methodology of the study. First explained were the development/adaptation of the six instruments, which assessing students' achievement, problem solving abilities, experimental strategy knowledge, perceptions on the quality of science lessons and motivation towards science. The translation and editing processes and validation of instruments were described in closing. Following that, an overview of statistical analysis was given.

The criterion validity of the model for problem solving used in this study, as estimated by experts has been demonstrated. As illustrated in Chapter 3.3.1, the developed model was used to construct lesson plans for teaching the topic of density and buoyancy. Further discussed was how instruments for science achievement, problem solving, experimental strategy knowledge, quality of instruction and motivation towards science could be developed and adapted for analysis use for students in Egyptian middle schools. The instruments were partly self-developed and partly adapted from newly performed studies on strategy knowledge and problem solving in Germany and the conditions in Egypt. Moreover, the criterion validity of those instruments was estimated by experts. To ensure the quality of the translation, the following steps were taken:

(1) Translate items into Arabic and back into English/German (two independent procedures).

(2) Assess the differences between the original version and the retranslated version of the items.

(3) Compile the final version of the test.

In order to determine the quality of the instruments, 44 Egyptian students in a pilot study (see results in Chapter 5.1) completed the questionnaires and tests during the school year 2010-2011. After that, those instruments and lesson plans were then used in the intervention study (see results in Chapter 5.2) with a sample that included 147 students in a general middle school in the city of Aswan, Egypt during the school year 2011-2012. The analysis of instruments' data and program's impact were done according to the Rasch model and classical test theory.

In the following chapter, the results of the pilot study will first be presented. Before the main results of this study are presented, the quality of the instruments in the main study will again be tested and discussed. After ensuring that the data resulting from these instruments are reliable, results from the intervention study will be presented, including the impact of the teaching program on students' problem solving abilities, experimental strategy knowledge, perceptions on the quality of science lessons, achievement, and motivation towards science for both the intervention group (IG) and control group (CG).

5. Results

In this chapter, all results of both the pilot and intervention studies will be presented, including an explanation of the quality criteria of instruments in each study. Starting in section 5.2.2., the evaluation of the teaching program's impact on students' problem solving abilities, experimental strategy knowledge, achievement, perceptions on the quality of science lessons and motivation towards science will then be analyzed in order to test the hypotheses H1.1 to H2 and H3.2.

5.1 Pilot study results

From October to November 2010, a pilot study was conducted on a sample of seventh grade students with a mean age of 12.2 years (SD .41). The sample included 44 students (girls) from one class in a general middle school in the city of Aswan, Egypt during the school year 2010-2011. As shown in Table 5 in this study four instruments were conducted: an experimental strategy knowledge test, a problem solving test, an achievement test and a motivation on science questionnaire. The two instruments for assessing the teacher's quality of instruction and students' perceptions on the quality of science lessons were tested later in the intervention study. Therefore, the quality of these instruments will be reported based on the data of the intervention study.

Table 5. Implementation of instruments

Instrument	Implementation	
	Pilot study	Main study
Science achievement test (SAT)	√	√
Problem solving test (PST)	√	√
Experimental strategy knowledge test (ESKT)	√	√
Motivation questionnaire (MQ)	√	√
Perception questionnaire (PQ)	__	√
Teacher quality of instruction questionnaire (IQ)	__	√

The data were processed by means of the Rasch model to evaluate a wide range of reliability and validity issues. Infit statistics (see Chapter 4.4.1) was used to assess whether a given performance or item is consistent with other performances or items (Dawson-Tunik, 2006). Firstly the quality criteria will be reported for every instrument and in closing it will be discussed how the final version of instruments have been compiled. The sample size was different for every instrument because of varying numbers of absent students at the time of administration.

To prove that only one underlying factor explains the variance in the results, a unidimensional Rasch analysis was conducted on all 30 items for the science achievement test (SAT). The analysis suggested eight misfitting items out of 30 items. Following the removal of these eight items, an additional Rasch analyses were conducted. The results suggested that the remaining 22 items are reliable in Rasch terms with satisfactory item fit parameters (.5<MNSQ<1.5; person reliability of .79) (Linacre, 2010). For the problem solving test (PST), all 62 items exhibited satisfactory item fit parameters with acceptable internal consistency (.7<MNSQ<1.4; infit; person reliability of .72). This suggested the problem solving test items (PST) were good indicators of a unified construct. For the experimental strategy knowledge test (ESKT), the analysis suggested one potentially misfitting item. This item was excluded from the analysis. An additional Rasch analyses were employed for the remaining items (34). The result confirmed the fit of the 34 items (.5<MNSQ<1.4; infit; persons reliability of .66) to the Rasch model. However, for the motivation questionnaire items (MQ), fit statistics provided evidence that 5 out of 60 items misfitted the Rasch model and therefore were excluded. Additional Rasch analyses were conducted for the 55 items. The result confirmed the fit of the items (.5<MNSQ<1.5; infit; person reliability of .68) to a Rasch model. This low reliability result for ESKT and MQ was acceptable due to the small sample size.

To demonstrate the discriminant validity of the instruments, a bivariate correlation as shown in Table 6 was estimated, revealing a moderate correlation between the problem solving (PST) and the science achievement test (SAT). This means that, generally, there is a relationship between problem solving abilities and scientific knowledge in the sample (Urdan, 2001). This can be expected because the underlying cognitive activities in Bloom's Taxonomy, like analyzing and evaluating, are also used for solving problems. The results show no significant correlations among the other instruments. This supports the discriminant validity of the instruments.

Table 6. Correlations between the four instruments (Pilot study)

Instrument		SAT	PST	ESKT	MQ
SAT	Pearson Correlation	1	.429**	-.107	-.007
	Sig. (2-tailed)		.006	.518	.969
	N		39	39	31
PST	Pearson Correlation		1	.058	-.038
	Sig. (2-tailed)			.710	.829
	N			44	34
ESKT	Pearson Correlation			1	.071
	Sig. (2-tailed)				.691
	N				34
MQ	Pearson Correlation				1
	Sig. (2-tailed)				
	N				

** =correlation is significant at the .01 level (2-tailed).

5.2 Intervention study results

From September to October 2011, an intervention study was conducted on a sample of seventh grade classes. Before the main results of this study are presented, the quality of the instruments will be tested for the main study sample. After ensuring that the data resulting from these instruments are reliable, the results of the intervention study will be investigated and described. These results include the impact of the teaching program on students' dependent variables. A discussion about these results will be provided in Chapter 6. In this study, the sample size was different for every instrument because of varying numbers of absent students at the time of administration.

5.2.1 Quality criteria of the instruments in the main study

Since the sample size in the main study was also small, results had to be carefully interpreted and restrictions towards the model fit indices had to be applied (Hessen, 2010).

The Science Achievement Test (SAT)

The aim of the test is to assess students' content knowledge on the topic *density and buoyancy*. This instrument is a multiple-choice test with 22 items. Items were developed in order to cover different cognitive activities and to estimate the difficulty of the items a posteriori based on the six categories of Bloom's Revised Taxonomy (see Chapter 4.1.1). The test was conducted in a pre/post design. For the control group, n=73 students participated while n=74 students participated in the intervention group. In order to investigate if the 22 items defined a single construct, a unidimensional Rasch analysis was conducted. The analysis suggested one potentially misfitting item. This item was then removed and additional Rasch analyses were conducted. The remaining 21 items fit the Rasch rating scale model at both the pre (.6<MNSQ<1.39; Infit; person reliability of .61) and post .76<MNSQ<1.39; Infit; person reliability of .80) time points. These analyses accepted using 21 achievement test items to define a single overall achievement variable.

The Problem Solving Test (PST)

The goal of this instrument is to measure students' abilities concerning problem solving (see Chapter 4.1.2) in a pre/post design. Tasks in problem solving were also all based on the topic of density and buoyancy. The test contains four items with 62 pair-wise comparisons. The control group included n=72 students while the intervention had n=71 students. All items exhibited satisfactory item fit parameters at both the pre (.65<MNSQ<1.32; Infit; person reliability of .69) and post (.69<MNSQ<1.29;

Infit; person reliability of.90) time points. As a consequence, 62 items of the problem solving test can be accepted for the intervention study.

The Experimental Strategy Knowledge Test (ESKT)

The aim of this test was to assess students' experimental strategy knowledge in a pre/post design (see Chapter 4.1.2). Again, tasks in experimental strategy knowledge were all based on the topic of density and buoyancy. A total of n=72 students participated in the control group and n=71 students participated in the intervention group. Item fit parameters were investigated and person measures computed. Fit statistics provided evidence that all 34 items fitted the Rasch model with infit values of .65<MNSQ<1.38 at the pre time points and an infit of .69<MNSQ<1.22 at pre time points. Analysis suggested a person reliability of .61 at the pre time point and a person reliability of.74 at the post time point.

The Motivation Questionnaire (MQ)

The fourth instrument used in this study is a questionnaire assessing the students' motivation towards science instruction again in a pre/post design. This instrument consists of 55 items and was adapted from the project of quality of instruction in physics (QUIP) and PISA 2006 questionnaires to the conditions in Egypt (see Chapter 4.1.3). In the control group n=71 students participated whereas n=70 students participated in the intervention group. The results confirmed that all 55 items fitted the Rasch model at the pre time points with infit values of Infit; .65<MNSQ<1.42; person reliability of .61, and an Infit of .88<MNSQ<1.28 with a person reliability of .79 at post time points.

The teacher's Quality of Instruction Questionnaire (QI)

The aim of this questionnaire was to ensure the comparability of teaching in both control group (CG) and intervention group (IG) using a standardized instrument (Clausen, 2002). This instrument consists of 112 items (see Chapter 4.1.4). During the teaching period for each lesson, the intervention was observed by a second teacher who had to evaluate instructional quality by rating different statements (items). One teacher participated as observer for both the control group (CG) and the intervention group (IG). The questionnaire was analyzed using classical test theory, due to poor Rasch-fit criteria for the partial credit model. Table 7 illustrates this instrument's quality criteria results. The results are separately reported for pre- and post-assessment of the control and the intervention group.

Table 7. Quality criteria of the teacher's quality of instruction questionnaire

No. of items	M (CG)	SD (CG)	Cronbach's Alpha (CG)	M (IG)	SD (IG)	Cronbach's Alpha (IG)
112	129.40	23.03	.845	132.40	29.77	.909

<u>Students' perceptions on the quality of science lessons questionnaire (PQ)</u>

The sixth instrument used in this study is a questionnaire assessing students' perceptions on the quality of science lessons. It was adapted from TIMSS 2007 (Olson et al., 2008) and Reyer (2004) questionnaires on the topic of density and buoyancy. This instrument consists of 24 items (see Chapter 4.1.4) and was implemented in a pre/post design arrangement. n=72 students participated for the control group and n=74 for the intervention group. The analyses demonstrated satisfactory item fit parameters of all 24 items at both the pre and post assessments with infit value of .78< MNSQ< 1.28; person reliability of .68 and .84< MNSQ< 1.29; person reliability of .70, respectively.

Discriminant validity of the instruments was verified by correlating the pre/post-assessments results of science achievement, problem solving, experimental strategy knowledge, motivation and perception instruments for the control and intervention group. Because the science achievement test (SAT) showed non-normally distributed data within each group in the pre and post assessments (explained in Chapter 5.2.2), both the Pearson r and Spearman r_s correlation coefficient were calculated. Tables 8, 9, 10, and 11 illustrate the related results. There is a small, two-tailed correlation at p≤.05 between the motivation and perception questionnaires in the post-assessment (Pearson r, Table 10). This shows that both the motivation and perception questionnaires are not completely independent. The questionnaires do, however, measure different constructs (Urdan, 2001). This result can be expected because some aspects related to perceptions on the quality of science lessons are also mentioned in the motivation questionnaire. In addition, the result showed no significant correlations towards the other instruments. This supports the discriminant validity of the instruments.

Table 8. Pearson correlations between the instruments in the pre-assessment

Instrument		SAT	PST	ESKT	MQ	PQ
SAT	Pearson Correlation	1	.126	-.011	-.088	.122
	Sig. (2-tailed)		.134	.893	.297	.141
	N		143	143	141	146
PST	Pearson Correlation		1	.029	.117	-.040
	Sig. (2-tailed)			.732	.167	.633
	N			143	140	142
ESKT	Pearson Correlation			1	.054	-.109
	Sig. (2-tailed)				.523	.196
	N				140	142
MQ	Pearson Correlation				1	-.124
	Sig. (2-tailed)					.146
	N					140
PQ*	Pearson Correlation					1
	Sig. (2-tailed)					
	N					

* = PQ perception questionnaire

Table 9. Spearman correlations between the instruments in the pre-assessment

Instrument		SAT	PST	ESKT	MQ	PQ
SAT	Correlation Coefficient	1.000	.127	-.031	-.078	.120
	Sig. (2-tailed)		.132	.716	.356	.148
	N		143	143	141	146
PST	Correlation Coefficient		1.000	.036	.108	-.072
	Sig. (2-tailed)			.670	.203	.396
	N			143	140	142
ESKT	Correlation Coefficient			1.000	.086	-.146
	Sig. (2-tailed)				.311	.084
	N				140	142
MQ	Correlation Coefficient				1.000	-.100
	Sig. (2-tailed)					.238
	N					140
PQ	Correlation Coefficient					1.000
	Sig. (2-tailed)					
	N					

Table 10. Pearson correlations between the instruments in the post-assessment

Instrument		SAT	PST	ESKT	MQ	PQ
SAT	Pearson Correlation	1	-.110	.024	-.042	-.022
	Sig. (2-tailed)		.186	.772	.624	.789
	N		147	147	141	147
PST	Pearson Correlation		1	-.043	.101	.105
	Sig. (2-tailed)			.608	.232	.205
	N			147	141	147
ESKT	Pearson Correlation			1	.043	.109
	Sig. (2-tailed)				.616	.188
	N				141	147
MQ	Pearson Correlation				1	.166*
	Sig. (2-tailed)					.049
	N					141
PQ	Pearson Correlation					1
	Sig. (2-tailed)					
	N					

*=correlation is significant at the .05 level (2-tailed)

Table 11. Spearman correlations between the instruments in the post-assessment

Instrument		SAT	PST	ESKT	MQ	PQ
SAT	Correlation Coefficient	1.	.015	-.129	-.037	-.076
	Sig. (2-tailed)		.858	.119	.663	.360
	N		147	147	141	147
PST	Correlation Coefficient		1.000	-.046	.053	.045
	Sig. (2-tailed)			.582	.536	.587
	N			147	141	147
ESKT	Correlation Coefficient			1.000	.140	.144
	Sig. (2-tailed)				.098	.083
	N				141	147
MQ	Correlation Coefficient				1.000	.123
	Sig. (2-tailed)					.147
	N					141
PQ	Correlation Coefficient					1.000
	Sig. (2-tailed)					
	N					

5.2.2 Evaluation of the impact of the teaching program

The following sections present the test on normal distribution of data within each group as well as the t-test and Mann–Whitney U-test results for all dependent and control variables in the pre-assessments. An investigation of the impact of the teaching program in the learning gains on students' problem solving abilities, experimental strategy knowledge, achievement, perceptions on the quality of science lessons and motivation towards science will be subsequently described. Figure 15 illustrates the data assessed within the study.

Pre-test	Intervention in Grade 7		Post-test
	Intervention group N= 74 students	Control group N=73 students	
- Age - Gender - Social background - Cognitive abilities - Motivation - Achievement - Problem-solving abilities - Experimental strategy knowledge - Perceptions on the quality of science lessons	Learning program (10 hours)	Regular teaching (10 hours)	- Motivation - Achievement - Problem-solving abilities - Experimental strategy knowledge - Perceptions on the quality of science lessons
	Teacher quality of instruction		

Figure 15. Data assessed within the study

Referring to the hypotheses H1.1 to H2 and H3.2, which assumed that – compared to the control group – students in the intervention group would exhibit larger gains in science achievement, problem-solving abilities, experimental strategy knowledge, perceptions on the quality of science lessons, and motivation towards science. To confirm theses hypotheses, parametric and non-parametric tests (t-test and Mann–Whitney U-test for independent samples) were used. Before conducting any analysis of variance (t-test or U-test for independent samples), as described in the following section, the normal distribution of all data within each group was checked separately.

Normal distribution of data

Because the analysis involves comparing two groups, then what is important is not the overall distribution, but rather the distribution within each group (Field, 2009). Hence, the distribution of data within each group in the pre/post- assessments and the calculated learning gains was checked separately and, since the sample size in this study is small ($3 < n < 2000$), the Shapiro-Wilk test was used (see Royston, 1982). As one can see in Tables 12 and 13 the Shapiro-Wilk test revealed that there is normal distribution of data at $\rho \leq .05$ within each group on the pre-assessment of the variables of age, cognitive abilities, social background, problem solving, experimental strategy knowledge, motivation, perception and teachers' quality of instruction. However, achievement variable showed not normally distributed data for both groups at $\rho \leq .05$ (Field, 2009).

In addition, the results showed that there are normally distributed data at $\rho \leq .05$ for both groups on the calculated gain scores (Posttest-pretest) of the motivation variable (see Table 13). And the results in the same table also demonstrated that there are no normally distributed data based on the leaning gains of the dependent variables of achievement, problem solving, experimental strategy knowledge and perception. Hence, in order to check for the differences between the two groups (CG & IG) based on the pre-assessments and the learning gains, the t-test was used with normally distrusted data, while the Mann–Whitney U-test was used with non-normally distrusted data (Field, 2009).

Table 12. Normal distribution of data on the pre/post assessments

Variable	Assessment	Group	Shapiro-Wilk Test			Normal distribution
			D	df	Sig. ρ*	√
Age	Pre	CG	.969	73	.07	√
		IG	.968	74	.06	√
Cognitive abilities	Pre	CG	.990	73	.82	√
		IG	.987	74	.64	√
Social background	Pre	CG	.986	73	.65	√
		IG	.978	74	.21	√
Achievement	Pre	CG	.933	73	.01	-
		IG	.967	74	.04	-
	Post	CG	.973	73	.12	√
		IG	.954	74	.01	-
Problem solving	Pre	CG	.972	72	.11	√
		IG	.980	71	.31	√
	Post	CG	.972	73	.10	√
		IG	.977	74	.18	√
Experimental strategy knowledge	Pre	CG	.975	72	.17	√
		IG	.981	71	.38	√
	Post	CG	.973	73	.11	√
		IG	.987	74	.67	√
Motivation	Pre	CG	.977	71	.23	√
		IG	.974	70	.14	√
	Post	CG	.974	71	.14	√
		IG	.978	70	.25	√
Perceptions	Pre	CG	.981	72	.35	√
		IG	.970	74	.07	√
	Post	CG	.975	73	.16	√
		IG	.973	74	.12	√
Teachers' quality of instruction	Pre	CG	.832	5	.15	√
		IG	.817	5	.11	√

*at the .05 level

(For ρ ≤.05 the null hypothesis must be rejected)

Table 13. Normal distribution of data based on the learning gains of each dependent variable

Variable	Group	Shapiro-Wilk			Normal distribution
		D	df	Sig. ρ*	
Achievement	CG	.952	73	.007	-
	IG	.963	74	.031	-
Problem solving	CG	.909	72	.000	-
	IG	.925	71	.000	-
Experimental strategy knowledge	CG	.925	72	.000	-
	IG	.960	71	.023	-
Motivation	CG	.983	71	.464	√
	IG	.990	70	.866	√
Perceptions	CG	.911	72	.000	-
	IG	.812	74	.000	-

*at the .05 level (For ρ ≤.05 the null hypothesis must be rejected)

T-test and U-test results in the pre-assessments

Both groups (CG & IG) were quite comparable regarding the gender variable. The control group contained two classes, one with 49.5% males and 50.5% females, while the intervention group consisted of two classes: 50% males, and the other 50% females. In addition, based on the results of the above Shapiro-Wilk test results and referring to Tables 14 and 15, the t-test and U-test results confirmed that, between the two groups (CG & IG), no significant differences at ρ ≤.05 existed in the pre-assessments of the aforementioned control or dependent variables. Accordingly, students in the control and intervention group were similar based on the measures of the control variables and the pre-assessments of dependent variables.

Table 14. T-test for dependent and control variables in the pre-assessments

Variable	Group	N	M	SD	Levene's test		t-test			HV**
					F	Sig. ρ	t	df	Sig. ρ*	
Age	CG	73	12.35	.47	1.09	.30	-.92	145	.36	√
	IG	74	12.43	.55						
cognitive abilities	CG	73	34.90	6.67	.04	.84	.73	145	.47	√
	IG	74	34.10	6.8						
social back-ground	CG	73	50.49	5.46	.05	.82	-1.92	145	.06	√
	IG	74	52.20	5.31						
problem solving	CG	72	31.56	18.37	3.13	.08	.29	141	.77	√
	IG	71	30.72	15.76						
experimental strategy knowledge	CG	72	21.96	6.97	.250	.62	.31	141	.76	√
	IG	71	21.60	7.24						
motivation	CG	71	30.82	9.94	1.11	.30	.-.30	139	.77	√
	IG	70	31.34	11.16						
perceptions	CG	72	19.70	7.65	.36	.55	-1.7	144	.09	√
	IG	74	21.86	7.42						

*at the .05 level (2-tailed), ** Homogeneity of variance

Table 15. U-test for achievement variable in the pre assessment

Variable	Group	N	M	SD	M rank	U	Z	Asymp. Sig. ρ*
achievement	CG	73	4.22	2.41	75.68	2578	-.481	.631
	IG	74	4.01	2.05	72.34			

*at the .05 level

Impact of the teaching program on students' achievement

The hypothesis H1.1 was tested, which assumed that the students in the intervention group would achieve significantly higher on content knowledge as the students of the control group. The comparison between control and intervention group based on the learning gains using U-test for independent sample are illustrated in Figure 16.

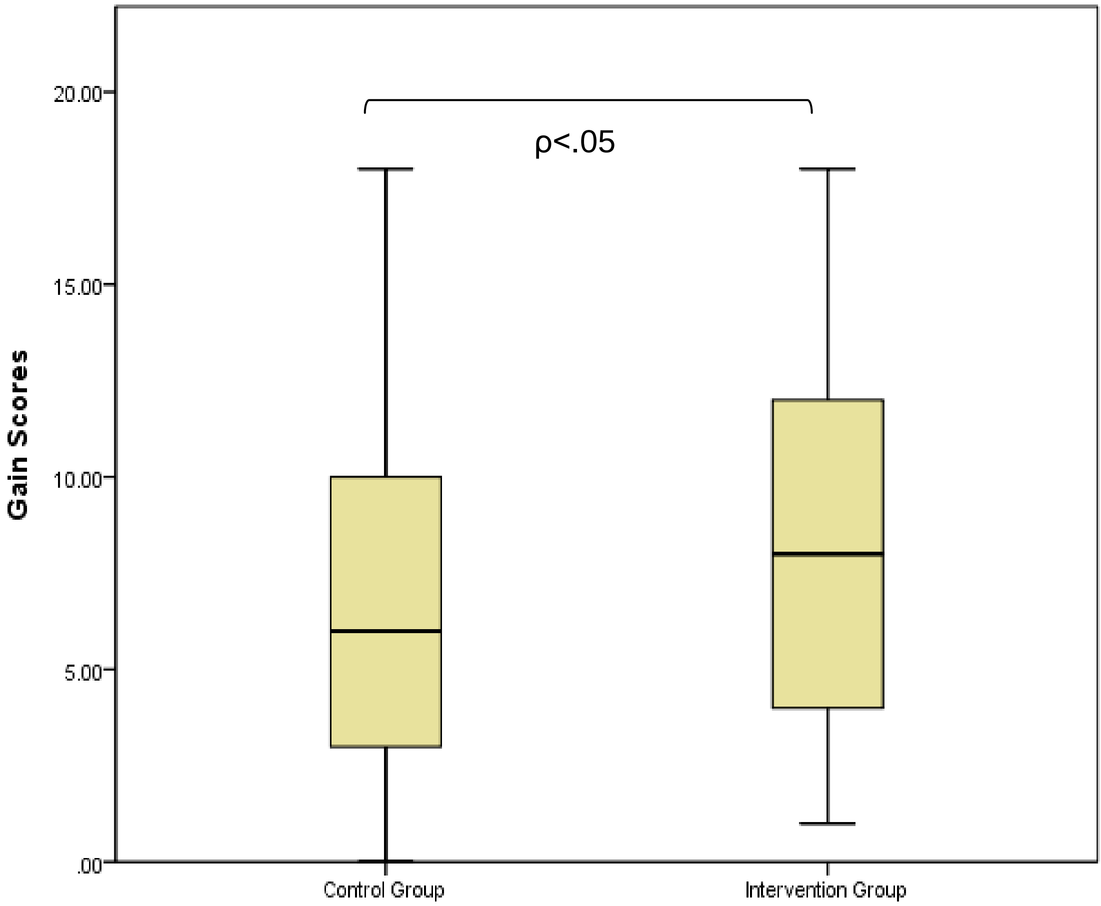


Figure 16. Comparison between control and intervention group on the increased scores of students' achievement

The result (U= 2155, Z=-2.121, $\rho<.05$) revealed that the teaching program had an impact on students' achievement, but with a small effect size of r=.18. Hence, the hypothesis H1.1 can be confirmed.

Impact of the teaching program on students' problem solving abilities

Hypothesis H1.2 was confirmed by finding a significant difference between CG and IG as shown in Figure 17.

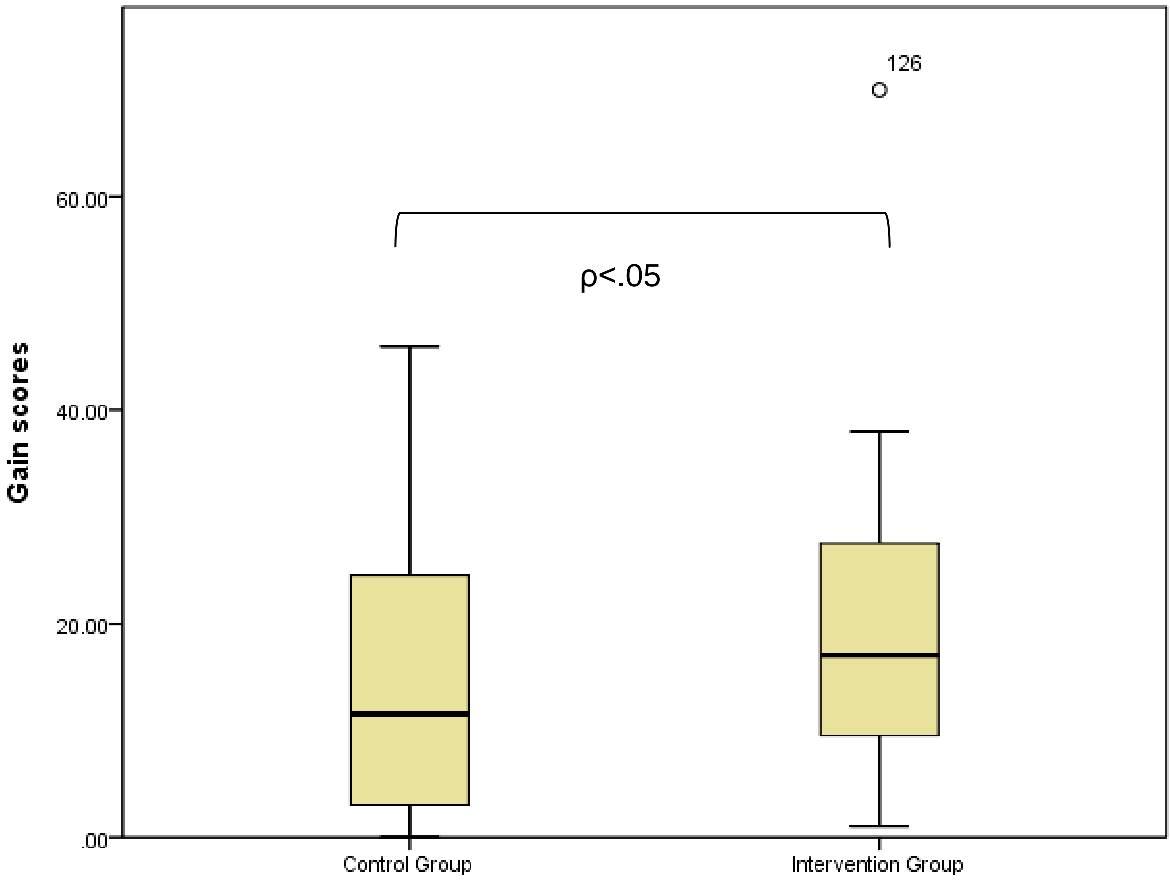


Figure 17. Comparison between control and intervention group on the increased score of students' problem solving abilities

The U-test (U=1940, Z=-2.49, $\rho<.05$) revealed that there is an impact again with a small effect size of r=.21of the teaching program on students' problem solving abilities.

Impact of the teaching program on students' experimental strategy knowledge

Hypothesis H1.3 assumed that the intervention group would reach larger gains in experimental strategy knowledge than the control group.

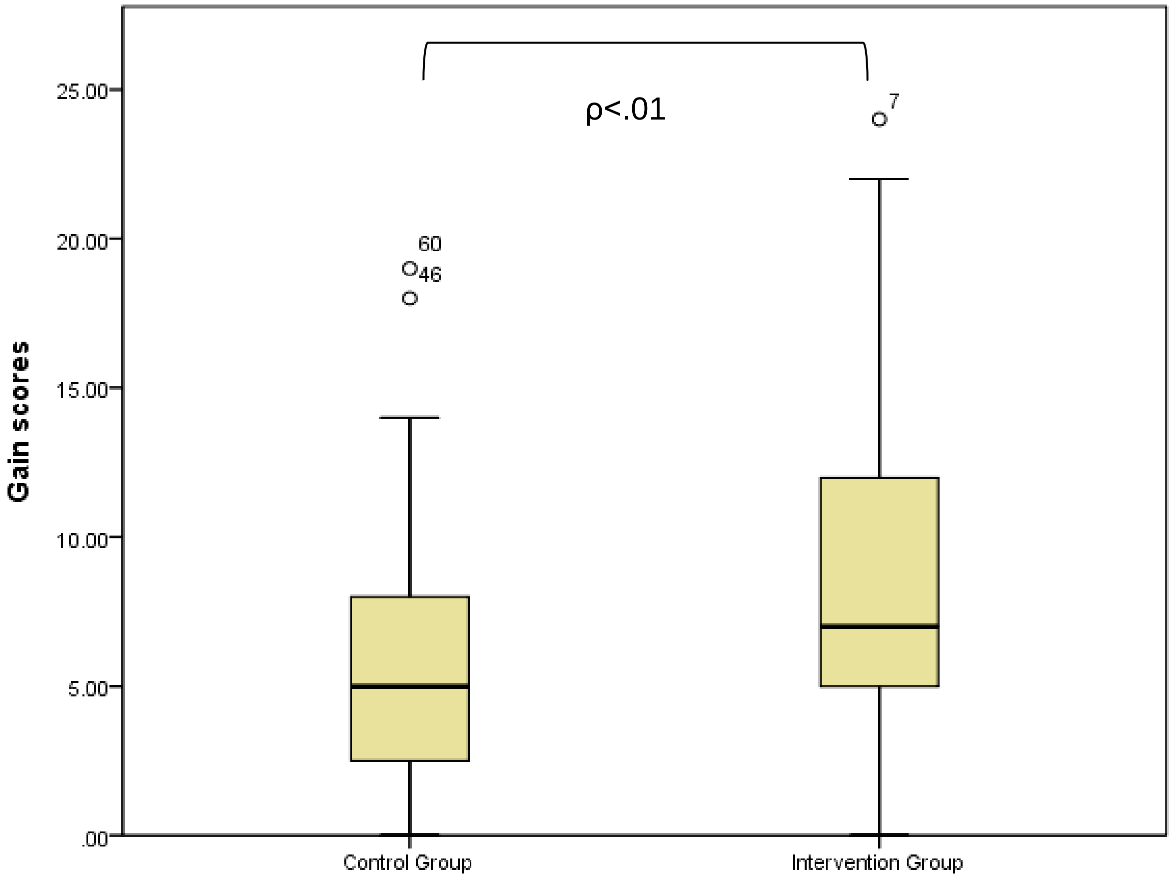


Figure 18. Comparison between control and intervention group on the increased scores of students' experimental strategy knowledge

As shown in figure 18 a significant difference between CG and IG could also be shown. The result of U-test ($U=1873.5$, $Z=-2.76$, $\rho<.01$) proved that there is an impact with a small effect size of $r=.23$ of the teaching program on students' experimental strategy knowledge. Accordingly, hypothesis H1.3 is also proven.

Impact of the teaching program on students' motivation

Hypothesis H2 assumed that the intervention group would show a significantly a larger pre/post difference of motivation towards science than the control group. Figure 19 illustrates the differences between the both groups (CG & IG) based on the increased scores of students' motivation using t-test:

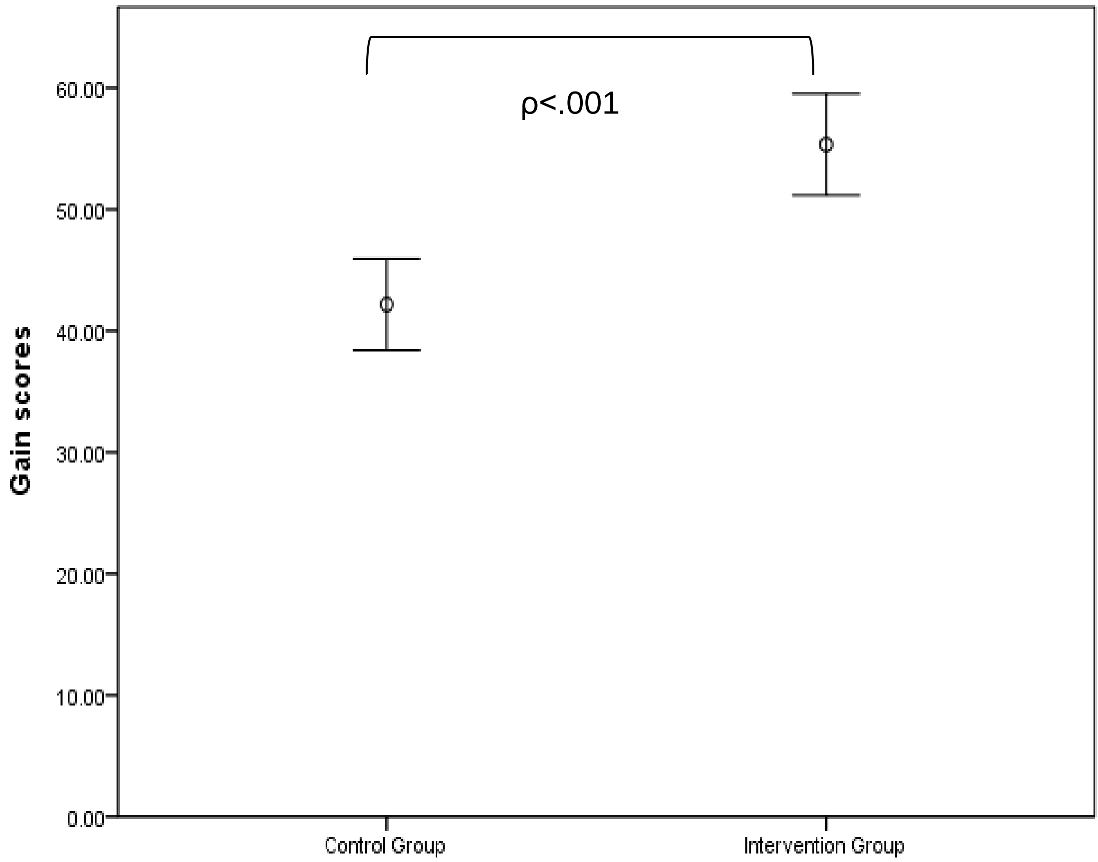


Figure 19. Comparison between control and intervention group on the increased scores of students' motivation

The result for motivation showed a significant difference between CG and IG. ($t(139)=-4.68$, $\rho<.001$) proved that there is an impact with a medium effect size of $d=.79$ of the teaching program on students' motivation towards science. Accordingly, the hypothesis H2 can also be confirmed.

Impact of the teaching program on quality of instruction

To determine the impact of the teaching program on quality of instruction, two main analyses were done. 1. A comparison between control and intervention group based on the teachers' quality of instruction was made. 2. To investigate the impact of the teaching program on students' perceptions on the quality of science lessons, a comparison between control and intervention group in the learning gains was conducted.

Teachers' quality of instruction

To ensure the comparability of teaching in both groups (CG & IG), a standardized instrument (Clausen, 2002) was implemented. In order to test the hypothesis H3.1 which assumed that there are no significant differences between the control and intervention group regarding the teachers' quality of instruction, a t-test for independent samples was used. The analyses confirmed, based on Clausen questionnaire, that there were no significant differences ($t(8)=-.297$, $\rho>.05$) between both groups (CG & IG) regarding the quality of the teacher's instruction. Thus, the hypothesis H3.1 is also proven.

Students' perceptions on the quality of science lessons

A U-test was used to investigate the impact of the teaching program on students' perceptions on the quality of science lessons. The hypothesis H3.2, assumed that, relative to the control group, the intervention group would exhibit more positive perceptions on the quality of science lessons. This hypothesis was also tested. Figure 20 illustrates the result of the comparison:

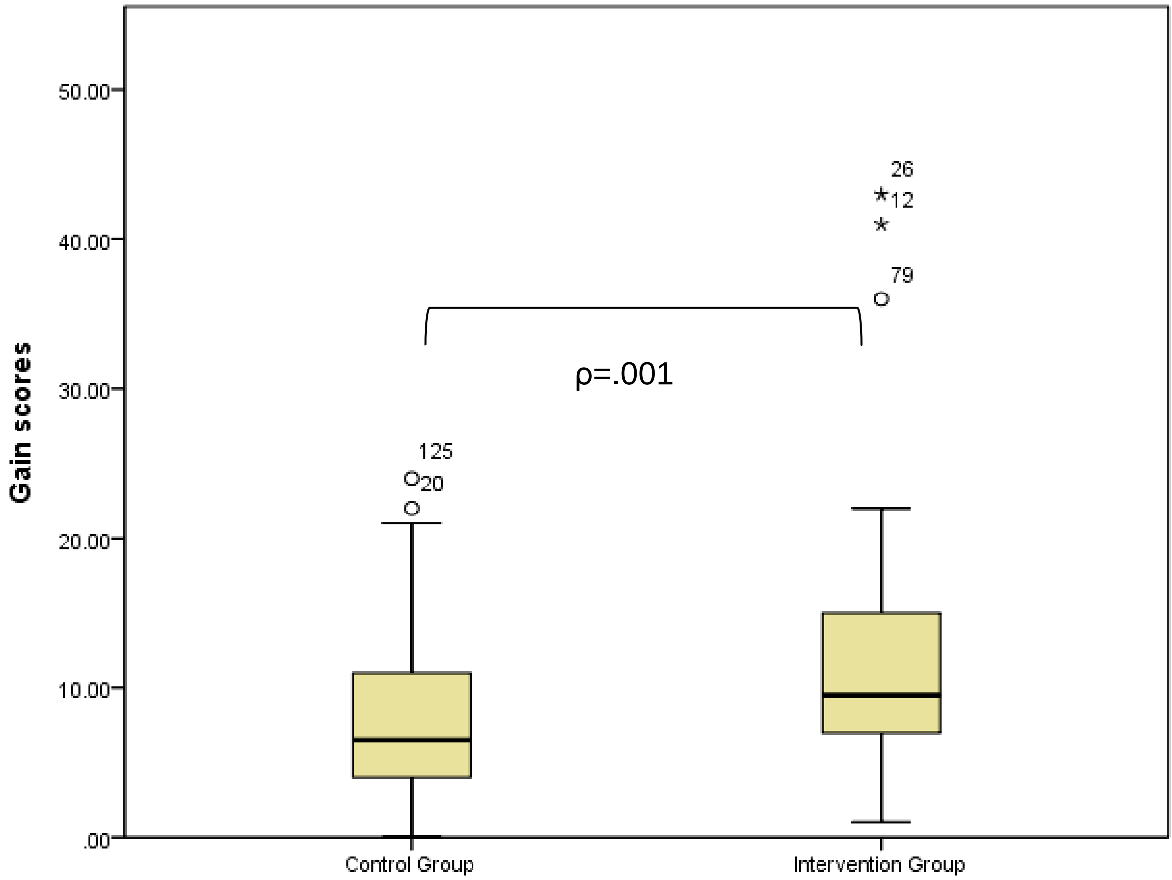


Figure 20. Comparison between control and intervention group on the increased scores of students' perceptions

U-test results (U=1734, Z=-3.65, ρ=.001) proved that there is a significant increase with a medium effect size of r=.30 of the teaching program on students' perceptions on the quality of science lessons. Thus, the hypothesis H3.2 is confirmed.

5.3 Summary of results

The main goal of this study was to develop and evaluate a teaching program based on a model for problem solving in order to foster students' problem solving abilities, experimental strategy knowledge, achievement, perceptions and motivation towards science. This program was developed using Helmke's (2003) model for quality of instruction. Helmke claims that teachers offer learning opportunities that are used by students in individual ways according to their abilities and knowledge. The intervention in the presented study refers to the teachers offer, how students perceive this offer and which influence it has on students' outcomes. Six instruments developed and adapted. The developed teaching program and instruments were then used in an intervention study, which was conducted from September to October 2011 on a sample of seventh grade students. These instruments were used to assess the students' dependent variables before and after a six-week unit on density and buoyancy at a middle school in Egypt. In this study, a two-group, pre/post-test, quasi-experimental design (Shadish et al., 2002) was used to determine whether students who are taught how to use the model of problem solving for an experiment dealing with the topic of density show improvement in the aforementioned dependent variables. Quality of instruments for the intervention study sample and comparability of groups have been shown. In closing, intervention group students were compared to control group students to determine possible differences in the learning gains.

The overarching hypothesis of the presented study assumed that relative to the control group, the intervention group would exhibit larger gains in science achievement, problem solving abilities, experimental strategy knowledge, perceptions on the quality of science teaching, and motivation towards science. In order to prove this hypothesis, both parametric and non-parametric tests (t-test and U-test for independent samples) were used. Results confirmed a significant impact at $\rho < .05$ with a small and medium effect size ($.1<r<.5$ & $.1<d<.8$) of the teaching program on students' outcomes.

In the next chapter, the discussion and perspectives of this study will be illustrated. This chapter includes sections about a discussion of the results, the limitations and the gains of the study, and finally, the recommendations for further research.

6. Discussion and perspectives

The small and medium effect sizes are in line with expectations as the sample size in this study was quite small (see e.g. Carvajal & Skorupski, 2010; Hancock & Freeman, 2001; Scherbaum & Ferreter, 2009). In addition, this could also be attributed to the short intervention period.

This study's findings are in agreement with those of several studies that have found a positive impact of problem solving as instructional strategy on students' achievement, problem solving abilities, strategy knowledge, and motivation towards science (see e.g. Afamasaga-Fuata'i, 2009; Gök & Sılay, 2010; Goldberg et al., 2010; Keil et al., 2009, Mason & Singh, 2010; Mercedes & Teresa, 2005; Sekuk et al., 2008; Valiotis, 2008).

One theoretical advancement made in the presented study is that a problem solving unit and a corresponding test was developed for teaching and assessing students' content knowledge on the specific topic of density and buoyancy. The unit covers the needs of the Egyptian school system and the Items were developed in order to cover different cognitive activities based on the six categories of Bloom's Revised Taxonomy, which include remembering, understanding, applying, analyzing, evaluating, and creating. Item difficulty increases from understanding to creating. Based on the fact that this topic is currently the first topic that is taught in middle school science curricula in Egypt, this test could be used in order to assess the pre-knowledge of students before enrolling in middle school and to rank students' performance according to pre-assessment (base score).

Another benefit of this study is that two instruments have been developed that can assess problem solving abilities and experimental strategy knowledge on the topic of density and buoyancy. The two tests were partly self-developed and partly adapted from newly performed studies on experimental strategy knowledge and problem solving in Germany. The results of this study proved that the teaching program has a significant, positive effect on students' problem solving abilities and experimental strategy knowledge. Therefore, these two instruments could very well be useful for further research in primary and secondary schools with other science topics.

Moreover, the added value of this study is also the adaptation of three instruments: One instrument is a motivation questionnaire and two instruments are related to students' perceptions on the quality of science lessons and teachers' quality of instruction. These instruments were adapted to the conditions in Egypt middle schools. As the author acknowledges, there is no current attempt being made to use or adapt these instruments in the Egypt education environment. Moreover, these instruments could very well be useful for further research in other school grades.

In addition to the development or adaptation of instruments, this study provides a valid and independent multidimensional model for problem solving. This model can be used for further research for developing other lesson plans in various topics at different school stages and for verifying its effect on students' outcomes. In addition, it could be very useful to apply this model in problem-solving process trainings for pre- or in-service Egyptian science teachers.

An additional theoretical gain of this study is that it developed and evaluated a problem solving-focused science teaching program. Based on an offer-use model for quality of instruction, it enables researchers to identify aspects such as teachers' learning offers, students' abilities and motivational prerequisites, and quality of teaching, which impact students' learning outcomes and which should be taken into account when developing lesson plans. The science teaching program may help and guide teachers and teacher educators on how to plan and to construct experimental learning situations of different content and implement it in classrooms, which is now relevant and top-of-mind in Egypt educator circles. The same, therefore, can be said of studies which provide guidance for Egyptian science teachers.

Another contribution of this dissertation is that it shows how instrument items can be successfully developed and adapted in an international study and applied in a different language and culture. In general, it has to be kept in mind that this study focuses on a very limited area of one topic and one grade. Moreover, the sample was selected only from one location – the city of Aswan. Results need to be discussed carefully due to the relatively small and restricted sample size; it illuminates some potentially fruitful possibilities for further research. Because students from the intervention group outperformed students from the control group in all dependent variables, these results can guide future research in the field of science teaching in

Egypt, having now provided evidence for an increase in students' achievement, problem solving abilities, experimental strategy knowledge, perceptions on the quality of science lessons, and motivation towards science. It would be beneficial if the study could replicate and validate on a larger and more representative sample in order to generalize the results to the Egyptian educational system. For implementation reason the description of the learning environment planned for this study integrates both a theoretical (Klahr & Dunbar, 1988) and a practical (Oser & Baeriswyl, 2001) model for problem solving. Further research must be performed on the acceptance and the effect of the unit and similar units of different content after being implemented on a larger scale. It should be additionally mentioned that, due to educational and political circumstances in Egypt during the intervention study, it was not possible to conduct a video study.

Accordingly,

First, conducting a video study in order to evaluate the quality of instruction in the Egyptian classroom should definitely be examined in a future study.

Second, the assumption that enacting reforms on current teaching approaches will influence students' outcomes, addressed at the beginning of this thesis, has to be further examined for both the primary and secondary school.

Third, using a new teaching approach that focuses on scientific inquiry and problem solving and its effects on students' outcomes for the other science topics should be further verified.

Fourth, using problem solving as instructional strategy and its effects on students' situational motivation needs to be further verified. The applied model of problem solving was shown to have a significant effect on students' general motivation towards science. It is possible that this model would also have an impact on students' situational motivation.

Fifth, the impact of problem solving strategy on students' outcomes should be examined in the context of other science disciplines. The applied model of problem solving was used only with a physics topic. Contexts related to other disciplines could shed light on whether problem solving as an instructional strategy is specific to the different disciplines or if it is specific to science in general.

Sixth, the relationship between students' general motivation towards science and perceptions on the quality of science lessons could be investigated in more detail. Such detailed analysis could provide useful aspects for developing science teaching in Egyptian classes.

Finally, the effects of the Egyptian class size for all elementary, middle and secondary schools on students' outcomes need to be evaluated. Such evaluation could lead to other active instructional programs or approaches that could at the very least deal with or quite possibly optimize handling the large size of Egyptian classes.

7. References

Afamasaga-Fuata'i, K. (2009). *Concept mapping in mathematics: Research into practice*. New York: Springer.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.

Anderson, L.W. & Krathwohl, D.R. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy*. New York: Longman.

Anderson, S.E. (2010). Moving change: Evolutionary perspectives on educational change. In A. Hargreaves, A. Lieberman, M. Fullan & D. Hopkins (Eds.), *Second international handbook of educational change* (Vol. 23, pp. 65-84). London: Springer.

Arora, A., Ramirez, M.J. & Howie, S.J. (2006). Using indicators of educational contexts in TIMSS. In S.J. Howie & T. Plomp (Eds.), *Contexts of learning mathematics and science: Lessons learned from TIMSS* (pp. 31-49). London: Routledge.

Ary, D., Cheser, J.L. & Sorensen, C.K. (2010). *Introduction to research in education, (*8th ed.). Belmont, CA: Wadsworth, Cengage Learning.

Babbie, E. (2011). *The basics of social research* (5th ed.). Belmont, CA: Thomson Wadsworth.

Bentley, D. & Watts, M. (2005). *Communicating in school science: Groups, tasks and problem solving 5- 16*. London: The Falmer Press.

Berg, C.A., Bergendahl, V.C., Lundberg, B. & Tibell, L. (2003). Benefiting from an open-ended experiment? A comparison of attitudes to, and outcomes of, an expository versus an open-inquiry version of the same experiment. *International Journal of Science Education, 25*(3), 351-372.

Bernholt, S., Neumann, K. & Nentwig, P. (Eds.). (2012). *Making it tangible: Learning outcomes in science education*. Münster: Waxman.

Blatchford, P., Bassett, P., Goldstein, H., Martin, C., Catchpole, G., Edmonds, S. & Moriarty, V. (2003). *The class size debate: Is small better?.* Philadelphia: Open University Press.

Bloom, J.W. (Ed.). (2006). *Creating a classroom community of young scientists (*2nd ed.). New York: Routledge.

Boone, W.J., & Scantlebury, K. (2006). The role of Rasch analysis when conducting science education research utilizing multiple-choice tests. *Science Education, 90*(2), 253-269.

Borghese, P. & Gronau, R.C. (2005). Convergent and discriminant validity of the universal nonverbal intelligence test with limited English proficient Mexican-American elementary students. *Journal of Psychoeducational Assessment, 23*, 128-139.

BouJaoude, S., Wiles, J.R., Asghar, A. & Alters, B. (2011). Muslim Egyptian and Lebanese students’ conceptions of biological evolution. *Science & Education*, DOI: 10.1007/s11191-011-9345-4.

Brislin, R.W. (1970). Back translation for cross-culture research. *Journal of Cross-Culture Psychology, 1*, 185-216.

Brislin, R.W. (1980). Translation and content analysis of oral and written material. In H.C. Triandis & J.W. Berry (Eds.), *Handbook of cross-culture psychology: Methodology* (pp. 389-444). Boston, MA: Allyn & Bacon, Inc.

Carroll, J.B. (1963). A model of school learning. *Teachers College Record, 64*, 723-733.

Carvajal, J. & Skorupski, W.P. (2010). The effects of small sample size on identifying polytomous DIF using the liu–agresti estimator of the cumulative common odds ratio. *Educational and Psychological Measurement, 70*(6), 914-925.

Carver, R.H. & Nash J.G. (Eds.). (2010). *Doing data analysis with SPSS® version 18*. Boston, MA: Brooks/Cole Cengage Learning.

Chang, H. (2011). How historical experiments can improve scientific knowledge and science education: The cases of boiling water and electrochemistry. *Science & Education, 20*, 317-341.

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.

Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive?* [Quality of instruction: A matter of perspective?]. Münster: Waxmann.

Cortina, J.M. (1993). What is coefficient alpha? Examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104.

Creemers, B.P. (1994). *The effective classroom.* London: Cassell.

Creemers, B.P. & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools.* London: Routledge.

D'Amico, J. & Gallaway, K. (2010). *Differentiated instruction for the middle school science teacher: Activities and strategies for an inclusive classroom*. San Francisco: JohnWiley & Sons, Inc.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1), Retrieved from (5.3.2012): http://epaa.asu.edu/ojs/article/view/392

Dawson-Tunik, T.L. (2006). Stage-like patterns in the development of conceptions of energy. In X. Liu & W.J. Boone (Eds.), *Applications of Rasch measurement in science education* (pp. 111-136). Maple Grove, MN: JAM Press.

De Gruijter, D.N. & Van der Kamp, L.J.T. (2008). *Statistical test theory for the behavioral sciences.* London: Chapman & Hall/CRC.

DeBoer, G.E. (2006). Historical perspectives on inquiry teaching science in schools. In L.B. Flick & N.G. Lederman (Eds.), *Scientific inquiry and nature of science: Implementation for teaching, learning and teacher education (*ix-xviii, pp.17-36). Dordrecht: Springer.

Deci, E.L. (1992). The relation of interest to the motivation of behavior: A self-determination theory perspective. In K.A. Renninger, S. Hidi & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 43-70). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Deci, E.L. & Ryan, R.M. (1985). *Intrinsic motivation and self-determination in motivation.* New York: Plenum.

Dogru, M. (2008). The application of problem solving method on science teacher trainees on the solution of the environmental problems. *Journal of Environmental & Science Education, 3*(1), 9-18.

Duschl, R. & Hamilton, R. (2011). Learning science. In R.E. Mayer & P.A. Alexander. (Eds.), *Handbook of research on learning and instruction* (pp. 78-107). New York: Routledge.

Elliott, A.C. & Woodward, W.A. (2007). *Statistical analysis quick reference guidebook: With SPSS examples*. London: Sage Publications, Inc.

Emden, M. & Sumfleth, E. (2012). Assessing experimental procedures through different formats–a comparative study. In C. Bruguière, A. Tiberghien & P. Clément (Eds.), *E-Book Proceedings of the ESERA 2011 Conference, Lyon France. Science learning and citizenship.* Part 10: Co-ed. R. Millar (pp. 23-29). Lyon, France: European Science Education Research Association.

Engelhard, G. (2011). Historical views of invariance: Evidence from the measurement theories of Thorndike, Thurstone, & Rasch. In N.J. Salkind (Ed.), *Sage directions in educational psychology* (Vol. V, pp. 76-97). Los Angelos: SAGE Publications, Inc.

Faessler, L., Hinterberger, H., Dahinden, M. & Wyss, M. (2006). Evaluating student motivation in constructivistic, problem-based introductory computer science courses. *World Conference on E-Learning*, Hawaii, 2006, 1178. Awarded Paper.

Feasey, R. (2007). *Primary science for teaching assistants*. New York: Routledge.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage Publications.

Fischer, H.E. & Horstendahl, M. (1997). Motivation and learning physics. *Research and Science Education, 27*(3), 411-424.

Fischer, H.E., Klemm, K., Leutner, D., Sumfleth, E, Tiemann, R. & Wirth, J. (2005). Framework for empirical research on science teaching and learning. *Journal of Science Teacher Education, 16*(4), 309-349.

Flick, L.B. & Lederman, N.G. (Eds.). (2006). *Scientific inquiry and nature of science: Implementation for teaching, learning and teacher education.* Dordrecht: Springer.

Frey, N. & Fisher, D. (2011). *The formative assessment action plan: Practical steps to more successful teaching and learning*. Alexandria: ASCD.

Gilbert, I (2002). *Essential motivation in the classroom*. London: Routledge.

Gök, T. & Sılay, I. (2010). The effects of problem solving strategies on students' achievement, attitude and motivation. *Latin-American Journal of Physics Education, 4*(1), 7-21.

Goldberg, F., Otero, V. & Robinson, S. (2010). Design principles for effective physics instruction: A case from physics and everyday thinking. *American Journal of Physics, 78*(12), 1265-1277.

Good, T.L. & Brophy, J.E. (Eds.). (2008). *Looking in classrooms* (10th ed.). Boston, MA: Person Education, Inc.

Goodwin, C.J. (2010). *Research in psychology: Methods and design* (6th ed.). San Francisco, CA: John Wiley & Sons, Inc.

Hambleton, R.K., Swaminathan, H. & Rogers, H.J. (1991). *Fundamentals of item response theory.* Newbury Park, CA: Sage.

Hancock, G.R. & Freeman, M.G (2001). Power and sample size for the root mean square error of approximation test of not close fit in structural equation modeling. *Educational and Psychological Measurement, 61*, 741-758. DOI: 10.1177/00131640121971491.

Hashimoto, K., Pillay, H.K. & Hudson, P.B. (2008) *Evaluating educational reform projects in developing countries: A case study of teacher educational reform in Egypt*. In: 15th International Conference on Learning. Chicago: University of Illinois (3-6 June).

Hassan, F. (1997). Science education in Egypt and other Arab countries in Africa and west Asia, *the interdisciplinary of study abroad, 3*, Retrieved from (22.10.2011): http://www.frontiersjournal.com/issues/vol3/vol3-11_Hassan.htm

Helmke, A. (2003). *Unterrichtsqualität erfassen, bewerten, verbessern* [Measuring, rating and improving the quality of instruction]. Seelze: Kallmeyer.

Hessen, D.J. (2010). Likelihood ratio tests for special Rasch models. *Journal of Educational and Behavioral Statistics, 35*(6), 611-628.

Holstermann, N., Grube, D. & Bögeholz, S. (2010). Hands-on activities and their influence on students' interest. *Research in Science Education*, *40*, 743-757.

Isaksen, S.G., & Treffinger, D.J. (1985). *Creative problem solving: The basic course*. Buffalo, NY: Bearly Limited.

Jackson, S.L. (Ed.). (2009). *Research methods and statistics: A critical thinking approach* (3th ed.). Belmont, CA: Wadsworth, Cengage Learning.

Jeong, H. & Songer, N. (2008). *Understanding scientific evidence and the Data collection process: Explorations of why, who, when, what, and how*. In C.L. Petroselli (Ed.), *Science education issues and developments* (pp. 169-200). New York: Nova Science Publishers, Inc.

Jimerson, S.R., Burns, M.K. & VanDerHeyden, A.M. (2007). Response to intervention at school: The science and practice of assessment and intervention. In S.R. Jimerson, M.K. Burns & A.M VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 3-9). New York: Springer.

Jonassen, D.H. (2011). *Learning to solve problems: A handbook for designing problem-solving learning environments*. New York: Routledge.

Kaplan, R.M. & Saccuzzo, D.P. (2009). *Psychological testing: Principles, applications, and issues* (7^{th} ed.). Belmont, CA: Wadsworth, Cengage Learning.

Keil, C., Haney, J. & Zoffel, J. (2009). Improvements in student achievement and science process skills using environmental health science problem-based learning curricula. *Electronic Journal of Science Education, 31*(1), 1-18.

Keziah A.A. (2010). A comparative study of problem-based and lecture based learning in secondary school students' motivation to learn science. *International Journal of Science and Technology Education Research, 1*(6),126-131.

Kim, M.C. & Hannafin, M.J. (2011). Scaffolding 6^{th} graders' problem solving in technology-enhanced science classrooms: A qualitative case study. *Instructional Science, 39*, 255–282.

Klahr, D. & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*(1), 1-48.

Klos, S., Henke, C., Kieren, C., Walpuski, M. & Sumfleth, E. (2008). Naturwissenschaftliches Experimentieren und chemisches Fachwissen-zwei verschiedene Kompetenzen [Natural science experimentation and content knowledge—two different components]. *Zeitschrift für Pädagogik, 54*(3), 304-321.

Kneeland, S. (1999). *Effective problem-solving: How to understand the process and practise it successfully*. Oxford: How To Books, Ltd.

Kuhn, D. (2005) *Education for thinking*. Cambridge, MA: Harvard University Press.

Lederman, N.G. & Lederman, J.S. (2012). Nature of scientific knowledge and scientific inquiry: Building instructional and capacity through professional development. In B.J. Fraser, K.G. Tobin & C. McRobbie (Eds.), *Second international handbook of science education* (Vol. 24, pp. 335-359). New York: Springer.

Linacre, J.M. (1994). Sample size and item calibrations stability. *Rasch Measurement Transactions, 7*(4), 328.

Linacre, J.M. (2010). *A user's guide to winstepsministep: Rasch-model computer programs*. Chicago: Winsteps.com.

Lipowsky, F., Rakoczy, K., Vetter, B., Klieme, E., Reusser, K. & Pauli, C. (2005). *Quality of geometry instruction and its impact on the achievement of students with different learning characteristics.* Paper presented at the Annual Meeting of the American Educational Research Association (AERA), Montreal, Canada.

Mansour, N. (2010). Impact of the knowledge and beliefs of Egyptian science teachers in integrating a STS based curriculum: A sociocultural perspective. *Journal of Science Teacher Education, 21*, 513-534.

Marschner, J. (2011). *Adaptives Feedback zur Unterstützung des selbstregulierten Lernens durch Experimentieren* [Adaptive feedback to support self-regulated learning through experimentation]. Doctoral dissertation, Universität Duisburg-Essen.

Martin, M.O., Mullis, I.V.S., Gonzalez, E.J. & Chrostowski, S.J. (2004). *TIMSS 2003 international science report. Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Martin, M.O., Mullis, I.V.S. & Foy, P. (2008). *TIMSS 2007 international science report. Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mason, A. & Singh, C. (2010). Helping students learn effective problem solving strategies by reflecting with peers. *American Journal of Physics, 78*(7), 748-754.

Meyer, J.P. (2010). *Reliability*. Oxford: Oxford University Press, Inc.

Ministry of Education-Egypt (MOE) (2008). *The development of education in Egypt* (2004-2008), a notional report, Cairo, Retrieved (12.9.2012) from: http://www.ibe.unesco.org/National_Reports/ICE_2008/egypt_NR08.pdf

Ministry of Education-Egypt (MOE) (2010). *The future vision of pre-university education,* (in Arabic), Retrieved (25.09.2012) from: http://portal.moe.gov.eg/AboutMinistry/Policies/Pages/politic_vision.aspx

National Research Council (1996). *National science education standards*. Washington, DC: The National Academies Press.

Neumann, K., Kauertz, A. & Fischer, H.E. (2012). Quality of instruction in science education. In B. Fraser, K. Tobin & C. McRobbie (Eds.), *Second international handbook of science education* (pp. 247-258). New York: Springer.

Newburgh, R. (2008). Why do We Lose Physics Students? In C.L. Petroselli (Ed.), *Science education issues and developments* (2nd ed.). New York: Nova Science Publishers, Inc.

Newton, L.D. & Newton, D.P. (2008). A problem-based approach to training elementary teachers to plan science lessons. In G.F. Ollington (Ed.), *Teachers & teaching: Strategies, innovations and problem solving* (pp. 55-74). New York: Nova Science Publishers, Inc.

OECD (2009). *Top of the class: High performers in science in PISA 2006*. Paris: OECD Publishing.

Ohle, A. (2010). *Primary school teachers' content knowledge in physics and its impact on teaching and students' achievement*, Doctoral dissertation, Universität Duisburg-Essen.

Olson, J.F., Martin, M.O. & Mullis, I.V.S. (Eds.). (2008). *TIMSS 2007 technical report.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Onetcenter (2008). O*nline.onetcenter.org/skills.* Retrieved (13.5.2011) from: http://www.onetcenter.org/content.html/1.A?d=1

Organization for Economic Co-operation and Development (OECD) & IBRD/the World Bank (2010). *Reviews of national policies for education: Higher education in Egypt* (2nd. ed.). Paris: OECD Publishing.

Organization for Economic Co-operation and Development OECD (2003b). *The PISA 2003 assessment framework: Mathematics, Reading, science and problem solving knowledge and skills*, Retrieved (13.5.2011) from: http://www.oecd.org/document/29/0,3343,en_32252351_32236173_33694301_1_1_1_1,00.html

Oser, F.K. & Baeriswyl, F.J. (2001). Choreographies of teaching: Bridging instruction to learning. In V. Richardson (Ed.), *AERA´s handbook of research on teaching* (4th ed., pp. 1031-1065). Washington, DC: American Educational Research Association.

Page, M.C., Braver, S.L. & MacKinnon, D.P. (2003). *Levine's guide to SPSS for analysis of variance* (2nd ed.). Mahwah, N.J: Lawrence Erlbaum Associates, Inc.

Peers, I.S. (2006). *Statistical analysis for education and psychology researchers*. London: The Falmer Press.

Posamentier, A.S. & Krulik, S. (2009). *Problem solving in mathematics, Grades 3–6: Powerful strategies to deepen understanding.* California, CA: Corwin.

Quintana, C., Reiser, B.J., Davis, E.A., Krajcik, J., Fretz, E., Duncan, R., Kyza, E., Edelson, D. & Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *The Journal of the Learning Sciences, 13*(3), 337-386.

Reyer, T. (2004). *Oberflächenmerkmale und Tiefenstrukturen im Unterricht*. [Characteristics of surface structure and deep structure in lessons]. Berlin: Logos.

Rice, J.K. (1999). The impact of class size on instructional strategies and the use of time in high school mathematics and science courses. *Educational Evaluation and Policy Analysis, 21*(2), 215-229.

Robertson, S.I. (2001). *Problem solving.* Philadelphia, PA: Psychology Press Ltd.

Roderick, M. & Engel, M. (2001). The grasshopper and the ant: Motivational responses of low-achieving students to high-stakes testing. *Educational Evaluation and Policy Analysis, 23*, 197-227.

Ronald, H.H. (2009). Teacher effectiveness and student achievement: Investigating a multilevel cross-classified model. *Journal of Educational Administration, 47*(2), 227-249.

Royston, J.P. (1982). An extension of Shapiro and Wilk's W test for normality to large samples. *Applied Statistics, 31*, 115-124.

Ryan, R.M. & Deci, E.L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*, 54-67.

Schauble, L., Glaser, R., Ranghavan, K. & Reiner, M. (1991). Causal models and experimentation strategies in scientific reasoning. *The Journal of the Learning Science, 1*(2), 201-238.

Schauble, L., Glaser, R., Ranghavan, K. & Reiner, M. (1992). The integration of knowledge and experimentation strategies in understanding a physical system. *Applied Cognitive Psychology, 6*, 321-343.

Scherbaum, C.A. & Ferreter, J.M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods, 12*(2), 347-367.

Schiefele, U., Krapp, A. & Winteler A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K.A. Renniger, S. Hidi & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 183-196). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Seidel, R.J., Kett, A.L. & Perencevich, K.C. (Eds.). (2007). *From principles of learning to strategies for instruction with workbook Companion: A needs-based focus on high school adolescents*. London: Springer.

Sekuk, G.S., Caliskan, S. & Erol, M. (2008). The effects of problem solving instruction on physics achievement, problem solving performance and strategy use. *Latin-American Journal of Physics Education, 2*(3), 151-166.

Shadish, W.R., Memphis, T., Cook, T.D. & Evanston, I. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.

Shaw, R. (2008). *Philosophy in the classroom: Improving your pupils' thinking skills and motivating them to learn*. London: Routledge.

Slavin, R. E. (2011). Instruction based on cooperative learning. In R.E. Mayer & P.A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 78-107). New York: Routledge.

Smith, R.M. (1998). *Ordered response category analysis: A best practice model*. Paper presented at the Winter 1998 Midwestern Objective Measurement Seminar, Chicago.

Spatz, C. (2008). *Basic statistics: Tales of distributions* (9th ed.). Belmont, CA: Thomson Wadsworth.

Stigler, J.W., Gonzales, P., Kawanaka, T., Knoll, S. & Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan and the United States.* Washington, DC: National Center for Education Statistics.

Sutman, F.X., Schmuckler, J.S. & Woodfieldand, J.D. (2008). *The science quest using inquiry/discovery to enhance student learning, grades 7*–12. San Francisco, CA: John Wiley & Sons, Inc.

Tang, X, Coffey, J.E., Elby, A. & Levin, D.M. (2009). The scientific method and scientific inquiry: Tensions in teaching and learning. *Science Education, 94*, 29-47.

Thillmann, H. (2008). *Selbstreguliertes Lernen durch Experimentieren: Von der Erfassung zur Förderung* [Self-regulated learning by experimentation: from controlling to fostering]. Doctoral dissertation, Universität Duisburg-Essen.

Thillmann, H., Künsting, J., Wirth, J. & Leutner, D. (2009). Is it merely a question of 'what' to prompt or also 'when' to prompt?-The role of presentation time of prompts in self-regulated learning. *Zeitschrift für Pädagogische Psychologie, 23,* 105-115.

Toth, E.E., Klahr, D. & Chen, Z. (2000). Bridging research and practice: A cognitively based classroom intervention for teaching experimentation skills to elementary school children. *Cognition and Instruction, 18*, 423-459.

UNESCO (2008). *Egypt's educational context and the government's education priorities and strategies*. UNESCO Cairo Office– Egypt.

Urdan, T.C. (2001). *Statistics in plain English*. Mahwah, NJ: Lawrence Erlbaum Association, Inc.

Valiotis, C. (2008). Improving conceptual understanding and problem solving skills in introductory physics courses using the Socratic dialogue method, *Proceedings of the 2008 American Society for Engineering Education Pacific Southwest Annual Conference.* Washington, DC: American Society for Engineering Education.

Van de Walle, J.A., Karp, K.S. & Bay-Williams, J.M. (2010). *Elementary and middle school mathematics: Teaching developmentally* (7th ed.). Upper Saddle River, NJ: Pearson Publications.

VanGundy, A.B. (2005). *101 Activities for teaching creativity and problem solving*. San Francisco, CA: John Wiley & Sons, Inc.

Vargas, J.S. (2009). *Behavior analysis for effective teaching*. New York: Routledge.

Walberg, H.J. (1981). A psychological theory of educational productivity. In F.H. Farley & N. Gordon (Eds.), *Psychology and education: The state of the union* (pp. 81-108). Berkeley, CA: McCutchan.

Walberg, H.J., Haertel, G.D., Pascarella, E., Junker, L.K. & Boulanger, F.D. (1981). Probing a model of educational productivity in science with national assessment samples of early adolescents. *American Educational Research Journal*, *18*, 233-249.

Weinstein, C.E. & Mayer, R.E. (1986). The teaching of learning strategies. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 315-327). New York: Macmillan.

Wendler, C.L.W. & Walker, M.E. (2006). Practical issues in designing and maintaining multiple test forms for large-scale programs. In S.M. Downing & T.M. Haladyna (Eds.), *Handbook of test development* (pp. 445-467). Mahwah, NJ: Lawrence Erlbaum Association, Inc.

White, B.Y., Shimoda, T.A. & Frederiksen, J.R. (1999). Enabling students to construct theories of collaborative inquiry and reflective learning: Computer support for metacognitive development. *International Journal of Artificial Intelligence in Education, 10*, 151-182.

White, B., Frederiksen, J. & Collins, A. (2009). The interplay of scientific inquiry and metacognition: More than a marriage of convenience. In D. Hacker, J. Dunlosky & A. Graesser (Eds.), *Handbook of metacognition in education* (pp. 175-205). New York: Routledge.

Wirth, J. & Leutner, D. (2006). Selbstregulation beim Lernen in interaktiven Lernumgebungen [Self-regulated learning in interactive learning environments]. In H. Mandl & H.F. Friedrich (Eds.), *Handbuch Lernstrategien* (pp. 172-184). Göttingen: Hogrefe.

Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität* [Interrater agreement and interrater reliability]. Göttingen: Hogrefe.

Wragg, E. (2005). *The art and science of teaching and learning: The selected works of Ted wragg*. New York: Routledge.

Yoon, C. (2009). Self-regulated learning and instructional factors in the scientific inquiry of scientifically gifted Korean middle school students. *Gifted Child Quarterly, 53*(3), 203-216.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review, 20*, 99-149.

8. Appendix

8.1 Science achievement test

1. Which term characterizes how tightly molecules are packed in a given volume?

   space ☐

   weight ☐

   density ☐

   mass ☐

2. An object has a mass of 20 g and a volume of 5 cm^3. What is its density?

   4 ☐

   15 ☐

   25 ☐

   100 ☐

3. The density of iron is given as about 8 g/cm^3. What is the mass of 48 cm^3 of iron?

   0.167 g ☐

   6 g ☐

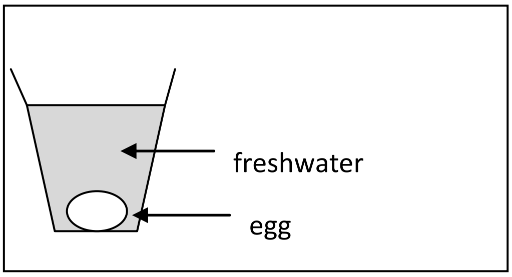
   0,038 g ☐

   384 g ☐

4. If the following solids all have the same mass, which one will have the greatest volume?

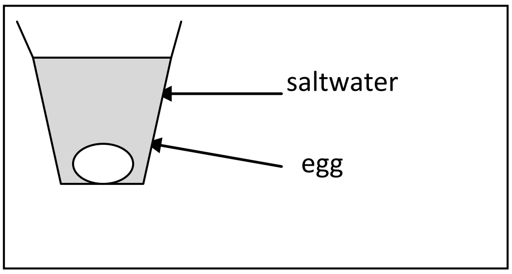
Solid	Density in g/cm^3	
Silver	10.5	☐
Glass	2.3	☐
Platinum	21.4	☐
Gold	19.3	☐

5. The following picture shows a fresh egg sinking in fresh water.

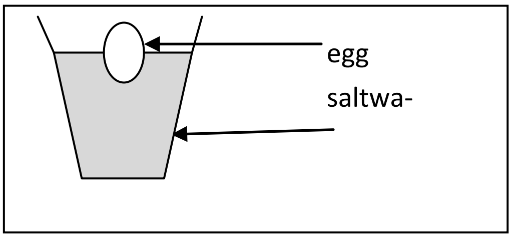


If this egg is placed in salt water from the ocean, which picture shows what would happen? Mark the correct picture.

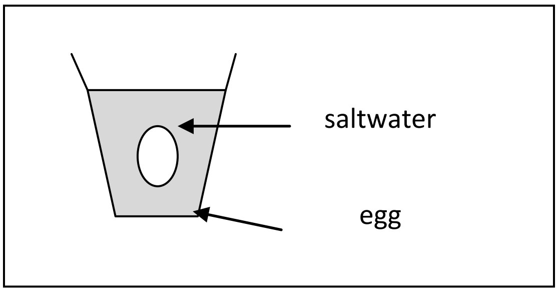
☐



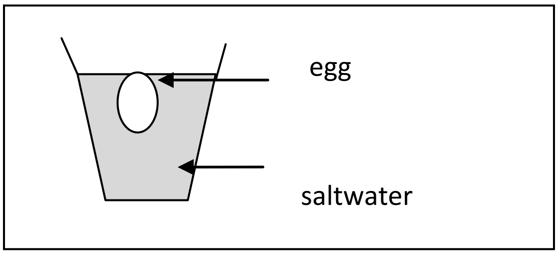
☐



☐



☐



6. An object has a mass of m = 7 g and a total volume of V = 18 cm^3. How will the object behave in water?

Sink	☐
Float or swim	☐
Neither float nor sink	☐
Sink at first, then float slowly	☐

7. What changes the density of objects?

mass and weight	☐
weight and volume	☐
volume and mass	☐
mass and shape	☐

8. How do we identify the substance of a homogeneous object?

By knowing its volume	☐
By knowing its density	☐
By knowing its weight	☐
By knowing its shape	☐

9. What happens to objects whose weight force is less than the buoyant force of the fluid they are floating in?

Sink	☐
Float or swims	☐
Neither float nor sink	☐
sink at first, then float slowly	☐

10. What happens to liquid with a density of 2 g/cm^3 when added to water?

- Floats or swims ☐
- Sinks ☐
- sinks at first, then floats slowly ☐
- Neither floats nor sinks ☐

11. A solid cylinder of plastic has a density of 1.6 g/cm^3. It is cut exactly in half. What is the density of each of the pieces?

- 1st half 1,3 g/cm^3 and 2nd half 1,3 g/cm^3 ☐
- 1st half 0,8 g/cm^3 and 2nd half 0,8 g/cm^3 ☐
- 1st half 1,6 g/cm^3 and 2nd half 1,6 g/cm^3 ☐
- 1st half 0,53 g/cm^3 and 2nd half 053 g/cm^3 ☐

12. What are the different states of matter?

- gas, solid and metal ☐
- solid, liquid and water ☐
- solid, gas and liquid ☐
- gas, liquid and oxygen ☐

13. Which state of matter has no fixed volume or shape?

- solid ☐
- gas ☐
- water ☐
- liquid ☐

14. A solid object floats in water when it is:

- light ☐
- heavy ☐
- more denser than water. ☐
- less dense than water ☐

15. According to the table, which wood will sink when placed in a fluid with a density of $\rho = 1.14\ g/cm^3$?

Type of wood	Density in g/cm^3	
African teakwood	0.98	☐
Balsa	1.14	☐
Cedar	0.55	☐
Ironwood	1.23	☐

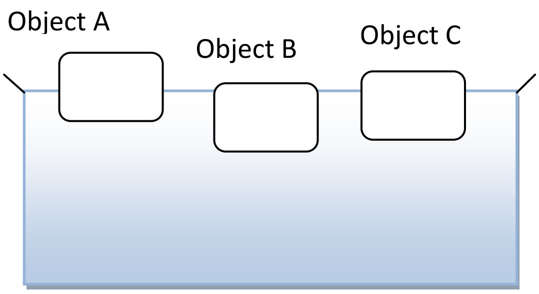
16. Archimedes' Principle helps to explain the relationship between

- energy and density ☐
- temperature and density ☐
- buoyancy and density ☐
- weight and density ☐

17. A piece of wood that is just below the surface of the water (not sinking or floating) has

- upward buoyancy ☐
- neutral buoyancy ☐
- downward buoyancy ☐
- no buoyancy ☐

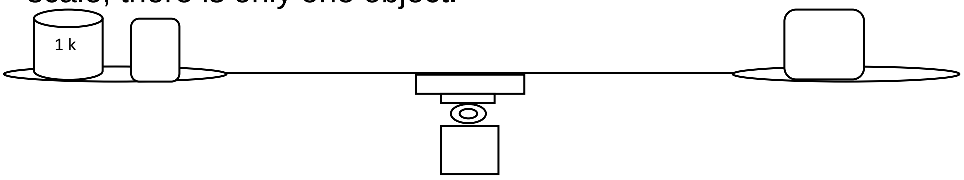
18. The picture shows three solid objects of the same volume floating in water.



Which object weights the most?

- Object A ☐
- Object B ☐
- Object C ☐
- They all weight the same ☐

19. The objects on the scale make it balance exactly. On the left pan there is a 1 kg weight (mass) and second object weighs half of that. On the right-hand side of the scale, there is only one object.



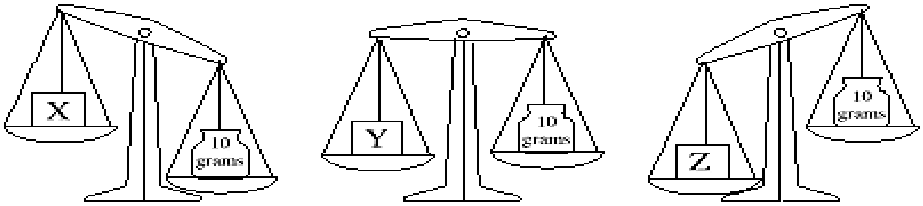
What is the mass (weight) of the object on the right-hand side of the scale?

- 1.5 k ☐
- 1 k ☐
- 2 k ☐
- 3 k ☐

20. Block sinks in the water in Container 1. When the same block is put in a big container with more water the block will ____________.

- Float ☐
- Sink ☐
- Sink at first, then float slowly ☐
- Neither float nor sink ☐

21. Which of the boxes X, Y, or Z has the LEAST mass?



X	☐
Y	☐
Z	☐
All three boxes have the same mass	☐

8.2 Experimental problem solving and strategy knowledge tests

Part (1): Experimental strategies knowledge items:

1. *You want to find out on which variables the property of buoyancy (whether a body sinks, floats or swims in water) is dependent.*

You consider the following procedures in order to answer the question. Review the procedures with the scale from *strongly agree* to *strongly disagree*:	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
a) I make a drawing with all possible variables.	❑	❑	❑	❑	❑
b) I check the individual variables one after the other.	❑	❑	❑	❑	❑
c) I try to explain the effect of each variable.	❑	❑	❑	❑	❑
d) I record the results of my experiments in a table.	❑	❑	❑	❑	❑
e) I check whether the results are correct or not.	❑	❑	❑	❑	❑
f) I define and represent the problem.	❑	❑	❑	❑	❑
g) I first formulate some hypotheses, which I then check in sequence.	❑	❑	❑	❑	❑

2. *Your class has conducted different experiments in small groups on the same topic. You shall now explain to the rest of the class what you have done in your group.*

You consider the following procedures in order to answer the question. Review the procedures with the scale from *strongly agree* to *strongly disagree*:	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
a) I describe in detail the results that are most important.	❑	❑	❑	❑	❑
b) I describe all experiments we did and what hypotheses we formed.	❑	❑	❑	❑	❑
c) I describe who did perfectly in our group.	❑	❑	❑	❑	❑
d) I describe what hypotheses we had and with which experiments we verified them.	❑	❑	❑	❑	❑
e) I choose one of the experiments and describe it in great detail.	❑	❑	❑	❑	❑

3. *You want to find out whether a room´s temperature changes if the door of the refrigerator in the room is kept open.*

You consider the following procedures in order to answer the question. Review the procedures with scale from *strongly agree* to *strongly disagree*:	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
a) I compare the temperature inside the fridge with the temperature outside the fridge.	❑	❑	❑	❑	❑
b) I measure the temperature of the room once while the door of the fridge is still open.	❑	❑	❑	❑	❑
c) I measure the temperature of the room before and after opening the door of the fridge.	❑	❑	❑	❑	❑
d) I measure the temperature of the room several times before opening, and several times after opening the door of the fridge.	❑	❑	❑	❑	❑
e) I measure the temperature of the room only when the windows and doors of the room are closed.	❑	❑	❑	❑	❑

4. *You want to explain to the rest of the class what you would do when faced with a science problem.*

You consider the following procedures in order to answer the question. Review the procedures with scale from *strongly agree* to *strongly disagree*:	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
a) I make a drawing with all possible variables.	❑	❑	❑	❑	❑
b) I check the individual variables one after the other.	❑	❑	❑	❑	❑
c) I try to explain the effect of each variable.	❑	❑	❑	❑	❑
d) I record the results of my experiments in a table.	❑	❑	❑	❑	❑
e) I define and represent the problem.	❑	❑	❑	❑	❑
f) I first formulate some hypotheses, which I then check in sequence.	❑	❑	❑	❑	❑
g) I check whether the results are correct or not.	❑	❑	❑	❑	❑

5. *You want to find out why, in saltwater, a heavy steel model of a ship floats but the same model in freshwater sinks?*

You consider the following procedures in order to answer the question. Review the procedures with scale from *strongly agree* to *strongly disagree*:	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
a) I make a drawing with all possible variables.	❑	❑	❑	❑	❑
b) I check whether the results are correct or not.	❑	❑	❑	❑	❑
c) I check the individual variables one after the other.	❑	❑	❑	❑	❑
d) I try to explain the effect of each variable.	❑	❑	❑	❑	❑
e) I record the results of my experiments in a table.	❑	❑	❑	❑	❑
f) I define and represent the problem.	❑	❑	❑	❑	❑

Part (2) Problem solving abilities items:

1. *Remember that the force of buoyancy influences whether an object sinks or floats. With respect to buoyancy, two important aspects of objects are mass and volume. You want to find out why, in water, a can of regular soda sinks, but a can of diet soda floats, so you'd like to measure the variables.*

You consider the following ideas in order to answer the question. Review the procedures with the scale from *strongly agree* to *strongly disagree*:	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
a) I compare the weights of the can of regular soda and the can of diet soda with my hands.	❑	❑	❑	❑	❑
b) I measure the mass of the can of regular soda and the mass of the can of diet soda using a scale.	❑	❑	❑	❑	❑
c) I measure the volumes of the can of regular soda and the can of diet soda using a scale.	❑	❑	❑	❑	❑
d) I measure the volume of water which the can of regular soda and the can of diet soda have displaced by using a scale.	❑	❑	❑	❑	❑
e) I think about the force of buoyancy and apply it to compare the weights of the can of regular soda and the can of diet soda.	❑	❑	❑	❑	❑
f) I compare the density of regular soda and diet soda.	❑	❑	❑	❑	❑

2. *You know that Archimedes's Principle states that the buoyant force on a submerged object is equal to the weight of the fluid that is displaced by the object. In order for an object to swim, it must displace enough water to equal its weight before it is fully submerged. An object will float if it weighs less than the weight of the volume of water it displaces. It will sink if it weighs more than the volume of the water it displaces. Ice has a density of 0.92 g/cm^3. Observing an ice cube floating in water (density = 1.00 g/cm^3), you want to find out if more or less of the ice cube will be submerged in seawater, so you decide to measure the variables.*
3.

You consider the following ideas that you want to verify experimentally in order to explain the observation. Review the procedures using the scale from *strongly agree* to *strongly disagree*:	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
a) I compare the density of the freshwater with the density of seawater.	❑	❑	❑	❑	❑
b) I compare the density of ice with the densities of freshwater and seawater.	❑	❑	❑	❑	❑
c) I measure the mass of the ice cube in freshwater and seawater by using a scale.	❑	❑	❑	❑	❑
d) I measure the volume of the ice cube in freshwater and seawater by using a scale.	❑	❑	❑	❑	❑
e) I think about the force of buoyancy and apply it to compare the weight of the ice cube in freshwater and seawater.	❑	❑	❑	❑	❑
f) I measure the density of the ice cube several times in freshwater and seawater.	❑	❑	❑	❑	❑
g) I compare the volume of water which the ice cube has displaced (in both freshwater and seawater).	❑	❑	❑	❑	❑
h) I compare the submerged part of the ice cube in freshwater and seawater.	❑	❑	❑	❑	❑

4. *Remember that the force of buoyancy is the upward force caused by fluid pressure that keeps things afloat. Something that is positively buoyant floats in water. Something that is negatively buoyant sinks in water. Things that are neutrally buoyant neither float nor sink in water. With regard to buoyancy, two important aspects of objects are mass and volume. In an experiment, it is shown that a heavy steel model of a ship floats, but a heavy solid block of steel sinks in freshwater. Which procedure can explain the observation?*

You consider the following ideas in order to answer the question. Review the procedures with the scale from *strongly agree* to *strongly disagree*:	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
a) I compare the weights of the solid block of steel and the steel model ship with my hands.	❑	❑	❑	❑	❑
b) I measure the mass of the steel ship model and the mass of the solid block and compare them.	❑	❑	❑	❑	❑
c) I measure the volumes of the solid block and the ship model by using a scale.	❑	❑	❑	❑	❑
d) I think about the force of buoyancy and apply it to compare the weights of the steel ship model and the solid block.	❑	❑	❑	❑	❑
e) I compare the volume of water which the steel ship model and the solid block have displaced.	❑	❑	❑	❑	❑
f) I compare the densities of the steel ship model and the solid block.	❑	❑	❑	❑	❑
g) I compare the densities of the steel ship model and the solid block with the density of freshwater.	❑	❑	❑	❑	❑

5. *Remember that Archimedes's Principle states that the buoyant force on a submerged object is equal to the weight of the fluid that is displaced by the object. In order for an object to swim, it must displace enough water to equal its weight before it is fully submerged. An object will float if it weighs less than the weight of the volume of water it displaces. It will sink if it weighs more than the volume of the water it displaces. You make the observation that a boat sinks lower into the water when you get into it. What procedures would you use to explain why this is?*

You consider the following ideas that you want to verify experimentally in order to explain the observation. Review the procedures with scale from *strongly agree* to *strongly disagree*:	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
a) I measure my mass and the mass of the boat and compare them.	❑	❑	❑	❑	❑
b) I measure the volume of the boat after I get onto it by using a scale.	❑	❑	❑	❑	❑
c) I think about the force of buoyancy and apply it to compare the weight of the boat and my weight.	❑	❑	❑	❑	❑
d) I compare the density of the boat after I get onto it and the density of water.	❑	❑	❑	❑	❑
e) I compare my weight and the weight of the boat by using a scale.	❑	❑	❑	❑	❑
f) I measure the density of the boat before and after I get onto it.	❑	❑	❑	❑	❑
g) I measure the volume of water which the boat has displaced before and after I get on.	❑	❑	❑	❑	❑

8.3 Motivation questionnaire

Deine Einstellung zu Naturwissenschaften und zum Lernen von Naturwissenschaften

1. Stimmst Du den folgenden Aussagen zu? *(Bitte in jeder Zeile nur ein Kästchen ankreuzen)*						
	stimme voll zu	*stimme zu*	*stimme eher zu*	*stimme eher nicht zu*	*stimme nicht zu*	*stimme gar nicht zu*
Es macht mir Spaß, mich mit naturwissenschaflichen Themen zu befassen.	❑	❑	❑	❑	❑	❑
b) Ich lese gerne etwas über Naturwissenschaften.	❑	❑	❑	❑	❑	❑
c) Ich beschäftige mich gerne mit naturwissenschaflichen Problemen.	❑	❑	❑	❑	❑	❑
d) Ich eigne mir gern neues Wissen in Naturwissenschaften an.	❑	❑	❑	❑	❑	❑
e) Ich bin interessiert, Neues über Naturwissenschaften zu lernen.	❑	❑	❑	❑	❑	❑

2. Stimmst Du den folgenden Aussagen zu? *(Bitte in jeder Zeile nur ein Kästchen ankreuzen)*						
	stimme voll zu	*stimme zu*	*stimme eher zu*	*stimme eher nicht zu*	*stimme nicht zu*	*stimme gar nicht zu*
a) In Zeitungen oder Zeitschriften Berichte über naturwissenschafliche Themen lesen.	❑	❑	❑	❑	❑	❑
b) Mich im Internet mit naturwissenschaflichenThemen beschäftigen.	❑	❑	❑	❑	❑	❑
c) In einem Lehrbuch, Nachschlagewerk (Lexikon) oder woanders nachlesen, wenn eine naturwissenschafliche Frage auftaucht.	❑	❑	❑	❑	❑	❑
d)In einer Gruppe mitmachen, die sich mit naturwissenschaflichen Inhalten beschäftigt.	❑	❑	❑	❑	❑	❑

3. Wie sehr stimmst Du den unten stehenden Aussagen zu? *(Bitte in jeder Zeile nur ein Kästchen ankreuzen)*						
	stimme voll zu	*stimme zu*	*stimme eher zu*	*stimme eher nicht zu*	*stimme nicht zu*	*stimme gar nicht zu*
a) Sich in Naturwissenschaften anzustrengen zahlt sich aus, weil mir das bei der Arbeit, die ich später machen möchte, helfen wird.	❑	❑	❑	❑	❑	❑
b) Was ich in Naturwissenschaften lerne, ist wichtig für mich, weil ich es für meine spätere Ausbildung brauche.	❑	❑	❑	❑	❑	❑
c) Ich lerne Naturwissenschaften, weil ich weiß, dass es für mich nützlich ist.	❑	❑	❑	❑	❑	❑
d) Naturwissenschaften zu lernen lohnt sich für mich, weil das Gelernte meine beruflichen Aussichten verbessern wird.	❑	❑	❑	❑	❑	❑
e) Ich werde in Naturwissenschaften viele Dinge lernen, die mir helfen werden, einen Job zu bekommen.	❑	❑	❑	❑	❑	❑

4. Die folgende Frage betrifft Deine Erfahrungen beim Lernen von Naturwissenschaften. Wie sehr stimmst Du den unten stehenden Aussagen zu? *(Bitte in jeder Zeile nur ein Kästchen ankreuzen)*						
	stimme voll zu	*stimme zu*	*stimme eher zu*	*stimme eher nicht zu*	*stimme nicht zu*	*stimme gar nicht zu*
a) Ich glaube, dass ich anspruchsvollen Stoff im Naturwissenschaftsunterricht leicht lernen kann.	❑	❑	❑	❑	❑	❑
b) Normalerweise kann ich Prüfungsfragen im Naturwissenschaftsunterricht gut beantworten.	❑	❑	❑	❑	❑	❑
c) Ich lerne neuen Stoff im Naturwissenschaftsunterricht schnell.	❑	❑	❑	❑	❑	❑
d) Den Stoff im Naturwissenschaftsunterricht finde ich einfach.	❑	❑	❑	❑	❑	❑
e) Wenn ich in Naturwissenschaften unterrichtet werde, verstehe ich neue Begriffe leicht.	❑	❑	❑	❑	❑	❑
f) Es fällt mir leicht, neue Ideen im Naturwissenschaftsunterricht zu verstehen.	❑	❑	❑	❑	❑	❑

Fragen zu Deinem Naturwissenschaftsunterricht

5. Wie oft macht Ihr die unten aufgelisteten Dinge in Eurem Naturwissenschaftsunterricht? (Bitte in jeder Zeile nur ein Kästchen ankreuzen)				
	jede/fast jede Stunde	*in der Hälfte der Stunden*	*In manchen Stunden*	*in keiner/ fast keiner Stunde*
a) Wir sehen der Lehrperson zu, wie sie einen Versuch durchführt.	❑	❑	❑	❑
b) Wir formulieren Fragen oder Vorhersagen, die überprüft werden müssen.	❑	❑	❑	❑
c) Wir arbeiten allein oder in kleinen Gruppen an einem Versuch.	❑	❑	❑	❑
d) Wir untersuchen die Auswirkungen von Naturwissenschaften auf die Gesellschaft.	❑	❑	❑	❑
e) Wir machen eine Exkursion oder ein Experte besucht unsere Klasse.	❑	❑	❑	❑
f) Wir lernen Dinge, die einen Bezug zum Alltag haben.	❑	❑	❑	❑
g) Die Lehrperson präsentiert uns die Inhalte im Vortragsstil.	❑	❑	❑	❑
h) Wir machen Notizen.	❑	❑	❑	❑
i) Wir lesen selbstständig im Lehrbuch.	❑	❑	❑	❑
j) Jeder arbeitet für sich an Aufgaben.	❑	❑	❑	❑
k) Wir arbeiten in Gruppen und präsentieren unsere Ergebnisse der ganzen Klasse.	❑	❑	❑	❑
l) Wir diskutieren in kleinen Gruppen über naturwissenschafliche Themen.	❑	❑	❑	❑
m) Die Lehrperson gibt uns Gelegenheit, unsere Ideen zu erklären.	❑	❑	❑	❑
n) Wir kontrollieren unsere Hausaufgaben.	❑	❑	❑	❑
o) Wir schreiben einen Test.	❑	❑	❑	❑

6. Wie gut findest Du den Naturwissenschaftsunterricht Deiner Lehrperson? *(Bitte in jeder Zeile nur ein Kästchen ankreuzen)*						
	trifft voll zu	*trifft zu*	*trifft eher zu*	*trifft eher nicht zu*	*trifft nicht zu*	*trifft gar nicht zu*
a) Die Lehrperson gestaltet den Unterricht spannend.	❑	❑	❑	❑	❑	❑
b) Die Lehrperson kann auch trockenen Stoff interessant machen.	❑	❑	❑	❑	❑	❑
c) Die Lehrperson kann die Klasse auch mal richtig begeistern.	❑	❑	❑	❑	❑	❑

7. Wie oft kommen die folgenden Dinge in Deinem Naturwissenschaftsunterricht vor? *(Bitte in jeder Zeile nur ein Kästchen ankreuzen)*						
	stimme voll zu	*stimme zu*	*stimme eher zu*	*stimme eher nicht zu*	*stimme nicht zu*	*stimme gar nicht zu*
a) Im Naturwissenschaftsunterricht wird ständig laut geredet.	❑	❑	❑	❑	❑	❑
b) Im Naturwissenschaftsunterricht wird andauernd Blödsinn gemacht.	❑	❑	❑	❑	❑	❑
c) Nach der Aufforderung leise zu sein, muss die Lehrperson lange warten, bis es auch wirklich ruhig ist.	❑	❑	❑	❑	❑	❑
d) Im Naturwissenschaftsunterricht ist es oft unruhig und laut.	❑	❑	❑	❑	❑	❑
e) Unsere Lehrperson muss uns oft daran erinnern, dass wir ruhig arbeiten sollen.	❑	❑	❑	❑	❑	❑
f) Im Naturwissenschaftsunterricht dauert es am Anfang der Stunde sehr lange, bis wir Schülerinnen und Schüler ruhig werden.	❑	❑	❑	❑	❑	❑

8. Wie geht Deine Naturwissenschaftslehrperson mit Störungen im Unterricht um? *(Bitte in jeder Zeile nur ein Kästchen ankreuzen)*						
	Stimme voll zu	*stimme zu*	*stimme eher zu*	*stimme eher nicht zu*	*stimme nicht zu*	*Stimme gar icht zu*
a) Unsere Lehrperson bemerkt alles, was in der Klasse vor sich geht.	❑	❑	❑	❑	❑	❑
b) Die Lehrperson greift gleich ein, wenn ein Schüler oder eine Schülerin anfängt zu stören.	❑	❑	❑	❑	❑	❑
c) Die Lehrperson merkt sofort, wenn jemand von uns Blödsinn macht.	❑	❑	❑	❑	❑	❑
d) Die Lehrperson bemerkt alles, was die Schülerinnen und Schüler machen.	❑	❑	❑	❑	❑	❑
e) Wenn jemand stört, macht die Lehrperson die Person so auf die Störung aufmerksam, dass nicht der ganze Unterricht gestört wird.	❑	❑	❑	❑	❑	❑
f) Die Lehrperson sorgt dafür, dass wir den Unterricht nicht stören.	❑	❑	❑	❑	❑	❑

9. Wie sehr stimmst Du den unten stehenden Aussagen über Regeln im Naturwissenschaftsunterricht zu? **(Bitte in jeder Zeile nur ein Kästchen ankreuzen)**						
	Stimme voll zu	*stimme zu*	*stimme eher zu*	*stimme eher nicht zu*	*stimme nicht zu*	*Stimme gar nicht zu*
a) Im Naturwissenschaftsunterricht gibt es bestimmte Regeln, an die wir uns halten müssen.	❑	❑	❑	❑	❑	❑
b) Mir ist klar, was man im Naturwissenschaftsunterricht machen darf und was nicht.	❑	❑	❑	❑	❑	❑
c) Alle Schülerinnen und Schüler kennen die Klassenregeln.	❑	❑	❑	❑	❑	❑
d) Ich weiß, was passiert, wenn wir die Regeln nicht einhalten.	❑	❑	❑	❑	❑	❑
e) Die Lehrperson hat uns genau erklärt, was im Naturwissenschaftsunterricht erlaubt ist und was nicht.	❑	❑	❑	❑	❑	❑

8.4 Teachers’ quality of instruction questionnaire

Regelklarheit	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Die Spielregeln, die eingehalten werden müssen, sind allen Schülern bekannt.	❑	❑	❑	❑
Der Lehrer hat klargemacht, was passiert, wenn die Schüler Regeln verletzten.	❑	❑	❑	❑
Den Schülern ist klar, was man machen darf und was nicht.	❑	❑	❑	❑
Gesamtbetrachtung Regelklarheit	❑	❑	❑	❑

Unterrichtsstörung (Disziplin)	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Im Unterricht wird andauernd Blödsinn gemacht.	❑	❑	❑	❑
In Naturwissenschaften wird der Unterricht sehr gestört.	❑	❑	❑	❑
Der Lehrer muss häufiger brüllen.	❑	❑	❑	❑
Im Unterricht wird fortwährend laut gequatscht.	❑	❑	❑	❑
Gesamtbetrachtung Unterrichtsstörung	❑	❑	❑	❑

Repetitives Üben	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Es wird sehr viel geübt und wiederholt, so dass die Klasse nur langsam vorankommt.	❑	❑	❑	❑
Es werden immer wieder fast dieselben Aufgaben geübt.	❑	❑	❑	❑
Gesamtbetrachtung Repetitives Üben	❑	❑	❑	❑

Anspruchsvolles Üben	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Unter den Übungsaufgaben sind oft Aufgaben, bei denen die Schüler sehen (können), ob sie etwas verstanden haben.	❑	❑	❑	❑
Beim Üben wenden die Schüler das Gelernte oft auf andere Dinge an.	❑	❑	❑	❑
Die Übungsaufgaben sind ähnlich, aber doch immer wieder anders, so dass die Schüler genau aufpassen müssen.	❑	❑	❑	❑
Gesamtbetrachtung Anspruchsvolles Üben	❑	❑	❑	❑

Diagnostische Kompetenz Sozialbereich	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer sieht nicht, wenn jemand Angst hat.	❑	❑	❑	❑
Der Lehrer spürt sofort, wenn zwischen Banknachbarn etwas nicht stimmt.	❑	❑	❑	❑
Der Lehrer merkt schnell, wenn jemand Kummer hat.	❑	❑	❑	❑
Der Lehrer spürt nicht, wenn jemand traurig ist und seine Gedanken woanders sind.	❑	❑	❑	❑
Der Lehrer sieht schnell, wenn es zwischen Schülern Streit gegeben hat.	❑	❑	❑	❑
Gesamtbetrachtung Diagnostische Kompetenz Sozialbereich	❑	❑	❑	❑

Diagnostische Kompetenz im Leistungsbereich	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer weiß, bei welchen Aufgaben die Schüler Schwierigkeiten haben.	❑	❑	❑	❑
Der Lehrer merkt sofort, wenn etwas nicht richtig verstanden wird.	❑	❑	❑	❑
Der Lehrer weiß sofort, was jemand nicht verstanden hat.	❑	❑	❑	❑
Gesamtbetrachtung Diagn. Kompetenz im Leistungsbereich	❑	❑	❑	❑

Positive Schülerorientierung	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer bemüht sich, die Wünsche der Schüler so weit wie möglich zu erfüllen.	❑	❑	❑	❑
Der Lehrer nimmt sich Zeit, wenn die Schüler etwas mit ihm bereden wollen.	❑	❑	❑	❑
Der Lehrer kümmert sich um die Probleme der Schüler.	❑	❑	❑	❑
Der Lehrer ist bereit, mit den Schülern zu reden, wenn diesen etwas nicht gefällt.	❑	❑	❑	❑
Gesamtbetrachtung Positive Schülerorientierung	❑	❑	❑	❑

Individuelle Bezugsnormorientierung	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Wenn sich jemand besonders angestrengt hat, lobt der Lehrer ihn, auch wenn andere Schüler noch besser sind.	❑	❑	❑	❑
Wenn ein Schüler seine Leistungen verbessert, wird er vom Lehrer gelobt, auch dann, wenn er im Vergleich zur Klasse unter dem Durchschnitt liegt.	❑	❑	❑	❑
Der Lehrer lobt auch die schlechten Schüler, wenn er merkt, dass sie sich verbessern.	❑	❑	❑	❑
Gesamtbetrachtung Individuelle Bezugsnormorientierung	❑	❑	❑	❑

Classroom Management	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer greift ein, bevor Unruhe und Störungen entstehen.	❑	❑	❑	❑
Der Lehrer reagiert zu spät, wenn Schüler Unsinn machen, so dass er dann massiver werden muss.	❑	❑	❑	❑
Der Lehrer unterbricht lange den Unterricht, wenn jemand Unsinn macht.	❑	❑	❑	❑
Gesamtbetrachtung Classroom Management	❑	❑	❑	❑

Sprunghaftigkeit	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer unterbricht die Einzelarbeit mit Aufforderungen, die an die ganze Klasse gerichtet sind, wenn ihm bei einem Schüler irgendetwas auffällt.	❑	❑	❑	❑
Der Lehrer lässt sich leicht ablenken, wenn ihm irgendetwas auffällt.	❑	❑	❑	❑
Der Lehrer fängt mitten in der Einzel- oder Gruppenarbeit an, etwas an der Tafel zu erklären.	❑	❑	❑	❑
Gesamtbetrachtung Sprunghaftigkeit	❑	❑	❑	❑

Motivierungsfähigkeit	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer gestaltet den Unterricht spannend.	❑	❑	❑	❑
Der Lehrer kann auch trockenen Stoff interessant machen.	❑	❑	❑	❑
Der Lehrer kann die Schüler auch mal richtig begeistern.	❑	❑	❑	❑
Gesamtbetrachtung Motivierungsfähigkeit	❑	❑	❑	❑

Klarheit und Strukturiertheit des Unterrichts	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer unterrichtet so, dass die Schüler auch schwierige Aufgaben bewältigen können, wenn sie sich anstrengen.	❑	❑	❑	❑
Der Lehrer fasst häufig noch einmal den Stoff zusammen, damit die Schüler ihn sich gut merken können.	❑	❑	❑	❑
Der Lehrer kommt vom *Hundertsten ins Tausendste*, und keiner weiß, was los ist.	❑	❑	❑	❑
Der Lehrer kann gut erklären.	❑	❑	❑	❑
Der Lehrer erklärt besonders an schwierigen Stellen ganz langsam und sorgfältig.	❑	❑	❑	❑
Gesamtbetrachtung Klarheit und Strukturiertheit des Unterrichts	❑	❑	❑	❑

Interaktionstempo	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer verlangt oft blitzschnelle Antworten.	❑	❑	❑	❑
Der Lehrer lässt bei Fragen kaum Zeit zum Nachdenken.	❑	❑	❑	❑
Der Lehrer geht gleich zum Nächsten, wenn ein Schüler nicht sofort antwortet.	❑	❑	❑	❑
Der Lehrer fragt oft unberechenbar in die Klasse und erwartet sofortige Antwort.	❑	❑	❑	❑
Gesamtbetrachtung Interaktionstempo	❑	❑	❑	❑

Pacing	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer geht im Stoff zügig voran, ohne die Schüler zu überfordern.	❑	❑	❑	❑
Der Lehrer bringt den Schülern auch beim Üben noch etwas Neues bei.	❑	❑	❑	❑
Gesamtbetrachtung Pacing (angemessen schnell)	❑	❑	❑	❑

Zeitverschwendung im Unterricht	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
In Naturwissenschaften dauert es lange, bis alle zur Arbeit bereit sind.	❑	❑	❑	❑
Zu Beginn der Stunde dauert es sehr lange, bis die Schüler ruhig werden und zu arbeiten beginnen.	❑	❑	❑	❑
In Naturwissenschaften enfehlt meist bei irgend jemandem etwas, wenn die Klasse anfangen will.	❑	❑	❑	❑
In Naturwissenschaften wird im Unterricht viel Zeit vertrödelt.	❑	❑	❑	❑
Gesamtbetrachtung Zeitverschwendung im Unterricht	❑	❑	❑	❑

Individualisierung des Unterrichts	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Die einzelnen Schüler bearbeiten oft verschiedene Aufgaben.	❑	❑	❑	❑
In Naturwissenschaften können schnellere Schüler schon zum Nächsten übergehen.	❑	❑	❑	❑
Der Lehrer stellt Schülern oder Schülergruppen unterschiedlich schwere Fragen, je nachdem, wie gut ein Schüler ist.	❑	❑	❑	❑
In Naturwissenschaften verlangt der Lehrer von guten Schülern deutlich mehr.	❑	❑	❑	❑
Gesamtbetrachtung Individualisierung des Unterrichts	❑	❑	❑	❑

Time on Task	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
In Naturwissenschaften arbeiten die Schüler intensiv mit.	❑	❑	❑	❑
In Naturwissenschaften machen die Schüler heimlich andere Dinge.	❑	❑	❑	❑
Gesamtbetrachtung Time on Task	❑	❑	❑	❑

Aggressionen gegenüber dem Lehrer	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Es gibt Schüler, die fast nie machen, was ihnen der Lehrer sagt.	❑	❑	❑	❑
Verschiedene Schüler geben dem Lehrer freche Antworten.	❑	❑	❑	❑
Manche Schüler äffen den Lehrer nach.	❑	❑	❑	❑
Bestimmte Schüler verhalten sich gegenüber dem Lehrer recht unverschämt.	❑	❑	❑	❑
Der Lehrer wird absichtlich geärgert.	❑	❑	❑	❑
Gesamtbetrachtung Aggressionen S >>>L	❑	❑	❑	❑

Aggressionen gegenüber den Mitschülern	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Einzelne Schüler ärgern ihre Mitschüler.	❑	❑	❑	❑
Es kommt vor, dass sich Schüler über ihre Mitschüler lustig machen.	❑	❑	❑	❑
Wenn ein Schüler etwas Falsches gesagt hat, lachen die anderen über ihn.	❑	❑	❑	❑
Die Schüler haben ein freundschaftliches Verhältnis zueinander.	❑	❑	❑	❑
In dieser Klasse helfen sich die Schüler gerne gegenseitig.	❑	❑	❑	❑
Gesamtbetrachtung Aggressionen S >>>S	❑	❑	❑	❑

Aggressionen gegenüber den Schülern	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Es kommt vor, dass sich der Lehrer über einen Schüler lustig macht.	❑	❑	❑	❑
Es kommt vor, dass der Lehrer einen Schüler nachäfft.	❑	❑	❑	❑
Manche Schüler werden vom Lehrer bevorzugt.	❑	❑	❑	❑
Es kommt vor, dass der Lehrer einen Schüler absichtlich benachteiligt.	❑	❑	❑	❑
Der Lehrer neigt zu Sarkasmus und zynischen Bemerkungen.	❑	❑	❑	❑
Gesamtbetrachtung Aggressionen L>>>S	❑	❑	❑	❑

Naturwissenschaften *Produktivität* der Schüler	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Formulierung von Ideen	❑	❑	❑	❑
aktives Frageverhalten	❑	❑	❑	❑
Steuerung des Unterrichtsgesprächs durch Schülerbeiträge	❑	❑	❑	❑
Gesamtbetrachtung Naturwissenschaften *Produktivität* der Schüler	❑	❑	❑	❑

Fehlerkultur	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer ist geduldig, wenn ein Schüler im Naturwissenschaftenunterricht einen Fehler macht.	❑	❑	❑	❑
Wenn der Lehrer selbst einen Fehler gemacht hat, gibt er dies offen zu.	❑	❑	❑	❑
Bei diesem Lehrer ist Fehlermachen nichts Schlimmes.	❑	❑	❑	❑
Der Lehrer achtet darauf, dass keiner ausgelacht wird, der einen Fehler macht.	❑	❑	❑	❑
Gesamtbetrachtung Fehlerkultur	❑	❑	❑	❑

Strukturierungshilfen	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer weist darauf hin, was sich die Schüler merken sollen.	❑	❑	❑	❑
Der Lehrer hebt wichtige Fakten hervor.	❑	❑	❑	❑
Der Lehrer fasst das Wichtigste nochmals zusammen.	❑	❑	❑	❑
Der Lehrer hält Rückblick auf das, was wichtig ist.	❑	❑	❑	❑
Der Lehrer fasst den Stoff nochmals zusammen.	❑	❑	❑	❑
Gesamtbetrachtung Strukturierungshilfen	❑	❑	❑	❑

Individuelle Lernunterstützung	trifft überhaupt nicht zu	trifft eher nicht zu	trifft eher zu	trifft voll und ganz zu
Der Lehrer muntert die Schüler auf, damit sie auch bei schwierigen Aufgaben nicht den Mut verlieren.	❑	❑	❑	❑
Der Lehrer bespricht schwierige Aufgaben auch einzeln mit den Schülern.	❑	❑	❑	❑
Der Lehrer hilft den Schülern auch einzeln, wenn sie bei einer Aufgabe nicht weiter wissen.	❑	❑	❑	❑
Der Lehrer kümmert sich darum, wenn ein Schüler beim Lösen von Aufgaben Probleme hat.	❑	❑	❑	❑
Gesamtbetrachtung Individuelle Lernunterstützung	❑	❑	❑	❑

8.5 Perceptions on the quality of science lessons questionnaire

How often do you do these things in your science lessons? *(Fill in **one** box for each line)*	Every or almost every lesson	About half the lessons	Some lessons	Never
1. We work in small groups on an experiment or investigation.	❑	❑	❑	❑
2. We design or plan an experiment or investigation.	❑	❑	❑	❑
3. The teacher organizes the discussion for groups.	❑	❑	❑	❑
4. We make observations and describe what we see.	❑	❑	❑	❑
5. We write up variables, control constants, hypotheses.	❑	❑	❑	❑
6. We conduct an experiment or investigation to verify our ideas and hypotheses.	❑	❑	❑	❑
7. In every group, each student is responsible for a certain task as a group member.	❑	❑	❑	❑
8. No student may finish group work until all teammates have mastered the subject as a group.	❑	❑	❑	❑
9. We ask our teammates for help before asking the teacher.	❑	❑	❑	❑
10. We think aloud when we solve a problem.	❑	❑	❑	❑
11. We write down plans for our experiments and predictions about the results of our activities on paper before conducting the experiments.	❑	❑	❑	❑
12. We present our ideas, variables, constants and hypotheses to the other groups.	❑	❑	❑	❑
13. We identify and formulate the problem.	❑	❑	❑	❑
14. We speak about what we already know about a given problem.	❑	❑	❑	❑

How often do you do these things in your science lessons? *(Fill in **one** box for each line)*	Every or almost every lesson	About half the lessons	Some lessons	Never
15. We formulate expected hypotheses.	❑	❑	❑	❑
16. We explore a possible way of solving a given problem.	❑	❑	❑	❑
17. We perform the solving process after we explore a way for solving a given problem.	❑	❑	❑	❑
18. We use scientific formulas and laws to solve problems.	❑	❑	❑	❑
19. We evaluate the final results with our hypotheses when we are solving a problem.	❑	❑	❑	❑
20. We give explanations about our results.	❑	❑	❑	❑
21. We listen to the teacher when he or she gives a presentation.	❑	❑	❑	❑
22. We work on problems on our own.	❑	❑	❑	❑
23. We listen to the teacher when conducting the solving processes.	❑	❑	❑	❑
24. We answer questions in class**.**	❑	❑	❑	❑

8.6 Lesson plans of the teaching program

Lesson 1: Density

<table>
<tr><td colspan="2">Title:
Density</td><td>Grade:
7</td></tr>
<tr><td>Curriculum:
Science</td><td>Concepts and Competencies:
Density</td><td>Duration:
90 minutes
(One session)</td></tr>
<tr><td colspan="3">Description:
Students will determine whether various objects sink or float in water. Whether an object sinks or floats in a liquid depends mainly on Archimedes´ Principle and density.</td></tr>
</table>

Standards:

Understand and apply concepts about the properties of matter regarding sinking, floating, and density.

Objectives:

By the end of this lesson, students will be able to:

- understand the concept of sinking, floating, and density.
- calculate density if given mass and volume.
- compare the densities of substances with the known density of water.

Science concepts:

- Objects which sink displace water of their own volume.
- Objects which float displace water of their own volume.
- Density equals mass per unit volume. The symbol of density is ρ.

Positive attitudes:

Students will…

- respect other students' opinions and predictions.

Background:

In this lesson plan, students become familiar with the concepts of density. Density, the ratio of the mass and volume of an object, is a physical property of matter. If two objects have the same volume but different mass, the object with more mass has a higher density.

Materials and Resources:

- Papers, pencils, worksheets
- 1 glass of water (filled to the very top); container filled with water; graduated cylinder or other measuring device; different boxes which contain the same volume of sugar, honey and flour; a balance scale; a ruler

Procedures:

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
7	**Preparation**	- Rank students' performance according to past assessments (base score). - Decide on the number of groups. - Assign students to homogeneous groups.	Form groups	Work group	Papers, pencils, worksheets	- Students will be conducting their own experiments. They will be required to record plans for their experiments and their predictions about the results of their activities on paper before conducting the experiments. - Each student should be responsible for a certain task as a group member. For example: writing notes, organizing the experiment, controlling the design, etc. - No students may finish group work until all teammates have mastered the subject as a group. - Students should ask teammates for help before asking the teacher. - Students should talk to each other softly and behave seriously.

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
7	**Introduction**	Challenge students to find common objects that will swim, float and/or sink in water.	Find objects that swim or float and sink in water	Work group	Paper, pencils, worksheets	- No students may finish group work until all teammates have mastered the subject as a group. - Students should ask teammates for help before asking the teacher. - Students should talk to each other softly and behave seriously.
15	**Problem**	Ask students the following question: Why is it that when we put different boxes filled with different kinds of substances such as sugar, honey and flour in water, some of them sink while others float?	-Write up their variables, control their constants, hypotheses -Record their data	Work group	Paper, pencils, worksheets, texts	- Each group of students will decide their problem and write it in their own words. - Teacher will monitor the students' work carefully (check if somebody is dominating or not participating and if all responsibilities are clear for students). - Teacher will offer assistance if the students ask.
15	**Ideas and hypotheses**	Ask each group to present their ideas, variables and hypotheses, and control their constants.	-Present their ideas, variables, control their constants, hypotheses -Other groups should give feedback.	Discussion	Paper, pencils, worksheets	- Teacher should organize the discussion for groups. - Every group should *think aloud* to show their teammates and the class how they solve their problem.

Time in minutes	Teaching Steps		Students activities	Work form	Materials	Comment
15	**Experiment**	Ask students to verify their ideas and hypotheses.	Complete the following experiment: investigate which boxes sink, float or swim in water	Work group	Worksheets, one container filled with water, different boxes with the same volume of sugar, honey, flour, a balance scale, a ruler, papers, pencils	- Students should have enough time to do the experiment. - The teacher will monitor the students' work carefully (check if somebody is dominating or not participating and if all responsibilities are clear for students). - Students measure and record the mass and volume of boxes. - Students should compare all the boxes at the same time in the water. - Teacher will offer assistance if the students ask.
20	**Explanation of the students**	Ask every group to present their results about the experiment to the other groups.	-Present their results -Other groups present feedback	Discussion	Worksheet, papers, pencils, posters	- Teacher should organize the discussion for groups. - Every group should *think aloud* to show their teammates and the class how they solve their problem.

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
20	**Explanation of the teacher**	Present and explain: -What swimming and sinking means -What the difference between mass and weight is -Density, the ratio of mass and volume of an object is a physical property of matter. If two objects have the same volume but different mass, the object with more mass has a greater density -We can calculate density using the formula: Density (ρ) = (m) Mass/ (V) Volume. So, volume (V) = Mass (m) / Density (ρ). Mass (m) = Density (ρ) X Volume (V). -Provide examples about density -- materials like steel are denser than water, so steel sinks in water. -Most woods are less dense than water, so they float. -Air and most other gases are less dense than water, so they float on the surface of water.	-Take notes -Give examples related to the densities of objects	Teacher explanation and discussion	Worksheet, papers, pencils, texts	- Teacher-centered discussion. The teacher should encourage students to discuss freely. - Students should read about the concept of density.

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
5	**Check understanding**	-Check students' understanding by asking the following question: -If you cut a piece of wood into 4 pieces what would happen to its density?	Read and answer the question from the worksheet	Work group	Papers, worksheet, pencils	- The teacher should encourage students to discuss the lesson freely (but quietly).

Lesson 2: The force of buoyancy

Title: Force of buoyancy		**Grade:** 7
Curriculum: Science	**Concepts and Competencies:** Density and buoyancy	**Duration**: 90 minutes (One session)
Description: Students will determine whether various objects sink or float in water. Whether an object sinks or floats in a liquid depends mainly on Archimedes´ Principle and density.		

Standards:

Understand and apply concepts about the properties of matter regarding sinking, floating, buoyancy and density.

Objectives:

By the end of this lesson, students will be able to…

- understand the concept of Archimedes´ Principle.
- understand that the buoyant force on an object is equal to the weight of the fluid it displaces.
- compare the densities of substances with the known density of water.
- explain that a totally immersed object displaces the volume of water equal to the volume of the object.

Science concepts:

Buoyancy is the upward force caused by fluid pressure that keeps things afloat. Something that is positively buoyant floats in water. Something that is negatively buoyant sinks in water. Neutrally buoyant neither floats nor sinks in water. An object becomes positively buoyant when it weighs less than the water it displaces and conversely is negatively buoyant when it weighs more than the water it displaces.

Positive attitudes:

Students will…

respect other students' opinions and predictions.

Background:

In this lesson plan, students become familiar with the concepts of Archimedes´ Principle (buoyancy). The Archimedes Principle states that the buoyant force on a submerged object is equal to the weight of the fluid that is displaced by the object. In order for an object to swim, it must displace enough water to equal its weight before it is fully submerged. An object will float if it weighs the same or less as the volume of water it displaces. It will sink if it weighs more than the volume of the water it displaces.

Materials and Resources:

- Papers, pencils, worksheets
- Ball, spring scale, bottle filled with water, different kinds of objects in water which could float, sink or swim

Procedures:

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
7	**Preparation**	Initiate group formation	Form groups	Work group	Paper, pencils, worksheets	- Each student should be responsible for a certain task as a group member. For example: taking notes, organizing the experiment, controlling the design, etc. - Students will be conducting their own experiments. They will be required to record their plans for their experiments and their predictions about the results of their activities on paper before conducting the experiments. - No students may finish group work until all teammates have mastered the subject as a group. - Students should ask teammates for help before asking the teacher. - Students should talk to each other softly and behave seriously.

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
7	**Introduction**	Review density, sinking, floating and ask questions related to it.	-Answer the questions -Take notes	Work group	Papers, pencils, worksheets	- Teacher should encourage students to think about the question before answering it and he or she should give students time to think.
15	**Problem**	Challenge students to explain: Why does the water rise differently with different objects that have not yet been fully submerged?	-Write up their variables, control their constants, hypotheses -Record their data	Work group	Paper, pencils, worksheets	- Each group of students will decide their problem and write it in their own words. - Teacher will monitor the students' work carefully (check if somebody is dominating or not participating and if all responsibilities are clear for students). - Teacher will offer assistance if the students ask

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
15	**Ideas and hypotheses**	Ask each group to present their ideas, variables, control their constants, hypotheses, etc.	-Present their ideas, variables, control their constants, hypotheses -Other groups should give feedback.	Discussion	Paper, pencils, worksheets	- Teacher should organize the discussion for groups. - Every group should *think aloud* to show their teammates and the class how they solve their problem.

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
15	**Experiment**	-Ask students to verify their ideas and hypotheses. -Ask students to do two tasks: a) Compare the weight and the mass of the ball with the weight and the mass of the water and also the weight which loss b) Compare the volume of the ball with the volume of the water	-Complete the following experiment: -Ball hanging from the spring scale is lowered into cylinder filled with water, which will overflow. -Complete the tasks Record data	Work group	Papers, pencils, worksheets, ball, spring scale, cylinder, Recording sheet, cube of wood, cube of metal, cube of glass	- The teacher will monitor the students' work carefully (check if somebody is dominating or not participating and if all responsibilities are clear for students). - The teacher should ask students to investigate different kinds of objects in water which could be (float, sink, swim). - Students should investigate the different influences of mass and volume separately. - The teacher will offer assistance if the students ask.

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
20	**Explanation of the students**	Ask every group to present their results about the experiment to the other groups.	-Present their results -Other groups present feedback	Discussion	Worksheet, papers, pencils, posters	- Teacher should organize the discussion for groups - Every group should use the laud thinking to show their way to solve the problem

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
20	**Explanation of the teacher**	• Present and explain: • If the weight of the displaced water is less than the weight of the object, the object will sink. • The object will float when the weight of the displaced water is equal to or less than the weight of the object (the buoyant force). • Archimedes´ Principle is experimentally applied to obtain the density ρ of a substance. • A body of weight W = m g has an apparent loss of weight when immersed in a fluid. This loss is equal to the weight of the volume of fluid displaced by the object or Wf = mf g. Hence, the apparent weight of the object W' submerged in the fluid is given by m' g = mg-m_fg , Since m = ρV one may replace mf of Eq. with p_fV where ρ_f is the density of the fluid and V the volume displaced. But the volume of fluid displaced is also the volume of the submerged object, V = m/ρ Thus, $m_f = \rho_f$ (m/ρ) and becomes, m'g = mg – ρ_f(m/ρ)g or w ' = w – (ρ_f/ρ)w and ρ = (w/w – w') ρ_f	-Read a text -Take notes -Lead a discussion	Teacher explanation and discussion	Worksheets, texts, papers, pencils	- Teacher-centered discussion. The teacher should encourage students to discuss freely. - Students should read about the concept of buoyancy.

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
5	**Check under-standing**	Check students' understanding by asking the following questions: A piece of metal has a volume of 40 cm^3 and a mass of 110g. 1) Calculate the density of the metal. 2) What type of metal is it?	Read and answer the question from the worksheet.	Work group	Worksheets, papers, pencils, calculators	The students should work in their groups quietly.

Lesson 3: Learning about problem solving and strategy knowledge

Title: Learning about problem solving		**Grade:** 7
Curriculum Area: Science	**Concepts and Competencies:** Problem solving	**Duration**: 90 minutes (One session)
Description: This lesson is an explanation of the stages of scientific inquiry and problem solving, and how they can be implemented in a scientific situation.		

Standards: Students should be able to…

1) understand problem solving processes and scientific inquiry by formulating usable questions and hypotheses, planning experiments, conducting observations, interpreting and analyzing data, drawing conclusions and communicating results.
2) apply the principles of problem solving.

Objectives:

By the end of this lesson, students should be able to …

- demonstrate an understanding of the problem solving model.
- understand the stages of the model.
- use the model of problem solving.

Positive attitudes:

Students should …

- respect other students' opinions and predictions.
- learn to withhold judgment and assumptions until they have sufficient information to reach a conclusion.

Background:

In this lesson plan, students become familiar with the concepts of problem and the model of problem solving.

Materials and Resources:

- Papers, pencils, worksheets,
- Small plastic cup, three different cubes (stone, metal, wood),
- Hard copies of the model of problem solving

Procedures:

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
7	**Preparation**	Initiate group formation	Form groups	Work group	Papers, pencils, work-sheets	Each student should be responsible for a certain task as a group member such as taking notes, organizing the experiment, controlling the design.
7	**Introduction**	Ask students for the problems they had to solve during the last lesson	Answer the question	Discussion	Papers, pencils, worksheet	The teacher should encourage the students to discuss the topic freely.
5	**Presentation of the model**	Explain the following stages of the model of problem solving: 1. Identifying and formulating the problem 2. Activating pre-knowledge related to the problem 3. Defining and representing the problem 4. Formulating an expected result (hypotheses) 5. Exploring a possible way of solving the problem (variable discrimination) 6. Performing the solving process 7. Fixing data and calculating 8. Looking back to the idea (hypotheses) and evaluating	Read the stages of the model of P.S. from the work-sheet. Take notes.	Discussion	Papers, pencils, tests, work-sheets	- The students should first quietly read the stages of the model from the worksheet. - The teacher should then encourage the students to discuss the topic freely. - The teacher should give the students an appropriate amount of time to finish reading and discuss the stages.

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
15	**Problem**	Present the following problem: Three objects: a rubber eraser, a wooden block and a potato were put into a tank of water. a) Why do you think the wooden block floats in the water? b) Why do you think the eraser and the potato sink in the water?	Write up their variables, control their constants, hypotheses Record their data.	Work group	Paper, pencils, worksheets	- Each group of students will decide their problem and write it in their own words. - The teacher will monitors the students' work carefully (check if somebody is dominating or not participating and if all responsibilities are clear for students). - The teacher will offer assistance if the students ask.
15	**Ideas and hypotheses**	Ask each group to present their ideas, variables, control their constants, hypotheses.	Present their ideas, variables, control their constants, hypotheses Other groups should give feedback	Discussion	Paper, pencils, worksheets	- Teacher should organize the discussion for groups. - Every group should *think aloud* to show their teammates and the class how they solve their problem.

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
15	**Experiment**	- Ask students to verify their ideas and hypotheses.	Complete the following experiment: put three objects - a rubber eraser, a wooden block and a potato into a tank of water. Record their observations.	Work group	Papers, pencils, worksheets, ball, spring scale, bottle, recording sheet, cube of wood, rubber eraser, potato	- The teacher will monitor the students' work carefully (check if somebody is dominating or not participating and if all responsibilities are clear for students). - Students should compare all the objects at the same time in the water. - Students should investigate the different influences of mass and volume separately. - The teacher will offer assistance if the students ask.
5	**Results of the students**	Ask every group to present their results about the experiment to the other groups.	Present their results Other groups present feedback	Discussion	Worksheet, papers, pencils, posters	- The teacher should organize the discussion for groups. - Every group should *think aloud* to show their teammates and the class how they solve their problem.
10	**Modeling of the teacher**	Describe how we can solve the problem.	Observe how the teacher solves the problem. Take notes Give feedback by using the stages of problem solving.	Discussion	Worksheets, papers, pencils, copies of model of P.S.	- The teacher should encourage students to discuss freely. - The teacher should focus on the importance of following the model to solve the problem. - The teacher should use *thinking aloud* method to model problem solving.

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
10	**Check under-standing**	- Select two students to engage in a dialogue about the model of prob-lem solving in front of the class.	Engage in a dia-logue Other groups give feedback by using the model of prob-lem solving.	Discussion	Texts, pa-pers, pen-cils, work-sheets	- The teacher will monitor the students' work carefully. - The teacher should organize the discussion. - The students should work in their groups quietly.

Lesson 4: Problem solving (activity I)

<table>
<tr><td colspan="2">Title:
Problem solving activity</td><td>Grade Range:
7</td></tr>
<tr><td>Integrated Curriculum Area:
Science</td><td>Concepts and Competencies:
buoyancy, density and sinking or floating</td><td>Duration:
90 Minute
(One session)</td></tr>
<tr><td colspan="3">Description:
This lesson is an example of how we can implement the practice of scientific inquiry and problem solving in the context of swimming and sinking.</td></tr>
</table>

Standards:

1) Apply concepts about buoyancy, density and sinking or floating.
2) Develop problem-solving and inquiry skills reflected by formulating usable questions and hypotheses, planning experiments, conducting observations, interpreting and analyzing data, drawing conclusions and communicating results.

Objectives:

By the end of this lesson, students will be able to…

- apply the model of problem solving to a new problem.
- test hypotheses and group objects by common characteristics regarding density, sinking, swimming or floating.
- make and test their predictions on whether various objects sink or float in water.

Positive attitudes:

- Respect other students' opinions and predictions.
- Learn to withhold judgment and assumptions until they have sufficient information to reach a conclusion.

Background:

In this lesson plan, students become familiar with the concepts of swimming and sinking and the problem solving processes.

Materials and Resources:

- hard copies of the model of problem solving,
- a glass full of water for teacher demonstration
- vegetable cooking oil, ice cube
- papers, worksheets, pencils, texts

Procedures:

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
5	**Preparation**	Initiate group formation	Form groups	Work group	Work-sheets, pencils, papers	Every student should be responsible for a certain task as a group member: taking notes, organizing the experiment, controlling the design, etc.
10	**Introduction**	Review the problem solving process and ask questions related to it.	Answer the questions	Discussion	Work-sheets, pencils, papers	The teacher should encourage students to discuss freely.
20	**Problem/ experiment**	• Ask students to read the tasks from the text and do the experiment. • Ask students the following question: Why does the cube of ice float on top of water and oil?	Do the following experiment: Fill the glass almost full of vegetable oil. Place the glass of oil on a flat surface and then add an ice cube. The cube of ice will immediately float on top of the oil. Record observations about the experiment.	Work group	A cup of glass, vegetable cooking oil, ice cube, Work-sheets, pencils, papers, text, copies of model of P.S.	- The students should organize the experiment, control the design, etc. - The students should have enough time to complete the experiment. - Students need to measure and record the mass and volume of the ice cube. - The teacher will monitor the students' work carefully (check if somebody is dominating or not participating and if all responsibilities are clear for students). - The teacher should ask students to use the model of problem solving.

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
15	**Applying the model of P.S. (working Phase)**	Ask students to apply the model of problem solving to solve the problem.	Apply the model of problem solving.	Work group	Hard copies of the model of P.S., Worksheets, pencils, papers	- Teacher should monitor the students' work carefully (check if somebody is dominating or not participating and if all responsibilities are clear for students). - The teacher should show other groups which group is working most effectively on the problem. - Students should work quietly.
25	**Presentation of the students**	Asking every group of students to present the procedures they used to solve the problem.	Present the procedures they have used to solve the problem. Other groups should give feedback by using the stages of problem solving.	Discussion	Poster, paper, pencil, worksheets	- Students should present their work as a poster or an oral presentation. - Students should include reading their step-by-step procedure out loud and show how they solved the problem. - Students should discuss problem solving processes.
15	**Check understanding**	Check students' understanding by asking the following question: Candles float in water. a) Please draw a picture of what you think would happen with a very big candle. Explain why you think this would happen. b) Please draw a picture of what you think the candle would do in a tank with a lot of water. Explain why you think this would happen.	Read and answer the question from the worksheet	Work group	Worksheets, papers, pencils	- Teacher should encourage students to discuss freely. - The students should work in their groups quietly.

Lesson 5: Problem solving (activity II)

<table>
<tr><td colspan="2">Title:
Problem solving activity</td><td>Grade Range:
7</td></tr>
<tr><td>Integrated Curriculum Area:
Science</td><td>Concepts and Competencies:
buoyancy, density and sinking or floating</td><td>Duration:
90 Minute
(One session)</td></tr>
<tr><td colspan="3">Description:
This lesson is an example of how we can implement the practice of scientific inquiry problem solving in the context of swimming and sinking.</td></tr>
</table>

Standards:

1) Apply concepts about buoyancy, density and sinking or floating.
2) Develop problem-solving and inquiry skills reflected by formulating usable questions and hypotheses, planning experiments, conducting observations, interpreting and analyzing data, drawing conclusions and communicating results.

Objectives:

By the end of this lesson, students will be able to…

- apply the model of problem solving to a new problem.
- test hypotheses and group objects by common characteristics regarding density, buoyancy.
- make and test their predictions on whether various objects sink or float in water.

Positive attitudes:

- Respect other students' opinions and predictions.
- Learn to withhold judgment and assumptions until they have sufficient information to reach a conclusion.

Background:

In this lesson plan, students become familiar with the concepts of swimming and sinking and the problem solving processes.

Materials and Resources:

- Hard copies of the model of problem solving
- One letter of Honey, one letter of oil, one letter of water, 50 mL graduated cylinder for teacher demonstration
- Worksheets, papers, pencils, texts

Procedures:

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
7	**Prepara-tion**	Initiate group formation	Form groups	Work group	Worksheets, pencils, papers	Every student should be responsible for a certain task as a group member (i.e. writing notices, organizing the experi-ment, controlling the design, etc.)
7	**Introduc-tion**	Review the problem solving pro-cess and ask questions related to it.	Answer the questions	Discus-sion	Worksheets, pen-cils, papers	Teacher-centered discussion. The teacher should encourage students to discuss freely.
20	**Problem/ experi-ment**	• Ask students to read the tasks from the text and do the exper-iment. • Ask students the following question: Why is it that the liquids will layer on top of one another (the oil on top, the water in the mid-dle and the honey at the bot-tom)?	Do the following experi-ment: Pour honey into a cylin-der. After that pour both liquids - water and then oil -SLOWLY into the container, one at a time. Make sure that the liq-uids do not touch the sides of the cylinder while you are pouring. As you pour, the liquids will layer on top of one an-other (the oil on top, the water in the middle and the honey at the bottom). Recording their observa-tions about the experi-ment	Work group	Honey, oil, water cylinder, work-sheets, pencils, papers, text	- Students should organize the experi-ment, control the design, etc - The students should have enough time to complete the experiment. - The teacher should ask students to use the model of problem solving. - Students should compare all the liq-uids at the same time. - Teacher will monitor the students' work carefully (check if somebody is dominating or not participating and if all responsibilities are clear for stu-dents).

Time in minutes	Teaching Steps		Student activities	Work form	Materials	Comment
20	**Applying the model of P.S. (working phase)**	Ask students to apply the model of problem solving to solve the problem.	Apply the model of problem solving to solve the problem.	Work group	Hard copies of the model of problem solving, worksheets, pencils, papers	- The teacher should ask students to share their predictions with the others in the same group. - The teacher should show other groups which group is working on the problem most effectively. - The teacher should monitor the students' work carefully (check if somebody is dominating or not participating and if all responsibilities are clear for students).
20	**Presentation of the students**	Asking every group of students to present their procedures to solve the problem as a poster or an oral presentation.	Present their procedures which they used to solve the problem. Other groups should give feedback by using the stages of problem solving.	Discussion	Worksheets, posters, papers, pencils	- The students should include reading their step-by-step procedure to show how they solved the problem. - Students should discuss problem solving processes.
15	**Conclusion**	• Ask students to summarize what they have learned about Archimedes' Principle. • Ask students to summarize what they have learned about problem solving.	Read and answer the question from the worksheet.	Discussion	Worksheet, posters, papers, pencils, texts	- Teacher should encourage students to discuss freely. - Students should work in their groups quietly.

8.7 Literature

Dingrando, L., Fisher, D., Gonya, J., Haase D.G., Klevickis, C., Turiel, I., K. Zorn, M. & Zike, D. (2007). *Focus on physical science: California grade 8: Interactive student edition*. New York: The McGraw-Hill Companies, Inc.

Kendig, K. (2008). *Sink or float? thought problems in math and physics*. Washington, DC: The Mathematical Association of America.

Yin, Y., Tomita, M. & Shavelson, R.J. (2008). Diagnosing and dealing with student misconceptions about "sinking and floating". *Science Scope, 31*(8), 34-39.

Websites:

Buoyancy (n.d.). Retrieved (5.4.2009) from http://teachers.oregon.k12.wi.us/sundstrom/Physical%20Science/Measurement/Buoyancy%20Submarine%20Lab.pdf

Collaborative problem solving (n.d.). Retrieved (27.3.2012) from http://www.phy.ilstu.edu/~ghr/phy108/ProblemSet11Key.pdf

California department of education (2008). California standards test, Retrieved 27.4.2009, from: http://www.cde.ca.gov/ta/tg/sr/documents/rtqgr8science.pdf

Density and buoyancy-chapter-4 (n.d.). Retrieved (27.5.2012) from:

http://www.bickfordscience.com/04-Density_and_Buoyancy/02-buoyancy.html

Density and buoyancy (n.d.). Retrieved (2.6.2012) from: http://geometry.sierramadre.pasadenausd.org/modules/locker/files/get_group_file.phtml?gid=1380344&fid=12638719&sessionid

Harvey, J. (2003). Buoyant force, Retrieved (16.5.2012) from: http://www.cmu.edu/gipse/materials/pdf-2003/10-12/buoyant-J_Harvey.pdf

OUSD Middle School Science (2002). 8th Grade physical science curriculum, Retrieved (25.5.2012) from: http://tlc.ousd.k12.ca.us/~acody/physci1.html

Part II: Chemistry & Matter (n.d.). Retrieved (21.7.2012) from: http://hpms.hpisd.org

Patterson, c, Kennedy, T. & Miller, T. (n.d.). Lesson plans on density for middle school teachers, Retrieved (27.4.2012) from: http://www.docstoc.com/docs/9453531/Lesson-Plans-on-Density-for-Middle-School-Teachers

Sniffen, M. & Hardy, M. (2009). Modeling the Trieste to explore density and buoyant force, *Marine Technology Society Journal*, Retrieved (3.4.2012) from: http://www.marinetech.org/rov_competition/2010/Trieste_Model_Detailed_Assembly_Instructions.pdf

The University of the state of New York (2009). Grade 4-elementary-level-science test, Retrieved (3.4.2012) from: http://www.nysedregents.org/Grade4/Science/home.html

Unit 1: Introduction to chemistry's measurement and problem solving (n.d.). Retrieved (5.4.2012) from: http://www.docstoc.com/docs/43575265/Unit-1-Introduction-to-Chemistrys-Measurement-and-Problem-Solving

School district 202 (n.d.) Chapter 11 practice test, Retrieved (18.5.2012) from: http://www.learningcommunity202.org/

8.8 List of figures

8.9 List of tables