Journal of Innovation in Psychology, Education and Didactics	Vol. 19, No. 2 2015	117 - 148

A STUDY ON CREATIVE THINKING SKILLS OF GIFTED CHILDREN AND THEIR DEVELOPMENT OVER HALF A YEAR IN THE BIOLOGY EDUCATION PROJECT "KOLUMBUS-KIDS"

CLAAS WEGNER, MAX BENTRUP, STEPHANIE OHLBERGER [a]*

[a] Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany

Abstract

This paper describes the development of a test concerned with creative thinking in science education. It serves as a measuring instrument in the science education project "Kolumbus-Kids" at Bielefeld University and is based on a test for the analysis of deductive and creative thinking. Creating the test included the age-appropriate adjustment of the items as well as the adaptation of the questions to the context of science education. The central thesis concerns the connection of creative thinking and intelligence. The relationship is of particular interest in the project "Kolumbus-Kids", since this project is catered to gifted pupils. Means for promoting creative thinking skills were studied in the literature and could be compared with the project's didactic concept, thus allowing the assumption that the lessons of "Kolumbus-Kids" indeed enhance the pupils' creative thinking. With regard to these findings, the central hypothesis stating that pupils of grades four and six will significantly increase their performance can be verified. The significance of gender-specific differences is also examined, but these seem to have no influence. Both boys' and girls' performance increased significantly over the six months of participation in "Kolumbus-Kids".

Key words: Biology, creative thinking, giftedness, Kolumbus-Kids

Introduction

According to empirical findings, thinking creatively often correlates with intelligence and is thus a very interesting construct with regard to programmes for gifted pupils. Since creative thinking is

* Corresponding Author: Stephanie Ohlberger, *E-mail address*: stephanie.ohlberger@uni-bielefeld.de

a broad construct, a preliminary definition will be given first. The concept's understanding on which this study is based is guided by the definition of Urban (2000) and Kuhl (2001). The creative thinking skills of the pupils in the project were measured with the help of a self-developed test for scientifically gifted pupils. The test is partly based on the "test for the analysis of deductive and creative thinking" (in German: *Test zur Analyse des Schlussfolgernden und Kreativen Denkens*, for short*: ASK-Test*) by Schuler and Hell (2005).

Being able to think creatively is a key competence for professionals in science and research, teaching, business management, technology and many other areas. The recent issue of skilled worker shortage can only be counteracted with diverse affirmative action (Koppel & Plünnecke, 2009, p. 10-11). Orloff (2005) considers the promotion of creative thinking skills a substantial part of this. When analysing the pedagogic-didactical alignment of the project "Kolumbus-Kids", it becomes obvious that the project involves many of the building blocks for the promotion of creative thinking suggested by Urban. Therefore, it seems reasonable to assume that the project's lessons can increase the participants' creative thinking skills.

The Project "Kolumbus-Kids"

The study is set in a project at Bielefeld University that promotes scientifically gifted pupils. The project itself and its pedagogical alignment are presented next.

"Kolumbus-Kids" was founded by Prof. Claas Wegner as an extracurricular project devoted to the promotion of gifted pupils and is located in the department of Biology Didactics at Bielefeld University (Wegner, 2008, p. 28). At the moment, the project comprises five courses with approximately 18 children each. Three of the courses are for pupils of grades five and six and two courses for primary school children (fourth grade). A course instructor and up to four university students oversee the courses (see www.kolumbus-kids.de). Normally, secondary schools in North Rhine-Westphalia have classes of 26 to 30 pupils being taught by one teacher (Ständige Konferenz der Kultusminister der Länder, 2013, p. 11). Thus, the project courses feature a comparatively high instructor-student ratio.

Working in the project involves the development and evaluation of teaching concepts that focus on the special educational needs of gifted pupils (Wegner, 2008, p. 28). The project thereby aims at preventing the pupils' mental underload and at fostering their abilities and potential instead. For that purpose, recent findings in psychology and pedagogy as well as results of the project's own studies are incorporated into the lesson planning (Wegner, 2008, p. 28).

Besides the teaching of certain content, individual development and the fostering of social competences, as put forward by Mönks (1987), are focused on (Wegner, 2008, p. 28). Special emphasis is put on primary experience which cannot be taken for granted in today's digitalised world (Wegner, 2008, p. 29). "Kolumbus-Kids" therefore aims at providing a space for pupils to gain primary experience in a university setting. Through the intensive examination of natural scientific phenomena, an emotional bond to nature and the sciences is meant to be created. The interest thereby developed is thought to increase the pupils' motivation. By compiling natural scientific questions, interest is aroused during group work with like-minded people, which contributes to the promotion of social competences. The project lessons are intended to be problem-oriented, solution-centred and action-oriented, since Wegner considers the following aspects relevant (Wegner, 2008, p. 29):

1. Confrontation of the pupils with natural phenomena. Researching the processes plays an important role in that regard (observing, seeing, knowing and interpreting).

2. To a great extent, autonomous analysis of natural phenomena and finding explanations on account of the acquired knowledge in the lesson.

3. Primary experience with plants and animals is to arouse pupils' interest and should be regarded as the basis for life-long learning.

4. The pupils learn by 'anchoring' authentic challenging situations in the form of interesting phenomena and problems.

5. The observation and comprehension of biological processes on a real object fosters creativity and self-initiative.

In order to reach these goals, the pupils are first familiarised with the scientific processes and learning techniques needed. Gathering information independently and understanding and presenting findings are key competences that are to be mastered by the children. This process is encouraged by relevant, motivating and practice-oriented topics and is continuously evaluated with the help of internally developed diagnostic methods and teaching concepts (Wegner, 2008, p. 29).

Theoretical Background

Creative thinking – probably few other psychological terms have been used in everyday speech in such a lustrous and diverse way. In everyday life and in the arts people often label themselves ‘creative’. Being creative is fashionable and can even be understood as a life motto (Urban, 2000, p. 117). But what does the term comprise from a scientific viewpoint? This is examined in detail in order to find a suitable definition. After that, a close investigation of the connection between giftedness and creative thinking is carried out, making use of a short excursus into business psychology. Since companies see creative thinking skills as an important component in their employees, particularly in the areas of product design, marketing or leading positions, the call for early promotion of these skills in and outside of school goes without saying. Torrance’s (1989) TALENT study showed that lessons organised around problem-solving and abstract thinking make for a significant increase in pupils’ creative thinking skills (Urban, 2004, p. 90). Nonetheless, measuring the construct of creative thinking remains a problematic issue – in this context, the ASK-Test by Schuler and Hell (2005), which served as a basis for testing pupils in the project “Kolumbus-Kids”, is presented.

Creative thinking

The word “create” stems from Latin *creare*, which means to bear, create or produce something (Linneweh, 1973, p. 15). This derivation already emphasises the fact that creativity is to be considered a dynamic, developing process that contains both the origin and the goal (Linneweh, 1973, p. 15). Since this description is applicable to many processes, a precise definition is difficult to arrive at. Urban (2000) states that there is no universally accepted theoretical construct of creativity available so far. This can partly be attributed to the term’s many shades of meaning in everyday speech and also to the existence of different academic creativity concepts (Urban, 2000, p. 123).

Various understandings of creativity exist in psychology alone (Urban, 2000, p. 123), such as cognition- or personality-oriented concepts. Definitions with a focus on personality particularly emphasise a person’s willingness of exertion, motivation and ambiguity tolerance (Urban, 2004, p. 47). Cognitive approaches bear on divergent thinking and on both general and specific knowledge bases (Urban, 2004, p. 47). In addition, further definitions of creativity can be found in philosophy and in chaos theory (Urban, 2000, p. 123). In philosophy, creativity can mean manifold, for instance connecting something that was not connected before, discovering analogies and similarities, choosing and transcending levels of abstraction, and dealing with standards and regulations (von Wissel, 2012, p. 13). Some approaches also take neurophysiological findings on the functioning of the human brain into account (Urban, 2004, p. 45).

As has become obvious, Ausubel's (1968) statement that the term creativity is one of the most confusing and ambiguous terms in psychology and pedagogy still holds today (as cited in Urban, 2000, p. 123). A preliminary definition on creative thinking skills thus seems necessary. Specialised literature often describes creativity as closely linked to problem solving (Wentura & Frings, 2013, p. 152). Problem-solving consists of several judgment and decision processes that involve creative thinking by finding and applying new and efficient operators (Wentura & Frings, 2013, p. 152). According to this view, doing or thinking of something unusual is not creative in itself, except when the ideas are actually productive (Wentura & Frings, 2013, p. 152). Creative thinking therefore exceeds the artistic and aesthetic dimensions of the term's everyday use and has to be regarded from a pragmatic viewpoint that takes into account its target-oriented character (Linneweh, 1973, p. 13). Urban (2000) proposes six aspects that qualify as a definition attempt for creative thinking. According to him, creativity is manifested as a new, surprising, and also to others innovative, meaningful project (Urban, 2000, p. 124). Further, creative thinking is characterised by the following six facets (Urban, 2000):

1. The encountered project is a result of a sensitively perceived or prescribed problem that is internalised and whose implications have been perceived sensorily.

2. This process should be based on a comprehensive and at the same time sensitive perception of existing and openly accessible or given information and data as well as on acquired knowledge.

3. Creative thinking is further characterised by analytical procedures, solution-oriented actions and a likewise highly flexible approach. Unpopular and innovative reorganisations and reformulations or combinations of information, data and imaginary elements have to be mentioned in this regard just as well.

4. The productive trait of creative thinking becomes apparent in the ability of shaping data, elements and structures by means of synthesising, structuring and composing in such a way that new solutions arise.

5. The result has to be a product or be elaborated in a product.

6. The developed product has to be established and communicated and has to be perceived as meaningful by others.

In the context of creative thinking, the word ‘problem’ has a broad meaning. It does not have to be a concrete question, but can also be understood as the idea to invent a new object with a certain goal in mind (Urban, 2000, p. 124). Also, a problem does not always have to be solved in a creative way (Urban, 2000, p. 124); in craftsmanship, work often has to be finished off directly and swiftly.

In addition to intuition, the process of thinking creatively involves cognition (Linneweh, 1973, p. 13). Thinking operations of many kinds could be mentioned in this context, such as digital and analogous, convergent and divergent, as well as productive and reproductive processes (Linneweh, 1973, p. 13).

Kuhl (2001) defines creative thinking with regard to his model of a systems-theoretical distinction of eight forms of achievement motivation (Kuhl, 2010, p. 588). He proposes a close link between creativity and curiosity. Several ideal motivational conditions for creative problem-solving can be found when there is a simultaneous activation of diverse curiosity and an active form of specific curiosity (Kuhl, 2001, p. 592). Originality, which is typically attributed to creativity, arises from the free exploration of remote solution possibilities, and specific curiosity makes for a purposeful diversive exploration (Kuhl, 2001, p. 592). This focus is given by a specific problem on the one hand, and the general purpose on the other hand, which becomes action-guiding once the extension brain has been activated (Kuhl, 2001, p. 592). Inhibitory and negative affects, however, cannot override the positive affect (Kuhl, 2001, p. 592). This sets Kuhl’s understanding apart from other definitions. Not only are the production of inventive associations and the purposefulness important, but also the tolerance of negative emotional states and the self-based testing of such states: creativity, thus, is inseparable from the development of the self-system (Kuhl, 2001, p. 592). According to Kuhl, the decisive competence is tolerating negative emotions that are associated with creativity. This additional definition shall also be taken into account in order to complement Urban’s six aspects. Figure 1 provides a graphical depiction of the preliminary definition and shows how the factors mentioned in the definition relate to the project “Kolumbus-Kids”.

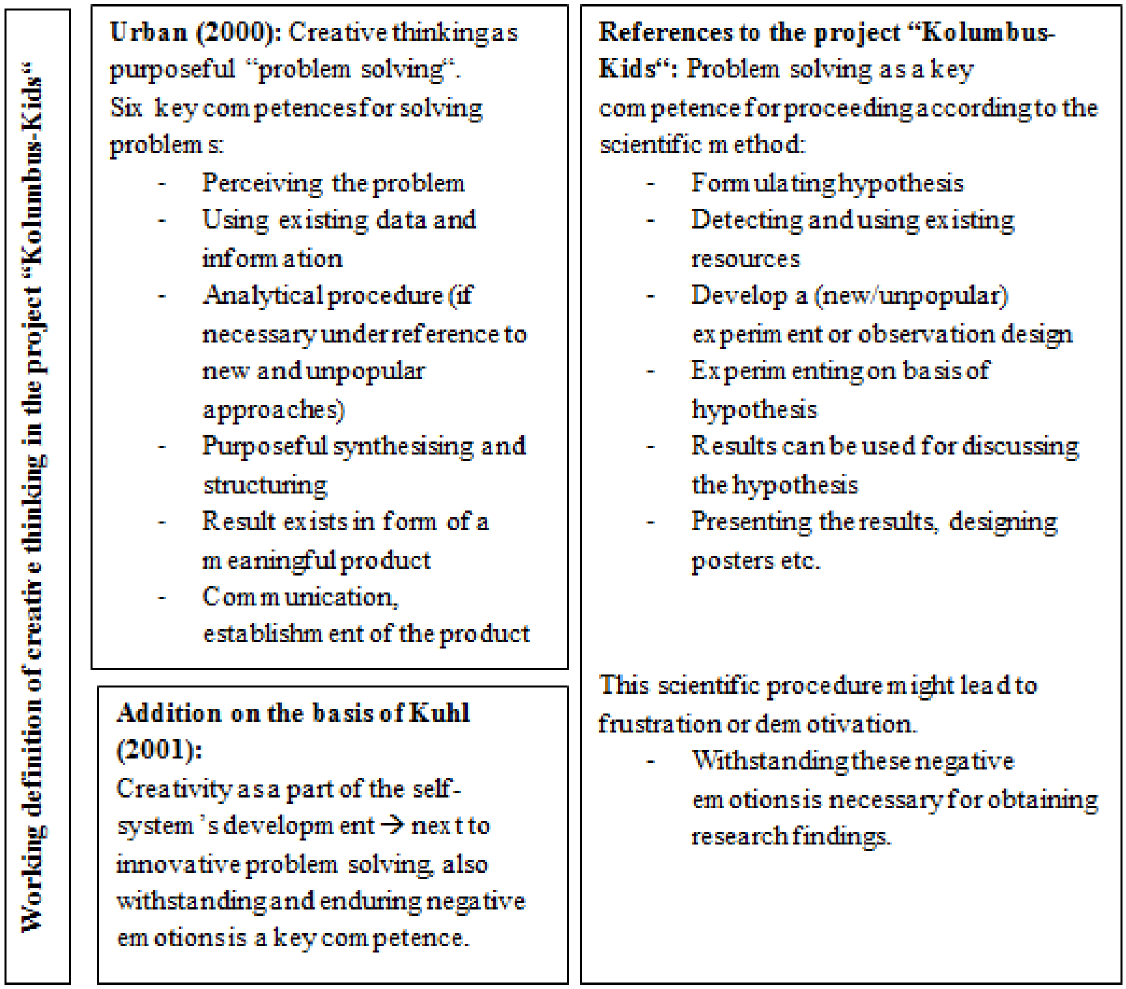


Figure 1: Deducing a definition of creativity

This extracurricular project focuses on teaching units that are oriented towards the scientific method which comprises Urban's aspects of purposeful problem-solving: The perception of problems is a crucial component in teaching concepts since the problem has to be worked out on the basis of observations, information or data and its possible solution then has to be formulated as a hypothesis. Available resources for the problem's solution have to be evaluated and used in a meaningful way so that pupils develop an experimental or observational design as a solution strategy. Even if they might seem unusual at a first glance, approaches that do not lead to an instant problem-solving but still provide important findings are indicators of creative work processes. The result should provide an explanation of the problem and can be presented by the pupils in a short talk or poster, thus disseminating the new insights. As scientific experiments or the observation of animals often do not immediately feature the effects wished for, this can be a frustrating and demotivating work process. As a consequence, pupils have to learn how to deal with these negative emotions in order to arrive at a result nonetheless, as is pointed out by Kuhl.

Creative thinking and intelligence

When Snyderman and Rothman (1987) asked experts to categorise given skills either as part of the construct "intelligence" or as separate constructs, there was a clear tendency to assign elements such as "abstract reasoning", "problem-solving skills" and "knowledge acquisition capacities" to the construct of intelligence (Snyderman & Rothman, 1987, p. 11). The opinions regarding "creative thinking", however, were not as uniform. Slightly more than half of the researchers subsumed creative thinking under intelligence, while the other half considered it a separate construct (Snyderman & Rothman, 1987, p. 12). The connection between intelligence and creative thinking was examined in studies by Sternberg et al. (1996) in the context of their Sternberg Triarchic Abilities Test (STAT). The findings revealed a correlation of the two constructs (Sternberg et al., 1996, p. 121-123). A sample of gifted people had to do a verbal, a figural and a quantitative short test testing the constructs "creative-cognitive intelligence", "analytical intelligence" and "practical-cognitive intelligence" as well as an essay task (Sternberg et al., 1996, p. 121-123). Haensly and Reynolds (1989) obtained similar results, stating that intelligence and creativity correlate throughout the whole range of both variables (Haensly & Reynolds, 1989, p. 249). Since this correlation is of high research interest, a study on creative thinking skills of gifted pupils in the project "Kolumbus-Kids" was conducted at Bielefeld University.

Pedagogical means for the promotion of creative thinking

On a macro level, the establishment of a creativity promoting environment seems hardly realisable, but it is possible to create certain situations that enhance creative thinking skills (Urban, 2000, p. 128). It is this study's purpose to show how children can be supported in this regard from a pedagogic viewpoint, in order to enable employment, possibilities and perspectives for the generations to come. In other words, creativity is a central component in sustainable education and thus should be fostered. Urban (2000) particularly criticises lessons lacking any awareness of problems that inhibit the conceptualisation of creative approaches in schools (Urban, 2000, p. 128). Learners' creative ideas often get lost and do not receive the credit they deserve. The strict distinction of learning, thinking and action processes in 45-minute intervals hampers the development of creative thinking skills, even though schools have the potential to encourage creative thinking (Urban, 2000, p. 129).

The development of creativity depends on the environment (Linneweh, 1973, p. 14). Therefore, schools should make it their responsibility to "create all possible and necessary preconditions for possibilities of creative learning and experience" (Urban, 2000, p. 129). Obviously, teacher personality plays an important role in this regard (Urban, 2000, p. 131-132). A first step is to avoid (un)conscious inhibition of creative thinking. Lessons are often designed in a way that supports common thinking but does not allow inventive and creative behaviour (Urban, 2000, p. 131). Teacher statements such as "Look for the right answer", "Work step by step", "That is not logical!" and "Follow the rules!" may be helpful for arriving at the learning goal, but they inhibit the pupils' ability to think creatively (ibid.). The approach by Urban (2000) comprises eight aspects for the promotion of creativity:

1. Teaching atmosphere. Learning should be anxiety-free and without fear of marks and sanctions. Even pupils' ideas that may seem strange in the first instance should be praised and taken seriously. Every pupil needs to have the feeling that he or she is taken seriously and accepted in the class.
2. The role of the teacher should be to consult, support and motivate pupils rather than evaluating and pressuring them. Teachers have to see themselves as motivators and encourage pupils to express ideas, statements, discoveries and thoughts freely.
3. Learning together. The teacher has to avoid jealousy among pupils and rather create an atmosphere of cooperation. This claim is based on the thesis that group pressure results in conformism, which prohibits creative and innovative thinking.
4. General acceptance of heterogeneity creates a working atmosphere that enables cooperative processes and individual creative thinking at the same time. The lessons will benefit from this acceptance since they yield diverse pupil utterances.
5. Rules for working together should not be prescribed, but developed and agreed on with the pupils. Thus, they work in an environment that they shaped themselves.
6. Lesson division in focused and defocused phases. Positive working stress can be productive for a certain amount of time, but should alternate with relaxed phases. A mixture of different phases makes for a working process that is productive for a longer time period and also supports creative thinking.
7. Creativity-supportive teacher-student-talks. Contributions by pupils that do not help in reaching the lesson goal but are rather of hypothetical nature should not be understood as a disturbing factor but as the individual effort for finding meaning and truth. Teachers should ask questions that require higher level thinking. Leading questions and polar questions are thus to be avoided. In case mistakes are not the result of guessing but of

serious thoughts, they should be regarded as distinct approaches to the problem. Mistakes can potentially represent one step on the way to problem solution.

8. Choice of material. Development of ideas requires diverse and inspiring material. Students should be allowed to choose from a variety of materials since evaluation and purposeful choice can be trained by this.

In the following, general remarks on the project "Kolumbus-Kids" are made in connection to Urban's eight aspects for lesson design. As "Kolumbus-Kids" promotes particularly gifted pupils and is an extracurricular project (Wegner & Ohlberger, 2014, p. 104), marks and examinations can be avoided. In working phases, pupils should develop their ideas with the help of the scientific method (Table 1).

Table 1. The scientific method with a practical example from the "Kolumbus-Kids" project (Wegner et al., 2014)

The scientific method	Lesson example from "Kolumbus-Kids"
1. Identification of the problem	1. A contribution from an internet forum on the topic "Are germs transmitted through coins?" is discussed in the form of a teacher-pupil-talk.
2. Formulation of the problem	2. The problem is analysed in partner work on the basis of informative texts on the lifestyle of bacteria.
3. Formation of hypotheses	3. The pupils form hypotheses and consider ways of checking them.
4. Methods (Planning the solution process)	4. The experiments are planned by the pupils with support from the teacher, taking into account the prescribed material (Petri dish with culture medium, coins, Bunsen burner, ethanol, liquid culture of bacteria).
5. Conduction	5. Conduction of the experiments. The Petri dishes are prepared and incubated for a week.
6. Integration of results (presentation of results)	6. In the following lesson, the pupils record their results, illustrate them and present them to their classmates.
7. Consolidation of results	7. The pupils make notes on their classmates' experiments and results.
8. Reflection of solution process	8. The hypotheses and experiment designs are discussed in teacher-pupil-talk.
9. Practising phase	9. The pupils work on additional tasks on the topic of bacteria.

The scientific method consists of different steps. First, the observation of a phenomenon leads to phrasing a research question, on the basis of which a hypothesis is formed (Barke & Harsch, 2011, p. 104-105). In order to verify or falsify the hypothesis, experiments of observational tasks are generated and conducted afterwards. Any results are recorded and then analysed and ultimately help to check the hypothesis (ibid.). In the project "Kolumbus-Kids" the scientific method is often applied in group processes. The pupils are to notice and formulate problems and design solution strategies. Due to the topics' complexity, cooperative work is necessary so that every pupil can integrate his or her ideas in the process.

Table 1 sketches the application of Urban's claims in a teaching concept that orients itself towards the scientific method. The focus always lies on the pupils and their ideas and suggestions, while the teacher takes a rather defensive and advisory role. He helps the pupils in case questions arise, prepares material or gives new impetus when the working process stagnates. This concept aims at the consideration of all the pupils' ideas, also of those that do not lead to the solution right away. As Urban recommends, every pupil should have the right to be taken seriously with his or her ideas and thoughts. A mix of different phases makes for a productive working process in the project lessons and thus supports creative thinking. The phases are sometimes rather dense, for instance when theoretical backgrounds have to be examined, but very open at other times when pupils can research freely on their own.

Brainstorming and thought experiments explicitly foster the consideration of an issue from different viewpoints and are part of general teacher-pupil-talk. Mistakes and wrong tracks are just as tolerated as innovative ideas. If wrong hypotheses are generated on the basis of misconceptions, these are corrected and modified through experiments or observations. The learning success arising from this is considerably stronger than in prescribed solution process (Barke & Harsch, 2011, p. 104-105). The project "Kolumbus-Kids" therefore offers pupils space for lateral thinking.

The demand for diverse and motivating material is met by supplying the pupils with material that they do not necessarily need for the solution, so that a purposeful decision for certain material is also part of the research process. Thus, the creative process begins with determining the conditions and handling the existing resources. The "Kolumbus-Kids" have to answer the questions "What is the problem?" and "How can I solve it with the help of the information, data and material available?" on their own or in a group.

The aspect of tolerating negative emotional states according to Kuhl (2001) is also considered in "Kolumbus-Kids". Tasks with research character in partner or group work are incorporated in phases that train the pupils' endurance. Sometimes the explorative working phases are long-term and extensive. A short creative phase, however, will lead the participants only partly to a sufficient solution.

As has been shown, there are several elements in the pedagogical concept of the project that support creative thinking. Their success has to be proven by a test that reliably measures creative thinking.

Measuring the construct "creative thinking"

Specific literature often states that creative thinking is hardly measurable (Gardner, 1996, p. 39-40). Reasons for this are its conceptual ambiguity and the various influencing factors and elements (ibid.). Still, Schuler and Hell (2005) developed a test for the analysis of deductive and creative thinking that fulfil the quality criteria of empirical social research and delivers reliable results (Schuler & Hell, 2005, p. 7). The test is based on the relevant factors general intelligence, deductive and creative thinking, with the latter area particularly focussing on divergent thinking processes, or lateral thinking (ibid.). The test covers the abilities to find similarities and connections between issues and to come up with ideas and solutions for problems (ibid.). These competences are surveyed in the module creative thinking by using a written, quantitative test, which is comprised of four groups of tasks: "combining sentences", "generating hypotheses", "interdependent conditions" and "forming categories" (Schuler & Hell, 2005, p. 9). Every group of tasks consists of two to four items (ibid.).

In order to measure creative thinking in the project "Kolumbus-Kids", the ASK-test had to be modified age-appropriately and limited to the task areas "generating hypotheses", "interdependent conditions" and "forming categories", since these task types relate to biology-specific competences (see Table 2). "Combining sentences" is of purely verbal nature. The following table depicts the task types used and relates them to teaching methodology.

Table 2. Description of relevant groups of tasks according to Schuler and Hell (2005) with references to biological teaching methodology

Group of task	**Description**	**Reference to teaching methodology**
Generating hypotheses	For the explanation of given phenomena (statistically proven correlations), as many potential reasons as possible have to be hypothesised (Schuler & Hell, 2005, p. 10).	Inductive approaches in terms of the scientific method include the generation of hypotheses as a key competence (Gentürk, 2009, p. 3).
Interdependent conditions	As many factors as possible influencing the given phenomenon are to be found, including their interdependency (Schuler & Hell, 2005, p. 10).	Recognising impacting factors and interactions is part of the concept 'governance and regulation' and important in the field of ecology (biotic/ abiotic factors) (Kremer, 2012, p. 5).
Forming categories	Different terms have to be put into meaningful categories. In doing so, as many categories as possible should be created (Schuler & Hell, 2005, p. 10).	The basic concept 'structure and function' describes the organisation of biological objects according to structural and functional commonalities. In biology, especially in classification, this ability belongs to the key competences (Kremer, 2012, p. 5).

The table shows that the types of tasks measuring creative thinking also test competences needed in the field of biology. This double-characteristic of the ASK-test makes it an ideal instrument to be used in a biology education project.

Research question and hypotheses

The study aims at the efficient measurement of creative thinking skills. Data collection takes place at two points of time, namely before (T1) and after (T2) a course term (half a year), in order to assess potential changes in creative thinking. As a measuring instrument, the ASK-Test by Schuler and Hell (2005) is employed. The test measures the competences deductive and creative thinking; however, the project uses a slightly modified version of the test. The two test modules deductive and creative thinking can be applied together or separately (Schuler & Hell, 2005, p. 9). As has been mentioned, the verbal task type "combining sentences" was excluded. The three task

types “generating hypotheses”, “interdependent conditions” and “forming categories” were used in their original form. However, the task difficulty had to be adapted.
Since the project employs measures to promote creative thinking, the following hypotheses are examined.

H1: The primary school pupils significantly improve in the domain ‘creative thinking’ throughout a “Kolumbus-Kids” course.

H2: The secondary school pupils significantly improve in the domain ‘creative thinking’ throughout a “Kolumbus-Kids” course.

The test comprised three task types that measure the construct of creative thinking. Since all three task types contain ways of thinking relevant in the field of biology, it seems reasonable to assume a significant improvement of the pupils in the individual task areas. The lessons are designed with regard to the scientific method, which is why pupils are potentially trained in the generation of hypotheses. The following hypotheses are based on these considerations:

H3: The primary school pupils significantly improve in the task area “generating hypotheses” throughout a “Kolumbus-Kids” course.

H4: The secondary school pupils significantly improve in the task area “generating hypotheses” throughout a “Kolumbus-Kids” course.

Since the detection of impacting factors and interactions is a central component in biology education, it seems reasonable to assume that the “Kolumbus-Kids” also significantly improve in the task area of “interdependent conditions”.

H5: The primary school pupils significantly improve in the task area “interdependent conditions” throughout a “Kolumbus-Kids” course.

H6: The secondary school pupils significantly improve in the task area “interdependent conditions” throughout a “Kolumbus-Kids” course.

"Kolumbus-Kids" lessons often focus on animals. Thereby, the children also learn about aspects of classification, which can be assumed to foster the competence of assigning structural or functional commonalities. Thus, a significant improvement in the task area "forming categories" can be expected:

H7: The primary school pupils significantly improve in the task area "forming categories" throughout a "Kolumbus-Kids" course.

H8: The secondary school pupils significantly improve in the task area "forming categories" throughout a "Kolumbus-Kids" course.

An age-appropriate test for pupils from grade four to six (aged 9-12) was developed for this study, which is why a higher score can be expected for the secondary school pupils (grade 5/6) when compared with the primary school pupils (grade 4).

H9: The secondary school pupils will achieve a significantly higher test score in the measurement when compared to primary school pupils.

Due to the purposeful promotion of the participants in the "Kolumbus-Kids" courses, assuming an improvement in both courses seems feasible (see H1 and H2). This leads one to ask what kind of influence the promotion of the pupils' abilities has when comparing age groups.

H10: Primary school pupils do not differ significantly from the secondary school pupils in the second measurement.

A prospective research focus in the project will be placed on the aspect of gender. The promotion of girls' scientific abilities will be focused on, since women are still underrepresented in scientific university courses and jobs (Kröll, 2010, p. 11). Due to the general interest in issues of gender, this study also examines possible differences in the test results of girls and boys. No evidence for gender-specific differences in creative thinking can be found in the literature. However, approaches for the promotion of girls in STEM-subjects similar to "Kolumbus-Kids" exist. Ziegler et al. (2012) found that the gender differences in the STEM-subjects are often accounted for by disinterest (Ziegler et al., 2012, p. 233). Certain promotion programmes show an empirical increase in interest, which, however, does not persist, as the environment counteracts as a neutralising factor (ibid.). Projects supporting STEM-learning can contribute to the promotion of gender equality in this area (Ziegler et al., 2012, p. 235-236). “Kolumbus-Kids” is such a learning sociotope. It thus seems reasonable to assume that girls’ regular participation in the project lessons leads to a higher interest in STEM-subjects. Accordingly, an increase in the girls’ competence might also be detectable in the measurements of creative thinking. Since these assumptions are purely hypothetical, measurable effects are not expected.

H11: Girls and boys do not achieve significantly different results in either measurement in the domain ‘creative thinking’.

Conception of study and data analysis

Method

The paper and pencil test was conducted according to the standardised guidelines of Schuler and Hell. Thus, necessary parameters such as a quiet environment, good lighting conditions and a sufficiently sized workplace are required (Schuler & Hell, 2005, p. 12). The researcher conducting the test has to follow a standardised procedure. He or she should be acquainted with the test in order to be able to answer potential questions (ibid.). An exact processing time for each part of the test was prescribed, adding up to a total test time of 30.5 minutes. The test for measuring creative thinking was conducted according to the standards, which is why the standardised analysis also applies.

At the beginning and end of every “Kolumbus-Kids” course throughout winter 2013/2014, this test was implemented with a sample of 35 primary school pupils and 50 secondary school pupils. The main quality criteria will not be discussed in detail, but the standardised procedure and the studies on reliability and analyses by Schuler and Hell have been shown to fulfil the main quality criteria of didactic research and the results can therefore be used to make statements about the creative thinking skills of pupils and their development over a school term.

The test

As has been stated, the evaluation instrument is an adapted version of the ASK-test by Schuler and Hell (2005). The test consists of three task areas that are described in detail by means of presenting the task formulation and individual items.

The task area "generating hypotheses"

The task formulation of Schuler and Hell's test booklet (2005a) was used with only a slight adaptation in the greeting, since the participants in this study are addressed informally due to their age (applies to all task areas). The first task is:

> "In the following, you will be presented a number of phenomena. You have to generate hypotheses on the cause of these issues. Formulate as many reasonable hypotheses as possible." (Schuler & Hell, 2005a, p. 6).

The instruction says that only reasonable solutions will be graded. Still, the instruction time for this task is 2.5 minutes instead of the pre-set 1 minute because not all pupils know exactly what a hypothesis is. Therefore, examples for hypotheses and non-hypotheses are discussed.

The task formulation, the examples and instructions are listed in the test booklet and thus visible for every participant. The three issues for the first task are:

Phenomenon 1: In the animal world, males often have a magnificent plumage, shining fur or colourful body parts, while the females are rather ordinary in their appearance.

Phenomenon 2: Some plants have colourful and showy blossoms, whereas others have ordinary and hardly noticeable blossoms.

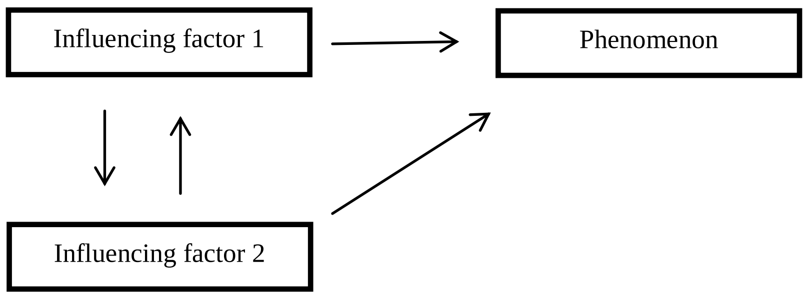
Phenomenon 3: Scientists predict a climate change for Central Europe which results in hotter summers and colder winters.

The task area "interdependent conditions"

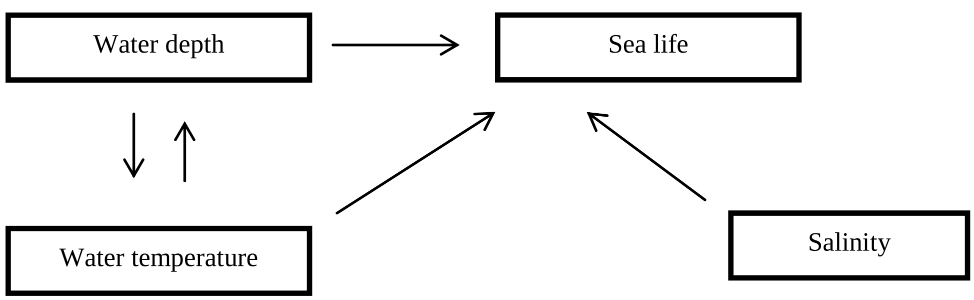
In this task area, the pupils have to allocate influencing factors to a certain situation and define their relations. The exact, quite comprehensive task formulation is:

> "Many biological issues are influenced by various factors or are at least influenceable by them. In this task you have to find various biological influencing factors and outline their impact. Please follow these rules:

> 1. Mark all influencing factors by frames
>
> 
>
>
> 2. Outline the influencing factors' impact on the phenomenon by using arrows. You can also use effect-arrows between the influencing factors."



This task was also taken from Schuler and Hell (2005a). In the test booklet, it is also focussed on the terms "phenomenon" and "influencing factor". The following example for the situation "sea life" is provided:



The items in the test were on the topics "desert life" and "nutrition".

The task area "forming categories"

For this task, the pupils are asked to group terms according to their meaning. The categories also have to be titled meaningfully. The exact task is (Schuler & Hell, 2005a):

> "In the following, there are groups of terms. You have to categorise the terms in a meaningful way and find as many categories as possible (see example). A category can just be

regarded as such if it is made up of at least two terms. They have to be found on the basis of content criteria and not due to formal aspects (e.g. all the words with three letters). Numbered terms are cited in the header of the following pages. If you find a category, give it a meaningful title and put it into the field "name of the category". (Schuler & Hell, 2005a).

For this type of task, the following example is provided:

Given terms: 1. Cat, 2. Mouse, 3. Earthworm, 4. woodlouse, 5. Water flea, 6. water bug

Category: mammals	**Representatives**: 1, 2
Category: invertebrates	**Representatives**: 3, 4, 5, 6
Category: isopoda	**Representatives**: 4, 6

There are 25 terms per task for the formation of categories in the test booklet.

Evaluation of the tests

The evaluation of the test is conducted in a standardised way as well, the first step being the assessment of the tasks through scores. In a second step, the scores are transferred into standardised values. Thirdly, the data is statistically processed using SPSS.

Assessment of the individual tasks

Schuler and Hell prepared concise guidelines on how the task types "generating hypotheses", "interdependent conditions" and "forming categories" are to be assessed. These are presented briefly.

Assessment of task area "generating hypotheses"

Only solutions that comply with the task formulation are assessed. Correctly phrased hypotheses yield one point, whereas senseless hypotheses are not taken into consideration (Schuler & Hell, 2005, p. 15). The hypotheses need to show an understandable reference to the phenomenon described and have to represent possible reasons or conditions for the situation (Schuler & Hell, 2005, p. 14). Different hypotheses that involve the same content are only counted once (ibid.). If two partial hypotheses are made from one hypothesis but are equivalent content-wise, only one hypothesis is counted (ibid.). For better understanding, see the following example: the issue "The birth rate in France has been considerably higher than in Germany for many years" might lead to the hypotheses "The Germans do not spend so much time with their partners" and "The French spend much time with their partners", which would then yield one point only (ibid.). If the partial hypotheses are however selective, they are assessed with one point each (Schuler and Hell, 2005, p. 14-15). Beyond that, objectively wrong hypotheses do not count (ibid.) Orthographic mistakes are not taken into account (Schuler and Hell, 2005, p. 15).

Assessment of task area "interdependent conditions"

Again, only solutions that comply with the assignment are counted as correct. The influencing factors defined by the participants need to have a meaningful relation to the issue. Consequences of the situation are not counted as valid answers. Each influencing factor mentioned scores one point. Orthographic mistakes are ignored.

Assessment of task area "Forming categories"

Only those solutions that suit the task formulation are considered in the assessment. A category has to be of content-related nature, thus categories such as "words with four letters" are not counted. The labelling of categories has to be meaningful. One term may be mentioned in more than one category, but a category itself is only acknowledged once it comprises at least two terms. Every category formed according to the guidelines scores one point. Orthographic mistakes are ignored.

Working with raw data

The pupils' scores for the seven items are considered the raw data in the standardised procedure according to Schuler and Hell. As a result, every pupil is ascribed a standardised facet score per test, including the corresponding percentile rank for creative thinking.

The first step of the procedure is the addition of the raw scores for the individual task types "generating hypotheses" (GH), "interdependent conditions" (IC) and "forming categories" (FC), giving the raw data sum (RDS) (Schuler & Hell, 2005a, p. 15). The raw data sums "GHRDS", "ICRDS" and "FCRDS" are standardised in a second step. Since the task type "generating hypotheses" consisted of only three instead of four items, the GHRDS value is multiplied by 4/3.

This is done assuming that the test pupil would have reached the same score in the fourth item as in the other three items.

For further calculations with the raw data sums, the GHRDS, ICRDS and FCRDS values are allocated to standardised raw data sums which can be taken from the respective tables (Schuler & Hell, 2005a, p. 51). Adding the standardised raw data sums yields non-standardised facet scores (Schuler & Hell, 2005a, p. 15-16), which are once again multiplied by 4/3 in order to project the data to a fourth task type that was not used in this study. Only by doing so, the assessment strategy by Schuler and Hell is applicable. By means of standards table, the non-standardised facet scores are matched to standardised facet score and the according percentile ranks (Schuler & Hell, 2005a, p. 53). The method of the assessment procedure is illustrated in Figure 2.

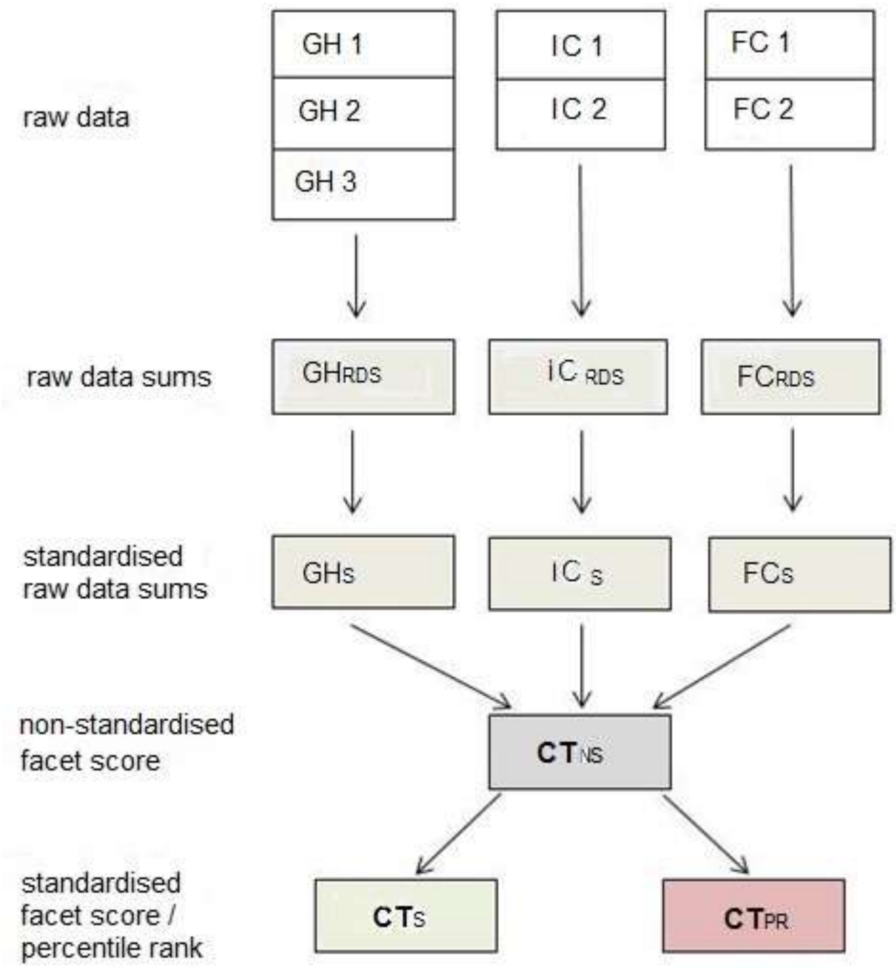


Figure 2. The assessment procedure based on Schuler and Hell (2005) for the Creative Thinking part of the ASK-Test

Statistical analysis of the test results

For the review of the previously stated hypotheses, the correlations among the test results have to be assessed using SPSS. Using the general linear model, the potential change of the results throughout repeated measurements is tested. A two-part test with the parameters time and age-group level, and time and gender, respectively is conducted. It is also tested whether there are significant differences between the age-group levels. The pupil's development in the single task areas is assessed in a three-stage test. The results will be presented in the following.

Results

First, the results of the reliability analysis will be outlined, before the results yielded on the basis of the general linear model are elaborated upon.

Results of the reliability analysis

The reliability analysis of the results at the first point of measurement with the data of all pupils (N = 78) was conducted using SPSS. In order to provide an overview, the results are presented in tabular form:

Table 3. Results of the reliability analysis (N = 78).

		Cronbach's α
Task areas	Generating hypotheses (GH)	0.747
	Interdependent conditions (IC)	0.605
	Forming categories (FC)	0.735
Total value Creative Thinking (CT)		0.728

As can be inferred from the table, the tasks "generating hypotheses" and "forming categories" achieved values of $\alpha \geq .70$, meaning these areas can be considered reliable (George et al., 2002, p. 231). The task area "interdependent conditions" yielded a value just under .70, making it an acceptable result nonetheless (ibid.). For the total value of creative thinking, Cronbach's α is clearly above .70.

Results of the calculations using the general linear model

Two-stage and three-stage analyses were conducted when checking the hypotheses using the general linear model. First, the data set was analysed regarding the development of pupils over time. For that purpose, the pupils' results from measurement point T1 (October 2013) and T2 (February 2014) are included as within-subject factors for the standardised total value CTS. 75

pupils took part in the tests at both T1 and T2. The age group allocations were considered between-subject factors.

When taking the test for the first time in October, the primary school pupils (N = 34) reached a mean score of 75.85 points with SD± 4.251, whereas pupils of the secondary schools (N = 41) gained 81.34 points with SD± 5.708. The average of all the pupils is 78.85 points (SD± 5.765). For the second test, both groups of pupils achieved higher mean scores. The primary school pupils' performance amounted to 87.12 points (SD± 3.453) and the secondary school pupils achieved 94.37 points on average (SD± 4.515). All results of the second test taken together amount to an average score of 91.08 points (SD± 5.435).

The increase of the standardised values for creative thinking at measurement points T1 and T2 is significant, $F(1.73) = 392.853$; $p<0.01$; $Eta^2 =.837$. Testing the between-subject effects confirms the difference between primary and secondary school pupils. The difference between pupils of year four and year six is thus statistically significant, $F(1.73) = 52.665$; $p<0.01$, $Eta^2 =.419$. Figure 3 is a visual depiction of the measured effects.

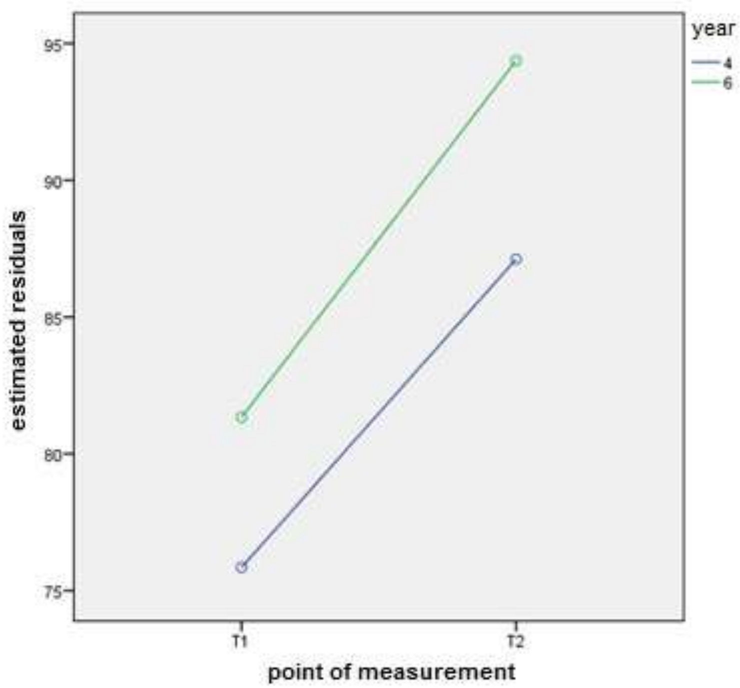


Figure 3. The development of pupil results (CTs total values) on the y-axis at the points of measurement T1 and T2 (x-axis). N = 75

In order to interpret the results of the study properly, the same analysis was conducted in a three-stage form with a scale on the standardised values of the single task areas. The results in a tabular overview (Table 4).

Table 4. Descriptive statistics of the three-stage analysis with the within-subject effects GHs, ICs and FCs, as well as the between-subject effects year 4 and 6. N = 75.

	year	mean	standard deviation	N
GHs_T1	4	79.6176	4.49252	34
	6	85.5854	5.81367	41
	sum total	82.8800	6.01790	75
ICs_T1	4	79.4706	3.79182	34
	6	83.3415	6.01918	41
	sum total	81.5867	5.45521	75
FCs_T1	4	87.5000	5.65284	34
	6	90.1463	6.57861	41
	sum total	88.9467	6.27714	75
GHs_T2	4	91.3529	4.66374	34
	6	96.0488	6.12353	41
	sum total	93.9200	5.95878	75
ICs_T2	4	85.4706	3.62847	34
	6	91.4390	5.21080	41
	sum total	88.7333	5.43098	75
FCs_T2	4	94.9706	3.75367	34
	6	99.9756	6.05181	41
	sum total	97.7067	5.68966	75

The top three rows of the table illustrate the results of the individual task types at T1. The lower three rows show the results at T2 respectively. Both groups of pupils achieved more points in the second test than in the first. A distinct difference between the results of the year 4 and year 6 pupils can be observed. The difference between the two points of measurement for the scale figures GHs, ICs and FCs is statistically significant for all cases, $F(2.146) = 7.174$; $p<0.01$; $Eta^2 = .089$. The differences between the age groups was statistically significant, too, $F(1.73) = 52.674$; $p<0.01$; $Eta^2 = .419$.

As table 4 shows, pupils achieved the highest scores in the area “forming categories”. At the beginning of the course, the primary school pupils gained an average of 87.5 points, whereas the sixth graders gained 90.15 points. At the end of the course, however, both groups improved significantly and achieved 94.8 and 99.99 points respectively.

For each group, “generating hypotheses” was the second most successful task area. The primary school pupils scored 79.62 points and the secondary school pupils achieved 85.59 points on average. A particularly distinct increase in points can be observed for the primary school pupils, who reached 91.36 points at T2. The secondary school pupils reached 96.05 points. The task area “interdependent conditions” came in last. Both groups achieved the lowest scores at both points of measurement. Based on the low original scores of 79.47 (primary school) and 83.35 (secondary school), the increase to 85.47 and 91.44 points respectively is significant. Also, a two-way analysis with the between-subject factor gender was carried out, with the within-subject factors being the standardised total values of CTs_T1 and CTs_T2. At one point of measurement, the participating number of pupils was 44 boys and 31 girls of both age groups. At T1, the boys scored 77.8 points, a slightly lower score than that of the girls, who achieved 80.35 points (in both cases, SD ± 5.6). When comparing gender, however, the effect is not as clear-cut. At T2, the boys achieved an average result of 90.68 points (SD ± 6.057), while the girls reached 91.08 points (SD ± 4.439). For both groups, the differences between T1 and T2 are statistically significant, $F(1.73) = 378.762$; $p<0.01$; $Eta^2=.838$. Testing the between-subject effects (male/female) did not show any significance, $F(1.73) = 2.364$; $p>0.01$; $Eta^2=0.031$. Therefore, it can be stated that there is a numeric difference between boys and girls concerning the standardised test results for creative thinking, but it is not significant, leading to the conclusion that boys and girls participating in “Kolumbus-Kids” did not differ regarding their creative thinking skills. The interpretation of the results and their integration in the context of the hypotheses follows in the discussion section.

Discussion

The interpretation of the results is done on two levels. On the one hand, the efficiency of the instrument is discussed, on the other hand the hypotheses is verified or falsified on the basis of the general linear model calculations.

Discussion of the test

Using the test by Schuler and Hell (2005) in an adapted version is a novelty and thus requires discussion. The pupils’ results show that the internal consistency of the measurement instrument is maintained, at least for the standardised total value of the construct “creative thinking” and the task areas of “generating hypotheses” and “forming categories”. For the task type “interdependent conditions”, Cronbach’s alpha was slightly below 0.7, which makes a more detailed discussion of this task type necessary. The pupils had to develop a structure of conditions regarding the topics “desert life” and “nutrition”. These two given phenomena were deliberately chosen with an obvious discrepancy of the topics so that pupils would be able to do the task irrespective of their knowledge in these areas. A child that has practically no previous knowledge of the habitat desert could not write as much as one that is highly interested in the topic. In a case like that the item would measure interest or lacking knowledge rather than the actual creative potential of a child. To reduce the dependence of the children on previous knowledge and in order to give them a chance to compensate in case of a score deficit, issues from two completely different areas (ecology and human biology) were chosen. This difference can affect the alpha coefficient since it can make for a greater disparity of the item scores.

Since the test was adjusted to a level of difficulty appropriate for year 4 and 6 pupils, one has to consider whether the pupils were possibly over- or under-challenged. The significant difference that is apparent in the results of year 4 and year 6 pupils at both points of measurement reveal that they were not under-challenged by the test. If the test's demands had been too low, such a difference between younger and older pupils would have been unlikely. In general, one would expect the results to correlate with results of other research instruments that were used together with the ASK-Test. In order to standardise the test, Schuler and Hell compared the results of the ASK-Test with the IQ of the test persons, which were students of social and economic sciences (Schuler & Hell, 2005, p. 16-17). The results of the "Kolumbus-Kids" test subjects exhibit a standard deviation that is distinctly lower at both measurement points. The SD found are 3.5 to 5.5 respectively with the test's significance level being 5%, while Schuler and Hell calculated a SD of 12.4 (Schuler & Hell, 2005, p. 18). This result may stem from the higher than usual scientific talent of the "Kolumbus-Kids", leading to a rather homogenous performance in the group. Schuler and Hell categorised the creative thinking total values into three levels: below-average test result (less than 93 points), average test result (93 – 107 points) and above-average test result (more than 107 points) (ibid.). Principally, this grading also has to apply to the adjusted test, since standardised evaluation and alike were considered. On that basis, the pupils in the project "Kolumbus-Kids" would have scored below-average for the first testing, which seems rather questionable since the pupils were shown to be gifted using a diagnostic test. The low test results might be explained by the fact that the test was too demanding for the pupils or just not suited for children of that age. Due to the natural scientific focus of the adjusted test, the instrument did not take the verbal part "combining sentences" of the ASK-Test into account. This, however, might have impacted the test's construct validity. Still, the natural scientific emphasis is reasonable since the pupils taking the test are considered scientifically gifted. Including the verbal section of the test would only make sense if the test group was heterogeneous with regard to special talents, as certainly was the case with Schuler and Hell's sample.

Looking at the results of the three-stage analysis, it becomes apparent that despite standardisation of the task groups, the pupils still handled the tasks differently. At both points of measurement, the participants achieved the highest scores in the task area "forming categories". As illustrated before, these kinds of tasks require the pupils to have good abilities of the basic concept structure and function, which are dealt with quite early in school. At the end of year 6, a pupil is supposed to have acquired the following competence: "The concepts of connections between structure and function are developed to the extent that simple relations on a phenomenological level can be outlined" (Kernlehrplan des Landes Nordrhein-Westfalen, 2008, p. 27). Thus, it is not surprising that the pupils were good at categorizing given terms according to structure or function. The project "Kolumbus-Kids" also promotes this basic concept so that the significant increase in the pupils' performance confirms expectations. The task area "generating hypotheses" came in second. Interestingly, particularly the year 4 pupils improved in this task. The significant difference of both groups at T1 and T2 was to be expected since the scientific method is the fundamental teaching concept in "Kolumbus-Kids". Lessons in the project are based on hypotheses that are formulated by the pupils, a competence which is trained in different contexts. The task area "interdependent conditions" also saw a significant increase in scores between T1 and T2. However, the pupils achieved the lowest scores in this task area, which might be explained by its complexity. The pupils did not only have to define factors of influence, but also interconnect them. The combination of defining and relating seems to have posed a problem for the children between 9 and 12 years. Still, "Kolumbus-Kids" promotes these combined competences since a positive development over the course of the project is detectable.

Discussion of the hypotheses

The hypotheses stated at the beginning relate to the development of creative thinking in pupils taking part in "Kolumbus-Kids" for half a year. Another aim was to find out whether there are differences between fourth and six graders and between boys and girls.

The first two hypotheses are:

H1: The primary school pupils significantly improve in the domain 'creative thinking' throughout a "Kolumbus-Kids" course.

H2: The secondary school pupils significantly improve in the domain 'creative thinking' throughout a "Kolumbus-Kids" course.

On the basis of the results, both hypotheses can be verified. This applies both to the general construct as well as to the different task areas. As the project “Kolumbus-Kids” conforms with Urban’s (2000) and Kuhl’s (2001) proposals for creativity-promoting lessons, the results are as expected. Interventions similar to “Kolumbus-Kids” have proven to be comparably effective. Bastian (2000), for instance, conducted a study based on a music-pedagogy and found that 6 to 12 year old children scored significantly better in creativity tests after the intervention (Bastian, 2000, p. 101-102).

Hypotheses H3 – H8 are concerned with the individual task areas. H3 and H4 on the development in the task area “generating hypotheses” are discussed first:

H3: The primary school pupils significantly improve in the task area “generating hypotheses” throughout a “Kolumbus-Kids” course.

H4: The secondary school pupils significantly improve in the task area “generating hypotheses” throughout a “Kolumbus-Kids” course.

The results make for a clear verification of both hypotheses. Again, the educational concept of the project explains the increase in competence. The lessons are geared to the scientific method, which includes the formulation of hypotheses regarding real phenomena.

Next, the focus is on the hypotheses for the task area “interdependent conditions”:

H5: The primary school pupils significantly improve in the task area “interdependent conditions” throughout a “Kolumbus-Kids” course.

H6: The secondary school pupils significantly improve in the task area “interdependent conditions” throughout a “Kolumbus-Kids” course.

H5 and H6 are also verified by the results. Both groups were able to improve significantly during participation in the project for one course term. Various biological contexts are established on the basis of the scientific method, which is why pupils certainly have improved their abilities to detect influencing factors and interactions. “Kolumbus-Kids” therefore promotes the capacity of defining and interconnecting several phenomena and influential factors.

Hypotheses H7 and H8 on the task area “forming categories” is discussed in the following:

H7: The primary school pupils significantly improve in the task area “forming categories” throughout a “Kolumbus-Kids” course.

H8: The secondary school pupils significantly improve in the task area "forming categories" throughout a "Kolumbus-Kids" course.

The results seem to verify H7 and H8. Both groups were able to improve significantly in the task area "forming categories". In the "Kolumbus-Kids" lessons, the competence of organising structural or functional similarities is fostered by activities such as having the children focus on reptiles and learning about different species, for example. That might explain the improvement in this task area.

Hypotheses H9 and H10 revolved around the difference between fourth and six graders. The hypotheses were:

H9: The secondary school pupils will achieve a significantly higher test score in the measurement when compared to primary school pupils.

H10: Primary school pupils do not differ significantly from the secondary school pupils in the second measurement.

The main purpose of H9 and H10 was to calibrate the test's level of difficulty. Since there was a significant difference at both points of measurement, it can be assumed not to have under-challenged any pupil. Moreover, the authors predicted that the primary school pupils would achieve basic competences due to assistance measures, which would compensate for the difference. Because of this, H10 was formulated reversely. As both groups improved significantly, it is understandable that there still remains a significant difference between both age groups at the second point of measurement. H9 therefore proves right, whereas H10 cannot be verified, meaning that "Kolumbus-Kids" promotes the abilities of all pupils equally. Looking at the absolute numbers, it even seems as if the six graders benefit more from the measures. This tendency, however, cannot be substantiated statistically. Yet, it can be said that both groups show clear possibilities for improvement. H11 was concerned with the aspect of gender:

H11: Girls and boys do not achieve significantly different results in either measurement in the domain 'creative thinking'.

The results are decisive. Regarding creative thinking, there are no huge differences between girls and boys. Both groups were able to improve significantly. The absolute scores even suggest a tendency in favour of the girls, who, on average, scored slightly higher, which should be investigated further in future studies. A seemingly necessary need of promoting girls' abilities in STEM-subjects thus cannot be deduced from the results, even though one has to consider that creative thinking is not the only component of success in STEM-subjects. Since there was no greater increase on the part of the girls detectable when compared to the boys, one can say that "Kolumbus-Kids" promotes both gender's competences equally. Since the project is coeducational, this is not surprising, even though "Kolumbus-Kids" might be considered a learning sociotope, which, according to Ziegler et al. (2012), supports competence acquisition of girls.

Conclusions

It is rarely the case that research interests correspond with the existing measurement instruments (Hunziker, 2008, p. 13-14), which also applies to this study. An instrument for collecting data on creative thinking skills in pupils of grades four and six, with a particular scientific focus, was not available. This is why Schuler and Hell's (2005) ASK-Test was used as a template for this study and slightly adapted to the needs of the project "Kolumbus-Kids".

This study aimed at analysing to what extent the project lessons influenced the children's ability to think creatively. Indeed, a significant increase could be measured over a period of half a year. This is true for both primary and secondary school pupils. Thus, it can be stated that the project "Kolumbus-Kids" seems to promote creativity. The score differences between boys and girls were not significant, which makes a special support programme aiming at fostering the creative thinking skills of girls in scientific contexts seem unnecessary.

In addition to the insights gained by interpreting the test results, this study also led to the development of a measurement instrument of great potential. In order to further evaluate the positive impact of the project "Kolumbus-Kids", the test will be used for a third time in the project courses, thus expanding the current study by T3. This is done in order to find out whether the pupils improve even further or whether they did reach a plateau after a certain time in the course. Another aspect that merits further investigation is the tendency that secondary school pupils' performance increased more than the performance of the primary school pupils. A follow-up study could try to replicate this result in order to determine whether the participants' age affects their ability to benefit from the course or whether the difference was simply due to inaccurate measurement.

Nevertheless, the test and the project's results contribute to the general research concerning the relation between intelligence and creativity. In future studies, the test could be used in performance-heterogeneous groups, for instance in schools, and the results could be compared to the outcomes observed in the project studied here. This might shed light on the general connection between intelligence and creative thinking.

References

Barke, H.-D. & Harsch, G. (2011). *Chemiedidaktik kompakt. Lernprozesse in Theorie und Praxis*. Berlin: Springer-Verlag.

Bastian, H. G. (2000). *Kinder optimal fördern - mit Musik. Ergebnisse einer sechsjährigen Langzeitstudie über Wirkungen von Musik und Musizieren auf die Entwicklung 6 – 12 Jähriger*. Mainz: Schott-Verlag.

Gardner, H. (1996). *So genial wie Einstein. Schlüssel zum kreativen Denken*. Stuttgart: Klett-Cotta-Verlag.

Gentürk, E. (2009). *Experimente im Biologieunterricht. Definition, Bedeutung, Kompetenzerwerb*. München: Grin-Verlag.

George, D. & Mallery, P. (2002). *SPSS for Windows step by step. A simple guide and reference, 11.0 update*. 3rd edition. Boston: Allyn and Bacon.

Haensly, P. A, & Reynolds, C. R. (1989). Creativity and intelligence. In J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), *Handbook of creativity* (pp. 112-132). New York: Plenum Press.

Hoyos, C. G. (2005). Wirtschaftsethik. In D. Frey, C. G. Hoyos, & L. von Rosenstiel, Lutz (Eds.), *Wirtschaftspsychologie* (pp. 407-414). Weinheim: Beltz-Verlag.

Hunziker, A. W. (2008). *Spaß am wissenschaftlichen Arbeiten: So schreiben Sie eine gute Semester-, Bachelor- oder Masterarbeit.* 3rd edition. Zürich: SKV-Verlag.

Kernlehrplan des Landes Nordrhein-Westfalen (2008). *Kernlehrpläne für das Gymnasium Biologie*. Retrieved from www.standardsicherung.schulministerium.nrw.de/lehrplaene/upload/lehrplaene_download/gymnasium_g8/gym8_biologie.pdf.

Koppel, O. & Plünnecke, A. (2009). *Fachkräftemangel in Deutschland. Bildungsökonomische Analyse, politische Handlungsempfehlungen, Wachstums- und Fiskaleffekte.* Köln: Dt. Inst.-Verl. (IW-Analysen, Nr. 46).

Kremer, M. (2012). *Grundbildung in den naturwissenschaftlichen Fächern - Basiskompetenzen. Denkanstöße und Empfehlungen zum Erwerb dauerhaft verfügbarer Kompetenzen bis zum mittleren Schulabschluss im Gymnasium und vergleichbaren Schulformen.* Neuss: Seeberger.

Kröll, D. (2010). "Gender und MINT". Schlussfolgerungen für Unterricht, Beruf und Studium: *Tagungsband zum Fachtag am 15.02.2010*, Universität Kassel and Max-Eyth-Schule Kassel.

Kuhl, J. (2001). *Motivation und Persönlichkeit. Interaktionen psychischer Systeme*. Göttingen: Hogrefe.

Linneweh, K. (1973). *Kreatives Denken. Techniken und Organisation innovativer Prozesse; kreative Denkprozesse, innovative Techniken, kreative Organisation, kreative Werbung, innovative Planung*. Karlsruhe: Nadolski-Verlag.

Orloff, M. A. (2005). *Grundlagen der klassischen TRIZ. Ein praktisches Lehrbuch des erfinderischen Denkens für Ingenieure*. 2nd edition. Berlin, Heidelberg: Springer.

Rohr, A. R., & Rohr-Dietschi, U. (1975). *Kreative Prozesse und Methoden der Problemlösung*. Weinheim, Basel: Beltz.

Schuler, H., & Hell, B. (2005a). *ASK. Testheft zum Modul Kreatives Denken (KD).* Bern: Hans Huber Verlag.

Schuler, H., & Hell, B. (2005). *ASK. Analyse des Schlussfolgernden und Kreativen Denkens*. Bern: Hans Huber Verlag.

Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist, 42,* 137–144, doi:10.1037/0003-066x.42.2.137

Ständige Konferenz der Kultusminister der Länder (2013). Vorgaben für die Klassenbildung. Schuljahr 2013/2014. Retrieved from http://www.kmk.org/fileadmin/pdf/Statistik/Dokumentationen/Klassenbildung_2013.pdf.

Sternberg, R. J., Ferrari, M., Clinkenbeard, P., & Grigorenko, E. L. (1996). Identification, instruction, and assessment of gifted children: a construct validation of a triarchic model. *Gifted Child Quarterly, 40*, 129–137.

Urban, K. K. (2000). Kreativität: vom Störfaktor zum Unterrichtsziel. In H. Wagner (Ed.), *Begabung und Leistung in der Schule. Modelle der Begabtenförderung in Theorie und Praxis* (pp. 117-138). 2nd edition. Bad Honnef: Bock.

Urban, K. K. (2004). *Kreativität. Herausforderung für Schule, Wissenschaft und Gesellschaft*. Münster: Lit-Verlag.

Von Wissel, C. (2012). *Wissenschaftliche Kreativität*. Ed. Hans Böckler Stiftung. Düsseldorf: Setzkasten GmbH.

Wegner, C. (2008). *Entwicklung und Evaluation des Projekts "Kolumbus-Kids" zur Förderung begabter SchülerInnen in den Naturwissenschaften.* Dissertation. Universität Bielefeld, Bielefeld.

Wegner, C., & Fischer, O. S. (2014). Schwermetallresistente Bakterien auf Euromünzen. *Unterricht Biologie*, Issue 391, 30-31.

Wegner, C., & Ohlberger, S. (2014). Giftedness and the project "Kolumbus-Kids": an example of putting theory in to practice. *Journal of Science Education, 15*(2), 104–106.

Wentura, D., & Frings, C. (2013). *Kognitive Psychologie*. Wiesbaden: Springer.

Ziegler, A., Stöger, H., & Heilemann, M. (2012). *Mädchen und Frauen in MINT. Bedingungen von Geschlechtsunterschieden und Interventionsmöglichkeiten*. Berlin: Lit-Verlag.