The Educational Potential of Inquiry Based -Workshops in a Science Museum

A descriptive analysis of IB education at Naturalis Biodiversity Center

Sofia J. van Santen
3220273
Master Thesis Science Education and Communication
Utrecht University
10-03-2014

Research Project SEC: 30 ECTS

Supervisor: D.J. Boerwinkel, Freudenthal Institute for Science and Mathematics Education, Faculty of Science, Utrecht University

Coordinator: S. Fechner, Freudenthal Institute for Science and Mathematics Education, Faculty of Science, Utrecht University

External supervisor: J. van der Brugge, content coordinator Naturalis Educational Development Department

RESEARCH REPORT

The Educational Potential of Inquiry Based -Workshops in a Science Museum

A descriptive analysis of IB education at Naturalis Biodiversity Center

Sofia J. van Santen

Table of Contents

Summary

This report provides the results on a structural evaluation of two Inquiry Based workshops, commissioned by Naturalis Biodiversity Centre. Through a descriptive analysis, using a mixed-methods approach, it was investigated to what extent these workshops contribute to IB education and this led to innovative recommendations. This provided insights on the potential of such structured IB workshops as an effective strategy to offer inquiry learning programs in science museums. As research on IB education focused on formal settings mainly, first a framework was developed bridging the gap between formal and informal science education. The two workshops, developed for upper level secondary school groups, were investigated throughout the three curricular levels (intended, implemented and attained) and evaluated in the light of inquiry learning and teaching theories. Since Naturalis did not clearly predetermine the objectives for the workshops, a method was developed to determine these in retrospect. Hypothetical learning trajectories were used throughout the study; first to reveal if and how the intended goals were addressed in the designs, then as an instrument to collect and analyze observational data providing results on implemented and attained level. Questionnaire data provided insights on students' experiences.

It has been indicated that school field trips to museums are often underused. And the results of this study seem to indicate that the evaluated programs indeed have greater potential than what is currently achieved. It is shown here how Naturalis could use its uniqueness, being a research institute with also an impressive collection. IB workshops with otherwise inaccessible objects could enhance student motivation while ensuring structure and guidance to support the inquiry learning. Physical and concrete examples of abstract, evolutionary phenomena can be provided while engaging the students in scientific inquiry in an authentic environment.

1. Introduction

Informal science education has a great potential in contributing to formal science education mainly due to the environment's unique objects, increasing the relevance of science subjects that are so often perceived by students as boring. Indeed studies have indicated that students appreciate class visits to out-of school settings (school field trips), and would recommend more trips as their number one priority to improve school science (Braund & Reiss, 2012). However, studies have shown that school trips are often underused; structure, prior knowledge and teachers' agendas have been indicated as important influential factors for learning results (Witt & Stoksdieck, 2008; Cox-Petersen, Marsh, Kisiel & Melber, 2003). Recognizing the curriculum programs and tight school schedules teachers have to deal with, the question rises how science education through field trips could be optimized.

A widely implemented strategy in formal science education is the use of 'inquiry', which led to the development and implementation of numerous derivative learning and teaching strategies in various forms. Inquiry Based (IB) education could support many aspects of science education; conceptual science ideas, knowledge, skills and understanding science as a way of knowing, all promoting scientifically informed citizens who have some appreciation of the beauty and wonder of science (Olson & Loucks-Horsley, NRC, 2000). A large body of research has accumulated on the implementation and attained outcomes of these various IB strategies. The degree of guidance and structure has been indicated as an aspect that should always be considered carefully in relation to the objectives the approach is serving (Olson & Loucks-Horsley, NRC, 2000).

The inquiry-oriented curriculum and policy developments have influenced the vision and practice of informal science education institutes as well. Biodiversity Center Naturalis in Leiden indeed adopted this view and states to aim at IB learning especially for the programs they develop for school groups. Naturalis is strongly affiliated to an IB approach, functioning as a research institute itself. This led to the development of IB workshops that could be a valuable contribution to in-school science education, offering inquiry with unique objects enhancing student engagement while ensuring structure and guidance to support learning.

Although Naturalis aims for the IB approach to guide their educational designs, it has not been clearly formulated what exact goals inquiry should serve, which forms of IB education would suit best, and through what design principles. As Bybee (2004) state, logic suggests that for the design of educational programs and teaching methods, educational outcomes should be identified first, then appropriate means could be implemented, and outcomes assessed. However, especially when it comes to IB education, these often seem clouded and obscure (Bybee, 2004) and indeed, in Naturalis' current practice this appears to be the case. As Thijs and van den Akker (SLO, 2009) put it, there are several levels of curriculum development; from ideal to practical situation (Table 1). And when implementing a learning or teaching strategy with intentions and goals, such as the IB approach, translations between these levels need to be evaluated carefully. To date no such structural evaluation has been performed for Naturalis' workshops and there is a demand for a study that provides a clear overview of the current programs and relates them to inquiry learning and teaching theories. Three aspects from literature will be examined; possible goals for IB education ('IB goals'), essential features for IB education ('IB features'), and the variations of these features.

Research on IB education and implementation has focused on formal settings mainly, therefore a framework needs to be developed first, as to bridge the fields of formal education evaluation, informal education and IB theories. This will provide an evaluation tool that will help to gain insight in Naturalis' educational practice taking into account the specific obstacles and possibilities of a museum setting.

Aim and Research Questions

This study aims to evaluate IB education in Naturalis through a descriptive analysis of two IB workshops developed for upper secondary school groups and thereby providing innovational recommendations. This might be of value for other informal education institutions aiming at IB education as well, revealing the potential of structured workshops as an effective strategy to offer inquiry learning programs. Besides the practical relevance, this study bridges the gap between formal informal and formal science education on inquiry learning and teaching, thereby contributing to theory forming on IB education and opportunities of science education through museums.

This study will seek to answer the following research question; to what extent do two Naturalis workshops designed for school visits contribute to Inquiry Based education? It is expected that the two workshops address inquiry from different angles and Naturalis is interested in gaining insight on how these different approaches relate to IB education and whether outcomes differ.

The following sub questions will contribute to answering this main question;

1. What are Naturalis' intentions for the workshops, and how do these relate to IB goals?
2. To what extent are Naturalis' intentions and the IB goals, features and variations addressed in the workshop designs, and through what design characteristics?
3. To what extent are the designs conducted in practice as planned and expected?
4. To what extent are Naturalis' intentions contributing to IB education attained?

Curriculum development is represented by Thijs and van den Akker (SLO, 2009) as consisting of three levels; the intended, implemented and attained level. This distinction, with subdivision to six levels, is useful for the investigation of implementation processes and helps to identify discrepancies between the curriculum forms. Table 1 indicates on what level the four research questions of this study take place.

For the investigation of Naturalis' intentions for the workshops, a method was developed to determine these in retrospect as Naturalis had not formulated predetermined IB goals. The method section will elaborate on this.

Investigation on the implemented level focusses on the operational aspect mainly, as the practice is evaluated by observations. The perceived aspect, that takes into account the interpretation of its users (the workshop supervisors in this case) is not directly studied here, however the observations may indirectly provide insights on this as well.

Measuring actual, attained learning outcomes, seems to be a disputable research objective, and will not be feasible at this point as it requires long term, large scale research. Therefore this study addresses the experiential aspect of the attained level, and provides insights on attained learning outcomes only implicitly.

Table 1
Forms of curriculum abstracted from Thijs and van den Akker (SLO, 2009), related to the research questions of this study

Curriculum levels		**Description**	**Research question**
Intended	Ideal	Vision (rationale or basic philosophy underlying a curriculum)	1
	Formal/written	Intentions as specified in curriculum documents and/or materials	1, 2
Implemented	Perceived	Curriculum as interpreted by its users (especially teachers)	
	Operational	Actual process of teaching and learning (also: curriculum-in-action)	3
Attained	Experiential	Learning experiences as perceived by learners	4
	Learned	Resulting learning outcomes of learners	(4)

2. Theoretical Framework

Goals for science education, ideas on how science is most effectively learned and how this should be supported have been subject for research for many decades. National Science Education Standards (NSES) have been developed by the National Research Council, part of The National Academies (NRC), aiming at the promotion of scientifically informed citizens. According to the latest framework as proposed by NRC in ‘A framework for K-12 science education’ it is of great importance that all students “have some appreciation of the beauty and wonder of science; possess sufficient knowledge of science and engineering to engage in public discussions on related issues; are careful consumers of scientific and technological information related to their everyday lives; are able to continue to learn about science outside school; and have the skills to enter careers of their choice, including (but not limited to) careers in science, engineering, and technology” (Quinn, Schweingruber & Keller, 2011, p.1).

These standards show the diversity and complexity of science education with learning objectives related to science specific knowledge and skills, understandings on the nature of science, as well as inducing active engagement and motivation for science and inquiry.

2.1 Learning Outside the Classroom

As Braund and Reiss summarize in the introductory chapter of their guidebook ‘Learning science outside the classroom’ (2012); science seems hard to learn, could be perceived as boring, and students may not always see its relevance for their personal lives. In the chapter ‘Learning Science at museums and hands-on centers’ it is stated that museums could be a valuable addition to science education in schools, as they may offer an environment in which students are motivated through handling real (otherwise inaccessible) objects, enable physical and concrete examples of abstract phenomena, engage in scientific inquiry and develop science interests. This point has been widely supported and reviewed, also by Rennie (2007), Rennie and McClafferty (1995) and Falk (1992). Research aiming to identify whether and to what degree this actually occurs is accumulating; the following aspects and conditions have recently being addressed.

Rennie and McClafferty (1995) emphasize the importance of activities before, during and after the visit, indicating the important tasks of teachers. They state that museum visits could complement learning in school for different kinds of learning objectives (motivation, introduction to new scientific topics, revise and consolidate the learning of concept), but highly depend on how the teacher integrates and prepares the visit.

In an extensive review on class visits to out of school settings, de Witt & Storksdieck (2008) also state that cognitive and affective learning can occur, though learning outcomes are strongly influenced by many factors such as structure, prior knowledge, teachers’ agendas amongst many others. They argue that despite their potential, field trips are often underused as learning experiences, and ways to maximize learning using the unique opportunities of the destinations need to be considered.

The degree and type of structure offered during the field trip itself is a controversial aspect in literature, and as de Witt and Stoksdieck (2008) summarize, “generally research suggests that in a museum setting, structured experiences can increase cognitive learning but may dampen interest overall or result in less positive attitudes” (p. 185). Guided tours or worksheets are often used by schools to ensure structure and content learning; however, these methods can highly conflict with active approaches as recommended by NRC serving NSES (Cox-Petersen et al., 2003). Educational institutions offering school programs are struggling to find the balance between structure, free choice and exploration to maximize both cognitive and affective outcomes.

Falk and Dierking (2012) did extensive research on science museum education and proposed an Interactive Experience Model (Falk, 1992) that emphasized the importance of personal background and interactions with the social and physical environment for museum experiences and learning outcomes. For school groups visiting museums, the sociocultural context is especially important referring to studies that show how learning is enhanced when students can socially interact in meaningful ways that relate to exhibitions (Falk &Dierking, 2012).

2.2 Informal Science Education

When evaluating science education outside the classroom, the framework developed by the Committee on Learning Science in Informal Environments might be of value (from the Board on Science Education of NRC, Bell, Lewenstein, Shouse & Feder, 2009). This framework covers the complete and diverse range of science education, articulating science-specific capabilities supported by informal environments. It was developed by adding two strands (1 and 6) to a four strands framework as developed previously for formal science education (NRC, 2007).

Table 2

Strands of Science Learning, adopted from Bell, Lewenstein, Shouse and Feder (NRC, 2009)

Strand	Learners in informal environments:
1	Experience excitement, interest, and motivation to learn about phenomena in the natural and physical world.
2	Come to generate, understand, remember, and use concepts, explanations, arguments, models, and facts related to science.
3	Manipulate, test, explore, predict, question, observe, and make sense of the natural and physical world.
4	Reflect on science as a way of knowing; on processes, concepts, and institutions of science; and on their own process of learning about phenomena.
5	Participate in scientific activities and learning practices with others, using scientific language and tools.
6	Think about themselves as science learners and develop an identity as someone who knows about, uses, and sometimes contributes to science.

The committee elaborates on how different informal settings (from everyday experiences to community based organizations, media and designed spaces like museums) may support learning on each strand. It is stated that literature shows evidence for learning in museums across all strands, with strongest support for strands 1; "learner excitement and strong positive emotional responses to experiences of science and the natural world", strand 3; "learners engage in exploration and interaction, 'doing and seeing', questioning, explaining and making sense of the natural and designed world", and strand 5; "engaging in science is also strongly supported especially in the general form of social interaction in which learners jointly explore and interpret the natural world" (Bell etal., NRC, 2009, p.161-162).

Museums offering science education profit from the high potential in learning in these strands (underscored in Table 2), but could possibly even broaden their educational value by addressing the other strands as well. By these means, they could provide curriculum relevant programs that meet with teachers agenda's, serving both cognitive and affective learning goals.

A strategy in formal science education that has been emphasized in reform documents over the past decades, as to deal with the wide range of science learning objectives, is the use of 'inquiry' (Lunetta, Hofstein & Clough, 2007; Anderson, 2002; Flick & Lederman (Eds.), 2004; Luft, Bell & Gess-Newsome, STA, 2008). And this has strongly influenced the educational vision of informal education institutions as well. In the following paragraphs IB education in the light of its goals, features and forms will be summarized. Considering the possible goals, a framework will be provided, relating the strands for informal science education as named above, to IB education. This will indicate the potential of IB education, related to the specific obstacles and possibilities of a museum setting. Based on this IB goal-framework, and the other IB aspects (features and forms), derivative evaluation instruments will help to investigate Naturalis' IB education.

2.3 Inquiry Based Education

Four components for science curriculum development have been indicated by Hodson (2003); 'learning science and technology', 'doing science and technology' and 'learning about science and technology' and 'engaging in sociopolitical action'. These components seem to relate to the NSES as cited previously, aiming at scientifically informed citizens. A widely implemented strategy in formal science education to meet with these objectives is the use of inquiry. In 2000 the NRC (National Research Council), dedicated a guide to 'Inquiry and the National Science Education Standards' which poses that "inquiry is at the heart of the NSES" (Olson & Loucks-Horsley, NRC, 2000, p. 6). "As pointed out in the NSES (NRC, 1996), students who use inquiry to learn science engage in many of the same activities and thinking processes as scientists who are seeking to expand human knowledge of the natural world" (Olson & Loucks-Horsley, NRC, 2000, p. 1), and through inquiry students will "acquire concepts of science, skills and abilities of scientific inquiry, and understandings about scientific inquiry" (Olson & Loucks-Horsley, NRC, 2000, p. 115).

The use of inquiry in formal science education led to the development of various learning and teaching strategies, and the body of research on implementation and attained outcomes of these approaches is enormous. A difficult aspect when addressing inquiry education is its definition, since various names, derivatives and interpretations are used and the shared understanding of what science as inquiry means, and what it should look like in practical education situations is still in development (Anderson, 2002; Crawford, 2007; Bybee, 2004; Colburn, 2000; Lunetta, Hofstein& Clough, 2007; Abd-El-Khalick et al., 2004; Luft, Bell, & Gess-Newsome, STA, 2008; Bybee, 2004).

As for this study, three aspects important for defining IB education will be examined; the possible objectives for using an IB approach, its essential features, and the possible forms. These aspects should naturally correlate, as the features that are emphasized, and the forms chosen, should depend on the learning objectives. These three aspects will be investigated when evaluating to what extent Naturalis' workshops contribute to IB education.

IB goals, the first aspect in defining IB education. The use of inquiry could serve several goals. These goals were related to objectives for science education and the strands for Informal Science Education by Bell et al. (NRC, 2009, Table 2), providing a framework (see Table 3) that bridges IB-, formal- and informal science education.

The most general division that can be made is inquiry as means, (inquiry as an instructional approach, in Dutch literature referred to as 'onderzoekend leren') and as ends (inquiry as a learning outcome, referred to as 'leren onderzoeken') (Abd-El-Khalick, 2004, Van Graft & Kemmers, SLO, 2007). These could subsequently be divided more specifically;

- ***Inquiry as means, or inquiry in science, intends to help students develop understandings of science content*** (Abd-El-Khalick, 2004). The process of inquiry is used for students to learn about a particular field, and investigations suit the field of science (Van Graft & Kemmers, SLO, 2007). Inquiry could then be used as the instructional approach to meet with Hodsons'(2003) curriculum component 'learning science and technology' which also relates to the informal science education strand two; inquiry could be used as a means to make students understand, remember, and use concepts related to science. Besides the use of inquiry as a means for content learning, inquiry may be challenging and can place concepts within a relevant context, thereby enhancing motivation for learning, engagement and a positive attitude towards science and inquiry (Stokking & van der Schaaf, 1999; Van Graft & Kemmers, SLO, 2007). Inquiry utilizes the natural curiosity of children (Van Graft & Kemmers, SLO, 2007) and could therefore be used to motivate active engagement with the subject (Stokkink & van der Schaaf, 1999) which also relates to Informal Science Education strand one; excitement, interest and motivation to learn about phenomena in the natural and physical world. Hodsons' fourth component 'engagement in socio political action' is also sorted under this IB goal. The action itself is not directly related to inquiry, however science learning through inquiry could be the first step for inciting what is described by Hodson (2003) as "acquiring the capacity and commitment to take appropriate, responsible and effective action on matters of social, economic, environmental and moral-ethical concern" (p.658). To encourage this, the inquiry would thus be used as a means. These three subdivisions of inquiry as means will be referred to as M 1, M2, and M3 (see Table 3).

- ***Inquiry as ends, or inquiry about science, refers to inquiry as an instructional outcome*** (Abd-El-Khalick, 2004). In this sense inquiry could contribute to both Hodsons' 'doing science' and 'learning about science' curriculum components. Doing science is described as; "engaging in and developing expertise in scientific inquiry and problem solving" (Hodson, 2003, p. 658). The objectives could relate to development of skills and abilities to perform inquiry (Olson &Loucks-Horsley, NRC, 2000; Stokking & van der Schaaf, 1999). This also relates to the curriculum goal to encourage a scientific attitude (not to be confused with the promotion of a positive attitude towards science), referring to scientific dispositions that could be fostered by doing inquiry (Van der Rijst, 2009, Osborne, 2003), or what is called the 'ethos of inquiry' by Stokking and van der Schaaf (1999). By doing inquiry, strands three and five of Informal Science Education are supported; relating to skills and participation in scientific activities, using scientific language and tools. Besides this, inquiry as ends also refers to inquiry about science. By these means it relates to Hodsons' 'learning about science and technology', which is described as; "developing an understanding of the nature and methods of science and technology, an awareness of the complex interactions among science, technology, society and environment, and a sensitivity to the personal, social and ethical implications of particular technologies" (Hodson, 2003, p. 658). Inquiry could be used to meet with this goal as it provides a critical context for discussion and reflection in which students develop a fuller understanding of the nature of science and technology (Schwartz & Crawford, 2004). This 'epistemological understanding on the nature of science' (nos) relates to the fourth strand Informal Science Education; to reflect on science as a way of knowing, a major aspect of scientific literacy (Wilson & Bertenthal, NRC, 2005). The inquiry could make students recognize that knowledge is developed by people (Stokking & van der Schaaf, 1999), and that they themselves use and sometimes contribute to science as well, meeting with the sixth strand of Informal Science Education. These two subdivisions of inquiry as ends are referred to as E1 and E2 (see Table 3).

This shows how IB education could attribute to different aspects of science education goals and science literacy. Although these IB goals are not entirely independent, for implementation and development of educational practices it should be clear what main goal is aimed for, as this should lead the development for the educational practices, emphasizing the right aspects of inquiry (Stokking

& van der Schaaf, 1999). This framework will be used as an instrument to identify Naturalis' goals for their practice, as a starting point for the evaluation of the IB workshops.

Table 3
Framework for IB goals; linking formal and informal learning objectives

	Inquiry as means			***Inquiry as ends***	
	M1	*M2*	*M3*	*E1*	*E2*
NRC (2000): 'three major learning outcomes of inquiry-based science teaching'	Conceptual understandings in science, develop knowledge of scientific ideas, concepts			Acquire abilities to perform scientific inquiry	Understandings about inquiry, understand the work of scientist
Hodson (2003): Curriculum components	Learning science and technology		Engaging in sociopolitical action	Doing science and technology	Learning about science and technology
Stokking and van der Schaaf (1999): possible objectives for inquiry	Develop content knowledge on school subject	Motivate active engagement with subject / practice independence		Gain knowledge on basics in inquiry; language, discourse, setting up an investigation	Recognize that knowledge is developed by people
				Development of skills and abilities	
				Ethos of inquiry; scientific attitude	
				Tools of scientist; technologies, values, methods, evaluation	
Strands for Informal Science Education NRC (2009)	2: Generate, understand, remember, and use concepts, explanations, arguments, models, and facts related to science.	1: Experience excitement, interest, and motivation to learn about phenomena in the natural and physical world.		3: Manipulate, test, explore, predict, question, observe	4: Reflect on science as a way of knowing
				5: Participate in scientific activities and learning practices with others, using scientific language and tools.	6: Think about themselves as science learners and develop an identity as someone who knows about, uses, and sometimes contributes to science.

IB essential features, the second aspect in defining IB education. The NRC guide 'Inquiry and the National Science Education Standards' (2000) recommends IB instruction and identifies specific features essential to inquiry-oriented teaching and learning centered on the learners mental activity (see Table 4) aiming at the development of scientific explanations (Bybee, 2004). The list of features sets the stage for discussion of instructional models and helps to structure student activities fostering student inquiry (Olson & Loucks-Horsley, NRC, 2000*)*. They apply across all grade levels, and it is stated that teaching approaches and instructional materials that make full use of inquiry, should address all these aspects. This is then referred to as ‘full’ inquiry and the path is followed, from “formulating scientific questions, to establishing criteria for evidence, to proposing, evaluating and then communicating explanations” (Olson & Loucks-Horsley, NRC, 2000, p. 27-28). ‘Partial’ inquiry may occur as well, when an essential feature is missing. For example if students are not engaged with a question but an assignment begins with an experiment. The sequential character of these elements, and often similarly represented (e.g., Van Graft & Kemmers, SLO, 2007*;* Justice et al., 2001*;* Primas Project guide, 2011; Spek & Rodenboog-Hamelink, SLO, 2011) is subject of discussion in literature since the nature of scientific inquiry is more complex and not as linear as proposed by such steps (Rens, Pilot & van der Schee, 2007; Lunetta, Hofstein & Clough, 2007). However, for the purpose of evaluating Naturalis’ workshops in the light of IB education, this set of essential features provides a useful tool as these elements could be checked on the different curriculum levels (from Table 1), and the design characteristics for these elements could be identified, thereby providing an insight on the IB education of Naturalis’ practice.

Table 4
Elements for inquiry-oriented teaching and learning adopted from ‘Inquiry and the National Science Education Standards’ (Olson and Loucks-Horsley, NRC, 2000, p. 29)

Essential Features of Classroom Inquiry
Learners are engaged by scientifically oriented questions.
Learners give priority to evidence, which allows them to develop and evaluate explanations that address scientifically oriented questions.
Learners formulate explanations from evidence to address scientifically oriented questions.
Learners evaluate their explanations in light of alternative explanations, particularly those reflecting scientific understanding
Learners communicate and justify their proposed explanations.

IB forms, the third aspect in defining IB education. Criticism on IB education that argues how the approach cannot be effective due to minimal guidance (e.g. Kirschner, Sweller, & Clark, 2006; Mayer 2004), has been refuted by the statement that the critics misinterpreted the fact that IB education is always minimally guided, and indeed depends on scaffolding (Schmidt, Loyens, Van Gog, &Paas, 2007; Hmelo-Silver, Duncan, & Chinn, 2007; Alfieri et al., 2011). As stated by Justice et al.(2001), IB learning relates to the educational theory of constructivism, therefore scaffolding and structured guidance should be ensured to support effective learning. This illustrates again the importance of a clear definition when using the term inquiry.

Ever since inquiry is used in science education, there have been countless variations of its use in the classroom, and as many attempts to frame and evaluate these various approaches. These evaluations led to several classification, or identification scales based on the level of a number of variables (e.g. student responsibility (Olson &Loucks-Horsley, NRC, 2000), teacher involvement (Crawford, 2000), guidance (Furtak, 2005), subject of inquiry (Khasnabis,2008). and other variables as evaluated by Spronken-Smith & Walker (2010), Colburn (2000), Cobern et al. (2009), Primas Project guide (2011), Prince & Felder (2006) and more). What becomes clear from these articles and different points of views, is that the 'forms', 'variations', 'levels', 'modes' or 'frames' as types of IB education chosen for an educational purpose, should depend on the predetermined learning objectives.

Staver and Bay (1987) provided an often used differentiation in IB forms; structured- (the problem and outline for solution is given, so-called 'cookbook' activities), guided- (materials and problem are provided, students must figure out the method) and open-inquiry (students formulate problem for themselves, most closely relating to 'real' scientific work) (as cited in Spronken-Smith et al., 2012; Colburn, 2000). According to Levy's (2009), this could be coupled to a scale from information- to discovery-oriented inquiry, with a focus on content learning towards a focus on process and attitude (as cited in Spronken-Smith, 2010). In an information framing, students experience research through already existing knowledge, acquiring a previously established body of knowledge. In a discovery framing, students understand and experience research through personal questioning, exploration and discovery in relation to new questions and lines of investigation. This seems to relate to the two general IB goals (as means, as ends) as described above and presented in the two columns of Table 3. Figure 1 gives an overview of the two extreme (structured- vs. open) forms based on the above named sources, with the guided form in between. This representation is oversimplified and it should be emphasized that in practice hardly one form occurs within one inquiry session.

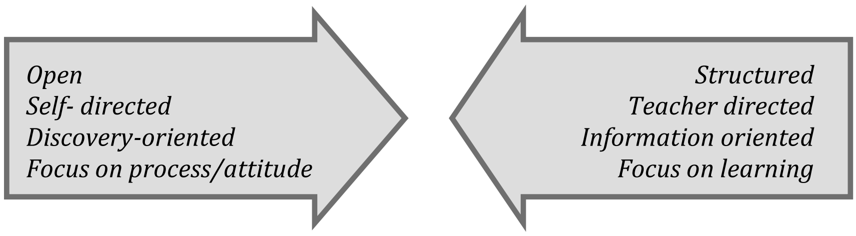


Figure 1 Overview of two extremes in IB educational approaches, with the guided form in the middle, based on reviews by Spronken-Smith et al. (2010) and Colburn (2000)

According to the NRC guide for inquiry, "investigations can be highly structured by the teacher so that students proceed towards known outcomes", "or investigations can be free-ranging explorations of unexplained phenomena". The variation depends largely on the educational goals for students, and because these goals are diverse, both highly structured and more open-ended inquiries should have their place in science education (Olson & Loucks-Horsley, NRC, 2000, p. 10-11,).

It should also be noted that even in the open, discovery-oriented forms of IB, were learners are mainly self-directed (going through full inquiry cycle; from engaging with topic, planning the investigation till evaluation of the research (Justice et al., 2002), teacher support is still necessary. It has been indicated that also for goals related to process, attitude and scientific dispositions, the teachers' role and their communication is of great importance (Lederman, 1992), as inquiry activities by students alone are not a guarantee for understanding science as a way of knowing (Van der Rijst et al., 2013*;* Swartz, Lederman & Crawford, 2004; Cobern et al., 2009; Alfieri et al., 2011).

For the evaluation of Naturalis IB practice the variations provided by the NRC guide on inquiry will be used. For each essential feature, four variations are described in a matrix (Table 5), with more to less learner self-direction, and less to more direction from the teacher or material. These variations could be seen as four forms between the two extremes as posed in figure 1. After checking the essential features in Naturalis' workshops, the variation could be determined. Then it would be interesting to see whether this suits to Naturalis' learning objectives.

Table 5

Variations of the five essential features, adapted from the 'Inquiry and the National Science Education Standards' (Olson and Loucks-Horsley, NRC, 2000, p. 29)

Essential Feature	**Variations from left to right: more to less learner self-direction, and less to more direction from teacher or material**			
	A	B	C	D
1. Learner engages in scientifically oriented questions	Learner poses a question	Learner selects among questions, poses new questions	Learner sharpens or clarifies question provided by teacher, materials, or other source	Learner engages in question provided by teacher, materials, or other source
2. Learner gives priority to evidence in responding to questions	Learner determines what constitutes evidence and collects it	Learner directed to collect certain data	Learner given data and asked to analyze	Learner given data and told how to analyze
3. Learner formulate explanations from evidence	Learner formulates explanation after summarizing evidence	Learner guided in process of formulating explanations from evidence	Learner given possible ways to use evidence to formulate explanation	Learner provided with evidence and how to use evidence to formulate explanation
4. Learner connects explanations to scientific knowledge	Learner independently examines other resources and forms the links to explanations	Learner directed toward areas and sources of scientific knowledge	Learner given possible connections	
5. Learner communicates and justifies explanations	Learner forms reasonable and logical argument to communicate explanations	Learner coached in development of communication	Learner provided broad guidelines to use sharpen communication	Learner given steps and procedures for communication

Inquiry in Naturalis. Naturalis aims for IB education, and workshops have been developed to serve this educational approach. Based on two reasons it will be posed how such workshops for inquiry could have great educational potential.

First, in relation to the strands of science learning, the workshops may address a wider range of curricular goals. As stated previously, evidence supports learning in museums mainly in strands 1, 3 and 5; relating to inquiry as a means (to motivate), and to inquiry as ends (to doing science), see framework Table 3. This means that strands 2 (content), 4 (science as a way of knowing) and 6 (science identity) are often less supported in science museums, whereas these might be of importance for teachers dealing with congested curricular programs and aiming to get the most out of their visit. The workshops, being educational programs with instructions from educators and providing supportive contexts, could support content learning (strand 2) and reflect on science as a way of knowing (strand 4) more than a regular museum exhibit. The IB workshops could be of additional value for museum learning, supporting learning on the strands otherwise left unaddressed in museums. However, while the workshops could, theoretically support learning on all strands and relating to all curricular aims, it is certainly not implied that this is desirable for each workshop. As it has been indicated above, for the development of educational programs the IB goals should be chosen, with a suitable IB form (Bybee, 2004; Thijs& van den Akker, SLO, 2009; Olson & Loucks-Horsley, NRC, 2000). Neither do the teachers' objectives for the museum visit relate to all strands, as the school trip should be reasoned and emphasize either affective or effective goals (Falk, 1992; Rennie, 1995). This implies the importance for clear communication between school teachers and museum.

Second, the IB workshop may provide a solution for the problem with guidance and structure for effective inquiry, and effective school trips. As stated previously, a balance between structure, free choice and exploration needs to be found for effective school visits to museums that support cognitive learning without impeding interests (DeWitt & Storksdieck, 2008). For IB education the right amount of guidance is essential as well, therefore museums aiming at effective science education through inquiry need to provide a structured form to ensure this. And as stated above, the goals for inquiry should be formulated so that a suitable form can be chosen.

By these means IB education developed for school groups in museum settings, could offer a solution for the problem with structure of field trips, while maximizing learning outcomes and meeting with teachers agenda's. Naturalis IB workshops could have great potential, offering schools a highly educative field trip with effective science learning. This study seeks to indicate to what extent these aims are being reached, whether this potential is used.

From this theoretical background, through evaluation of literature on both IB education, effective school trips and museum education one conclusion can already be drawn; predetermined goals (for both Naturalis and the school teacher), and the communication about these objectives, is a condition for desirable outcomes of participation in the workshops.

3. Study Design and General Methods

Methodology

This study adopted a mixed methods approach; using both qualitative and, to a lesser extent quantitative strategies addressing the research questions. Sequential methodological triangulation was planned; results from certain approaches were used to plan and develop the next (Morse, 1991). In addition, data triangulation was used when different sources address the same question. Both quantitative and qualitative data were used for descriptive analysis.

Procedure

As mentioned previously, the research questions addressed in this study could be linked to the curriculum levels of Thijs and van den Akker (SLO, 2009); intended (research questions 1 and 2), implemented (research question 3) and attained (research question 4). Table 6 gives an overview of what is investigated on each level, related to the research question and it briefly describes the research actions.

Starting on the intended level, first Naturalis' learning objectives were determined in retrospect, and it is investigated to what extent these relate to possible goals for IB education (IB goals) as studied in literature. Then the designs were evaluated to determine if these goals were addressed. Besides IB goals, the essential features for IB education and variations of these features were also examined. When moving to the implemented and attained levels, the study addressed how the workshops function in practice and (to some extent) whether intentions contributing to IB education were attained. Each curricular level will be discussed in a separate section of this report with a method, results and conclusion related to the particular research question.

The case selections

In accordance with the content coordinator of Naturalis Education Department, two of Naturalis educational workshops were selected for this study. Human Evolution (HE) and Fossils and Evolution (FE) both address the topic of evolution, were developed for the same target group (upper secondary school) and both aim for IB education. The workshops are thus comparable. However, they were developed by two different designers, and are expected to address inquiry from different angles. The two evaluations gained insights on how these different approaches relate to IB education and whether outcomes differ.

During the FE workshop, student groups work on different topics, there are five different assignments. One of these assignments, indicated by the designer as exemplary for what is intended generally for the workshop, was chosen as main object of investigation. This assignment addressed elephant evolution. On intended level, results will focus on this design, tough others will be briefly addressed as well. On the implemented level, only the groups working on the elephant topic were followed, so also implicit assessment on attained level concerns these groups. Analysis of students' experience obtained though the questionnaire included all participants of the FE sub groups.

Table 6
Overview of research questions on each level (Thijs & van den Akker, SLO, 2009) (first and second column), with brief description of what is investigated (third column) and how (fourth column)

Curriculum level	Related research question	Data/material	Research actions
Intended	*1. What are Naturalis' intentions for the workshops, and how do these relate to IB goals?*	Educational Ruler Learning objectives from script	Selection of principles according to designer Id IB goals of selected principles/ objectives
		Workshop design (materials and script)	Selection of principles addressed in designs Check material for script objectives
	2. To what extent are Naturalis' intentions and the IB goal, features and variations addressed in the workshop designs, and through what design characteristics?	Designer interview	Verify principles addressed in material with designer, come to agreed selection Verify IB goals with designer
		HLT	Produce HLT, determine indicators related to intentions, learning goals
			Check and describe characteristics for IB goals and features, determine IB variation of addressed features
Implemented	3. *To what extent are the designs conducted in practice as planned and expected?*	Observation material (guided by HLT)	Check HLT , student/supervisor activities
Attained	*4. To what extent are Naturalis' intentions contributing to IB education attained?*	Observation material (guided by HLT)	Implicit assessment learning goals and IB goals, features and variations
		Questionnaire	Check student experience related to selected principles from Educational Ruler and IB goals

Documents

Educational Ruler and the Basic Document.

Two documents have been composed by Van Keulen concerning Naturalis' educational principles. The first ('Onderzoekend leren in Naturlis', 2011) will be referred to as the 'Basic Document IB Naturalis'. Based on this document van Keulen developed a so-called 'Educational Ruler', an enumeration of goals as to guide the development of educational materials (see Appendix I). The list contains ten main principles, each consisting of three to ten aspects. Due to its size, this demanding list can hardly function as design manual, and it is indeed mainly used for general inspiration by designers. It will however be used as a starting point in this study, as it is the only document with written goals used by the education developers. Only the principles of interest were taken into account, namely those related to learning goals, student activities, attitude and addressing the right target group. Organizational aspects were left out.

The principles form this Educational Ruler were related to the possible IB goals, the underpinning for this allocation is provided in Appendix II. The workshop designs were evaluated to indicate if principle aspects from the Educational Ruler were addressed. To determine this, criteria were formulated and the designs were checked for these criteria. See Appendix III for underpinnings of these criteria.

Workshop designs

The HE and FE designs were considered as the developed materials including; objects, assignment sheets, worksheets and information brochures (the latter only for FE) and the workshop scripts; documents composed for the workshop supervisors including a time schedule and learning objectives. For FE, the materials for the elephant evolution group were investigated mainly.

Interviews

Semi structured interview with the HE and FE designers were audio recorded, backed up by written notes and then transcribed. Quotes from transcripts were used as illustrations, to emphasize aspects of the descriptive results or indicate designers' nuances. This was valuable especially to understand designers' interpretations of certain principle aspects, and to identify differences between the two designers.

The semi-structured interviews addressed several aspects; principle selections in the Educational Ruler (both as intended and apparent in the design), descriptions of desirable student behavior, learning goals from the script and IB goals. Results on these different aspects are considered in the relevant sections of the report.

Hypothetical Learning Trajectories

HLTs are a known tool for education design as well as education research. Simon (1995) state that (cited in Simon &Tzur, 2004): "An HLT consists of the goal for the students' learning, the tasks that will be used to promote student learning, and hypotheses about the process of the students' learning". As described by Bakker, and Van Eerde (in press), HLTs are useful for all phases of design-based research (DBR); preparation and design (to guide the design), teaching experiment (to guide observation and data collection), and retrospective analysis (to guide data analysis). Although this research is not DBR, HLTs were used in a similar manner here. However, since the design phase has already passed, HLTs were made to capture the expectations of the designs in retrospect. In education research it seems uncommon, to determine the predictions after the design process, it is however often used in product design processes when it's called 'reverse engineering'. It is predicted what an existing product should do, followed by analysis of its actual performance, leading eventually to redesign of the product (Otto & Wood, 2001).

In a sense this is what happens here as well; the HLTs are produced first to reveal if and how the intended goals were addressed in the designs, and later as an instrument to collect and analyze observational data. After evaluation on implementation and attained outcomes guided by the HLTs, innovation recommendations for the workshops could be given. Specifics on the use of the HLTs on the different levels and the implicit assessment approach will be discussed in the method sections on these levels.

Systematic observations

Through observation, both non-verbal and verbal elements (student's discussions and supervisors' interactions) were obtained. Visual data provided insights on student's activities relating actions from the HLT such as object examinations. Both Naturalis' supervisors as school teachers were asked for permission to record the workshop and students were informed about the study and given opportunity to object.

Observed workshops took place in the regular workshop areas; these were different for the two workshops. Due to noise disturbance in the HE setting, the observed group was separated from the rest of the group. The three students were places on another table in the sane workshop space, and for the group discussions they joined the rest.

Participants and contextual aspects

Based on convenience sampling, the booked workshops in November and December 2013 were observed. The Human Evolution workshop was observed four times; all fifth graders, two pre-university education and two senior general secondary education groups, each consisting of 18 to 20 students, from two different schools. The Fossils and Evolution workshop was observed two times; both fifth grade gymnasium groups of 17 and 12 students, from the same school. Different Naturalis supervisors gave the workshops, all experienced staff members.

Data collection consisted of video recording, audio tapes and field notes. First one group of students (in all cases groups of three) working on the assignments was recorded, then group discussions and whole-group assignments (for HE workshop), or the museum tour (in case of the FE workshop) were recorded.

Data analysis

For analysis of observational data on implemented and attained level, the HLTs were used and duration and frequencies of events were determined. Group discussions were transcribed. Descriptions with quotes from transcripts were used to illustrate exemplary or notable events. A second examiner analyzed the video recording of one HE group and used the HLT observation instrument. Student actions in the HLT were formulated as definite observable student behavior and a brief instruction was provided with descriptions of certain HLT elements to minimize room for interpretation. Results were compared and the similarity score was 90%, which validates the HLT as observation instrument to guide data collection.

Questionnaire

As certain IB goals, and intended principles from the Educational Ruler are highly intangible and hard to observe, these were addressed through a questionnaire. Mainly students' opinions related to curiosity, motivation and interest were obtained. And also their experiences on learning about certain aspects will be asked. Besides this, a large part of the observations focused on only the three students working on the assignment, and the questionnaire was used to collect data from all participants, providing a more complete picture. Specifics on the questionnaire will be addressed in the method section on the attained level, as results on the intended level gave input for the questionnaire development.

Participants, data collection, and contextual aspects

Students from all observed workshops, (the four HE, and two FE groups) and three additional groups for HE (consisting of 16 and 19 fifth grade gymnasium students, and 22 fourth grade senior general secondary education students) completed the questionnaire, right after the workshops were finished. Total of 136 HE students and a total of 29 FE students completed the questionnaire.

Data analysis

Descriptive analysis on questionnaire data was executed, using SPSS Statistics 20. Outcomes within the results of the two workshops were checked. As all participants were target groups no differences were expected. Indeed comparing the results between the seven HE groups no striking differences were found and results from all participants (fourth and fifth graders of the different education levels) were taken together for further analyses.

The answers of several open questions were categorized.

Language

When citing from written documents and materials used for this study, Dutch was translated to English. Also quotes from designer interviews and students' responses from the questionnaires were translated.

4. Intended level

4.1. Method

This section first addresses the question; '*What are Naturalis' intentions for the workshops, and how do these relate to the IB goals?'* Since Naturalis did not clearly predetermine the objectives for the development of the IB workshops, a method has been developed to determine and investigate these in retrospect.

Determining Naturalis' intentions for the workshops in retrospect

The designers were asked to select from the Educational Ruler which principle aspects were supposed to be addressed in the workshops. A checklist was composed for this purpose. In this checklist, three components of the first principle 'being a researcher' were further specified: the conditions related to 'a range of inquiry skills', the 'higher order reasoning' and the 'inquisitive attitudes'. Hereby the designers could select precisely which skill, reasoning form or attitude was intended to be supported by the workshop.

Through this selection of principle aspects in the Educational Ruler, the intentions for the workshops were made explicit, so they could be evaluated in the light of IB education, and on the implemented and attained levels. These selections for the two workshops were also compared to indicate differences in intentions. Besides this, the workshop specific learning objectives as stated in the scripts were evaluated.

Relating the intentions to IB education

As described in the theoretical framework, IB education could be used 'as a means', or 'as ends' and these two goals could be divided as well (Table 7). Inquiry as means is divided in three subcategories; M1 (means for content knowledge and science concepts), M2 (means to motivate), M3 (means to engage in sociopolitical action), inquiry as end is divided in two subcategories; E1(doing science; skills, attitude etc.) and E2 (learning about science; reflect on science as a way of knowing) (see Table 2 for complete framework).

Table 7
Possible goals for IB education

Inquiry as means			**Inquiry as ends**	
M1	M2	M3	E1	E2
Content knowledge, conceptual understandings	Motivation, interest	Sociopolitical action	Doing science (abilities; skills, attitude, dispositions)	Learning about science and inquiry (nos)

To evaluate Naturalis' intentions in the light of IB education, the selected principles from the Educational Ruler and the learning objectives from the script were related to these possible goals for

IB education. See Appendix II for elaborate underpinning of the allocation of IB goals to the principles. Only the first principle ‘being a researcher’ relates specifically to inquiry. The other principles of the ruler were linked to IB goals if it was clear that the translation of this principle in the workshop would contribute to one of these goals. For example, a principle stating the importance of appreciation for objects as a condition for motivation to learn is coded as M2 since the workshop uses the inquiry as a means to enhance motivation through the objects of investigation.

Based on the selections of the principles by both designers, linked to the IB goals, it was indicated which IB goals are addressed by each workshop.

Next to the IB goal identification of the designers’ intended selections and the learning objectives from the scripts, semi-structured interviews were held with the designers to verify which inquiry goals they find most important.

Design evaluation

To answer the second research question ‘*To what extent are Naturalis’ intentions and the IB goals addressed in the workshop designs, and through what design characteristics?*’ the following methods were used.

First, the designs (workshop scripts and materials) were checked for the Educational Ruler principles. For underpinning of principle indicators in the designs see Appendix III, both Van Keulens documents (Basic Document and the derivative ruler) were used. For several principle aspects it was hard to determine whether they were truly addressed in the design; some could only be checked on implemented or attained level.

Second, during the semi-structured interviews with the designers, the discrepancies between the two selections (the intentional selection by the designer, and the design check) were discussed. Consensus was found, and the agreed selection of principles addressed in the materials, with related IB goals, could be evaluated on the implemented and attained levels. This discussion also provided insights on designers’ intentions (contributing to the first research question) and how certain aspects were not addressed in the workshops. After conduction of both interviews, it became clear that the two designers had different interpretations of some principle aspects. These aspects regarded higher-order reasoning forms, and inquisitive attitudes. Indeed the Educational Ruler and the Basic Document are very concise in describing what exactly these forms and attitudes mean and other literature was consulted for the design evaluation of these aspects.

When discussing intangible aspects from the ruler, the designers were asked to describe students’ desired behavior, together with indicators from literature, this could be checked during observation.

Third, Hypothetical learning Trajectories (HLTs) were composed for the HE and FE workshops, based on the workshop assignments, the learning objectives from the script, the agreed principle selection from the Educational Ruler and interview results. Student behavior was described using behavioral verbs; action words that connote observable student behavior (Kizlik, 2002). For aspects related to inquiry activities and attitudes, student behavior indicators described in Van Graft and Kemmers' SLO document on IB education (2007) and Van Rijst dissertation on scientific research dispositions (2009) were used. Descriptions for desired student behavior from the interviews with designers were also incorporated. The designers verified the final HTLs.

The complete HLTs will be provided as a result on the design analysis, and based on this the designs were evaluated in the light of IB education. First the previously determined intended IB goals were evaluated for both designs. Then the essential IB features, and the variations of these features were identified, examining the intended self-directedness of students versus the amount of directions provided by the materials. Through description of these IB elements, the design characteristics for Naturalis IB education were provided, answering the second research question.

4.2. Results Intended Level

Table 9 gives an overview of results on the intended level related to the Educational Ruler selections for both workshops. Only principles of interest for this study will be shown here. For these principles, first the related IB goals are indicated. Then for each workshop, the intended principle selection is shown (‘x’ in the ‘Int’ column indicate designer selection), followed by the agreed selection for the design (‘x’ in the ‘Agr’ column indicate that the principle is addressed, ‘/’ indicates that this aspect could not be checked in the design). When agreed selection differed from the intended selection, brief underpinning for consensus is given. For underpinning of IB goal identification see Appendix II, for underpinnings of design evaluation for the principle aspects see Appendix III.

The latter columns indicate how principles of interest, or principles that could not be checked in the designs, were investigated on the implemented or attained levels. For those aspects that will be observed (o), it is indicated what behavior is aimed for (result from the interview with the designer), or could be expected based on literature.

4.2.1. Results regarding the first research question; ‘What are Naturalis’ intentions for the workshops, and how do these relate to IB goals’

The following sections (a-c) provide the outcomes of the different approaches to answer this question.

a. The Educational Ruler principles related to IB education

Table 8 gives an overview of the linked IB goals for all ten principles from the Educational Ruler.

Table 8

Ten main principles from the Educational Ruler by Van Keulen (2011) with related IB goals

Principle from Educational Ruler	Inquiry goal
1. Being a researcher	E1/E2/M1
2. Experience real things	M2
3. To raise wonder	M1/M2
4. Learning about diversity and change	M1
5. Exemplary work’	x
6. Heart, head, hands’	M2
7. Take differences into account’	x
8. Social interaction’	M1, E1
9. Give supervisors a role	x
10. Agreements with the school	x

The first principle ‘being a researcher’ relates specifically to inquiry, the topic of inquiry by students as an educational approach is named in Van Keulens documents. However, it seems that the goals for inquiry as named in the ‘Basic Document’ and the guiding Educational Ruler itself have some discrepancies. When looking at the aspects and descriptions from the ruler, E1 is addressed mostly, referring to activities, skills, abilities, attitudes conditional for proper inquiry. Except for one aspect ‘there is room for uncertainty’, related to E2 ‘learning about science’. When examining the ‘Basic Document’ by Van Keulen, it seems that the description for the use of inquiry relate mainly to E2 and

M1; ‘Students can develop a better understanding of various biological and geological concepts and the certainties and uncertainties, if they themselves work as a researcher’. All these different goals could indeed be the intentions for education; however it seems that the underpinning for inquiry does not corresponds to the deduced conditions in the Educational Ruler.

When looking at the principles of the Educational Ruler overall, it seems that if inquiry is used while meeting with, or even to specifically serve these principles, it would mainly be as a means for motivation and concept learning (M2and M1). Learning about science, addressing the nature of science (E2), was only linked to one principle aspect, although this aspect is emphasized in Van Keulens Basic Document. Engaging in socio-political action (M3) is not addressed in the Educational Ruler as matters of social, economic, environmental and moral-ethical concern were not named.

Besides the result that certain IB goals seem to be addressed more than others in the Educational Ruler, it was also noted that the inquiry aspects addressed seem adopted from curriculum documents for primary education, whereas these workshops target upper secondary school students.

b. Intended selection Educational Ruler by designers

See ‘Int’ columns in Table 9 for the selected principles by the designers on the principles of interest for this study. Both designers selected aspects in all ten principles, the HE designer generally selected more aspects.

The principle aspects related to the inquiry goal E1 ‘doing science’, mainly related to principle one ‘being a researcher’ and were equally selected by the two designers. Also aspect ‘there is opportunity or necessity to collaborate and discuss’ from principle ‘social interaction’ (not shown in the table) was indicated as E1 and selected by both designers. The only aspect related to E1 not selected by the designers was ‘the inquiry cycle is followed’. Aspects three, four and five of principle one ‘being a researcher’, list a number of elements (certain skills, higher-order reasoning forms and inquisitive attitudes) and these were all separately selected but not shown in the results Table 9. The intentional selections for these specific elements of the two workshops were quite similar; both selected the following skills; observe, collecting data, interpreting, explain, generalize, apply, verbalize, evaluate and present. Only the FE designer selected modelling, only the HE designer selected visualize.

During the interview with the designers, it turned out that they had interpreted the higher-order reasoning forms and inquisitive attitudes quite differently (compared to each other and literature) therefore the intended selections cannot be compared. These aspects will be discussed later.

E2 (learning about science, science as a way of knowing) had only one aspect related to it in the Educational Ruler, and was selected by both designers.

The HE designer selected more principles related to inquiry as a means (M1 and M2) compared to the FE designer. These were related to aspects in the principle ‘raise wonder’ (M1 and M2) and ‘heart head hands (M2).

Table 9 *Overview results intended level; Int (intentional selection by designer), Agr (agreed selection after design evaluation) for both workshops with underpinning for agreed selection if this differed from intentional selection. The final column indicates for both workshops how principle aspects are investigated (q= in questionnaire, o = through observation; then indicators are stated as well).*

Principle from educational ruler	IB goal	Human Evolution			Fossils and Evolution			HE and FE	
		Int	Agr	Underpinning	Int	Agr	Underpinning	Method of investigation	Observation / input for HLT
1. Being a researcher									
1. There is a research question	E1	x	i	not stated	x	e			
2. The inquiry cycle is followed (from question to searching process for result)	E1					i	main aspects from question to result are followed	o	are students following intended steps from inquiry cycle
3. The following skills are used …	E1	x	x		x	x		o	are students predicting, observing, interpreting etc.
4. Use higher-order reasoning forms…	E1-M1	x	x		x	x		o	are students reasoning in terms form-function etc.
5. The cultivation of inquisitive attitudes	E1	x	/		x	/		o, q	are students posing questions etc.
6. There is room for uncertainty	E2	x	x		x	x		o	are uncertainties addressed during group discussion,
2. Experience real things									
1. Work with real objects.	M2				x	x			
2. The objects raise questions and can focus attention.	M2	x	x			x	objects can arouse curiosity (aspect 3.1, 3.2) and thereby raise questions	o	are students posing questions when first handling the objects
3. There are rich and varied opportunities for direct sensory experience (seeing, feeling, hearing, smelling).	M2	x		only two					
4. Use of supplemental materials(replicas, model representations, text panels, videos, sensors) help to focus attention.	M2	x	x		x	x		o	are attributes used, showed, handled (HE), is extra information used (FE)
5. The objects are placed in an authentic context that helps children to interpret their experiences and give meaning.	M2				x		context could be a scientific one, though it is too simplified		
6. The objects have a story that can be discovered.	M2	x	x		x	x			
3. To raise wonder									
1. Confrontation with surprising objects or phenomena	M2	x	x		x	x		q	
2. Students' attention is focused by aspects that arouse curiosity and can be used to build upon	M2	x	x		x	x		o	do supervisors react and build on expressed curiosity of students
3. Natural curiosity of children is being used	M2	x	x			x	served by 3.1, 3.2		
4. Confrontation leads to reflection and questions from children	M1	x	/			/		o	are students discussing what they see, are they posing questions within the group
5 . These questions initiate a searching process that leads to knowledge / insight	M1	x		the investigation is not initiated by students' questions					
4. Learning about diversity and change									
1. Students prepare on Naturalis' focus on biodiversity	M1		/			/			
2. Students experience out-of school, informal learning	M1		/			/		q	
3. Focus on diversity and origin and meaning of this diversity	M1	x	x		x	x		o	are students posing explanations for differences they observe, changes over time
4. Attention for dynamics in natural environment of objects	M1	x	x		x	x		o	are environmental changes that caused certain developments discussed
5 . Students get an image of, and form an opinion about diversity and change'	M1	x	x	image forming only	x	x	image forming only		
6. Heart, head, hands									
1. Students are challenged to think	M2	x	/		x	/		q	
2. There is opportunity to act	M2	x	/		x	/			
3. Students get the space to appreciate objects and phenomena and express this	M2	x	/			/			
4. Students can personally identify with stories and objects	M2	x	/			/		q	

c. Learning objectives from the workshop scripts

Table 10 gives the learning goals as stated in the workshop scripts, and the IB goals related to them, after verification with the designers. Based on this, IB goals M1 (content knowledge), and E1 (doing science) seem to be addressed in both workshops. E2(learning about science) is also addressed in the HE script goals. M2 (motivation) and M3 (socio-political action) are not related to the goals from the script of either workshop.

Table 10
IB goal identification of learning objective from the Human Evolution and Fossils and Evolution workshop scripts

Workshop	Learning objective from script	IB goal
HE	Knowledge about human evolution	M1/E1
	Understanding scientific processes	E2
	Applying logical reasoning skills	E1
FE	Knowledge about evolution of [...]	M1
	Understanding scientific processes	E1
	Using skills: Identify and describing differences or similarities when selecting criteria for object determination Processing data in graphs and tables; obtaining information from graphs and tables.	 E1 E1 E1

d. Designer interviews

Human Evolution designer:

According to the HE designer, inquiry is mainly used as a means to motivate students, trigger interest and marvel them. Indeed almost all M2 related principle aspects were selected from the Educational Ruler and 'experience real things' and 'raising wonder' principles were indicated as being very important. It was stated that the museum collection should be used as a starting point to meet with these aims.

"*Based on these objects [from the collection] you may start to wonder, looking at things, and perhaps, posing questions[…] truly observing[…] you want them to practice skills for this purpose, aside from what they learn from it.*

In first instance the objects, the skull casts and the topic of human evolution should raise interest and motivate the students. Then there should be time and attention for careful investigation of objects, so through inquiry students will be motivated to engage with the subject, and experience excitement, interest and motivation to learn about phenomena in the natural and physical world (M2).

The designer also aims to enhance an inquisitive attitude, and several attitudes from the Educational Ruler were indeed selected. A curious attitude was especially emphasized during the interview and this is something he hopes students take home as well. These inquisitive attitudes form the ruler, as part of the principle ‘being a researcher’ were initially indicated as E1, contributing to abilities for ‘doing science’. The interview indicated however, that the attitudes, as well as the skills from the Educational Ruler are actually mainly addressed to enhance excitement and motivated to learn about certain phenomena, being a M2 goal again.

“I want student to pick up things from their surroundings a bit more often and have a look at it, wonder about it.”

Other learning objectives did relate to inquiry as end goal; to doing science (E1) and learning about science (E2). The first was addressed when discussing the learning objective from the script ‘knowledge about human evolution’. This seems a typical M1 goal, and the knowledge refers to the characteristics of the skulls that can be used to construct the human evolution lines. However, the designer emphasized the goal is not really for students to be able to reproduce these characteristics, the goal is for them to learn to reason logically.

“It's more about the process, and which features can be used and which cannot. It is about reasoning and suggesting possible explanations. Pupils should be rewarded for good reasoning”.

So the M1 goal, from the script, refers to both topic related concepts and argumentations (indeed M1), and to develop the required abilities; reasoning skills (E1). The selected higher-order reasoning skills from the Educational Ruler were also indicated das being important, the underpinning related both to E1; stimulating the use different forms of reasoning as abilities for object investigation, and to M1; using these forms as different angles to perceive and learn about these objects and the underlying phenomena. Besides this, and learning about tree reconstruction, M1 goals are also addressed and selected in the Educational Ruler ‘learning about diversity and change especially; “forming an image of the diversity of objects and changes through time” was being emphasized related to this principle.

Learning about science (E2), was stated as a goal in the script ‘understanding scientific processes’ and was confirmed as an objective for the workshop. These processes refer to theories about human evolution and mainly, how these are constructed and how they will always be developing due to new scientific discoveries. The related principle aspect ‘there is room for uncertainty’, which was selected by the designer was stated to suit very well with the intentions for this workshop.

‘T*his is really the program that could address this point of uncertainty, and it suits well with the level of the target group, showing the scientific practice’.*

Fossils and Evolution designer*:*

The main goal was indicated as doing science (E1); to let students participate in inquiry and enhance a scientific way of thinking. The intention is that students develop inquiry abilities, mainly to reason about their observations and measurements. Students should become aware of relations that can be discovered through investigation of objects. Then by discovering relations, sequences of fossils that represent trends in evolution, interest and motivation should be enhanced (M2).

"I want students to be dared to draw conclusions based on their observations, to explain and underpin what they see."

"*It's the kind of thinking that I find most important, and I hope they become enthusiastic then, and that they feel that they are encouraged to reason.[…] a way of looking at your environment, discovering relationships. […] indeed that's the most important. And then, to get excited about these relations.*

The designer was a bit more skeptical about the principle aspects from 'experience real things', 'raising wonder', and 'heart, head, hands' relating to the confrontation of objects provoking questions, curiosity, and personal identification with the topic. These aspects were not selected. So the M2 goal should really follow after the first discoveries have been done.

"You're dealing with adolescents and fossils and evolution is simply not part of their world. So no .. I don't really expect to give some turmoil there."

Although the principle aspect related to predicting was selected by both designers, the FE designer really emphasized this point. The HE workshop has a hypothesis assignments, and two reasons were given by the designer; first, to meet with the phases for scientific inquiry (E1), and second, it helps students to focus and observe the objects before they start with the measurements (M2).

As with the HE script objective related to M1, the knowledge about evolution of (elephants, horses etc.) is not to be taken literally. Students are supposed to learn about certain general evolution processes, and how evolution lines can be reconstructed, not specifically for certain animals. The higher-order reasoning skills as stated in the Educational Ruler could be used for this. The first skill as learning objective from the script, 'identify and describing differences or similarities when selecting criteria for object determination' is not really an intention for the workshop according to the designer. It was stated that it is more important that students learn that objects contain a variety of information, and at the same time, not all this information is appropriate for reconstruction of evolution lines. As with the HE, this goal relates both to M1, conceptual knowledge on evolution reconstruction, and to E1, as this should be revealed through inquiry. By doing science, students will acquire the right abilities and attitudes (E1). This is also what is meant by the script learning objective ‘understanding scientific processes.

In contrast to the HE designer, where the corresponding objective related to reflecting on science as a way of knowing, for the FE workshop, this objective relates more to the processes to retrieve information, which will be learned by participating in such inquiry (E1). The FE designer did address the ‘learning about science’ (E2) goal as well, not specifically related to nature of science and it’s uncertainties, but to understanding the work of paleontologists. How their investigations lead to knowledge about evolution. This illustrates another aspect of the E2 goal; making students aware that knowledge is constructed by people, and to a certain extent how they could be a scientist as well.

“H*ow objects from your environment can be studied to retrieve information. […]. You want them to experience what paleontologists do with fossils”*

4.2.2. Conclusion on the first research question: ‘What are Naturalis’ intentions for the workshops, and how do these relate to IB goals?’

What this study indicated so far, is that the learning objectives are not very precisely defined, and serve all kinds of IB goals. Intentions for the two workshops have many similarities, as both designers emphasize the aim to make students wonder, and to stimulate a curious attitude that students hopefully take home and use in their daily life as well. For HE this is really the main goal. Overall it could be stated that inquiry for the HE designer relates strongly to IB goal ‘as a means to motivate’ (M2), whereas for the FE designer, ‘doing inquiry’ comes first (E1).

The topic of human evolution and the objects (skull casts) are intended to enhance motivation, engagement and interest for evolution, science and the assignment. The FE designer was a bit more skeptical about the natural curiosity being stimulated by the topic or fossils; motivation and interest is intended to be provoked after students investigate and discover relationships. Inquiry abilities and skills need to be trained by participation in investigative activities, including the formulation of hypotheses. Specifically the scientific way of thinking was emphasized. Then, students may come to appreciate research and the information revealed by it.

Both designers indicated logical reasoning, as ability for doing inquiry, as an important goal. This was indicated as an E1 goal relating to a scientific way of thinking supported by doing science.

Learning how scientific knowledge is constructed, showing the nature of science including its uncertainties related to the IB goal ‘learning about science’ (E2) and was also indicated as an intention for HE. For FE the learning about science related more to showing how investigations, the work of paleontologists, contribute to knowledge development.

Both designers, to a lesser extent aim for conceptual knowledge about evolutional phenomena, to support imgae forming on diversity and change (M1). And students should learn that not all characteristics of the studied objects are useful for tree reconstruction.

Overall, this shows that both workshops relate to most of the possible IB goals (M1, M2, E1 and E2), with slightly different emphasis. Only M3, the IB goal related to socio-political action, was not indicated by either designer as an intention. This aspect was not related to any principle in the Educational Ruler, or in the workshops script nor was it mentioned during the interviews.

4.2.3. Results regarding the second research question; 'To what extent are Naturalis' intentions, the IB goals and essential features addressed in the workshop designs, and through what design characteristics?'

a. Agreed principle selection

See 'Agr' columns in Table 9 for the agreed principle selections of the workshops. For the principle aspects that could be checked in the designs, the intended selections by the designers were very similar to the agreed selections; intentions from the Educational Ruler, were mostly addressed in the designs. Appendix III provides the underpinning for the principle selection in the design. Discrepancies between intentional selection by the designer, and agreed selection will be discussed in this section, as intentions not addressed in the materials can have great consequences.

The principle aspect "children are challenged to verbalize their stories, experiences and opinions" (from principle eight, not shown in the results) was selected by both designers, though not visible in either design. This aspect emphasized the importance of knowledge anchoring and real concept understanding through verbalization. The enhancement to talk about the concepts, related personal story telling or sharing experiences, is however not specifically addressed in the assignments nor in the scripts for supervisor instruction. Both designers stated during the interview they indeed find this important, and that the group presentations are used for this purpose. However, during the presentations only one student gets to speak, and according to the design this only addresses their main findings on the investigation; no stories, opinions or experiences. During group discussions students could also tell personal stories, based on their experiences, pervious knowledge. Whether this will occur will be checked during observations.

Aspect 'the inquiry cycle is followed' from being a researcher was not selected by the designers, though for FE it was decided that it is implicitly addressed, as a there is a search process with a question and through steps of inquiry results are found and evaluated. The aspect 'there is a research question' is indicated as implicitly present in the HE workshop, as the research question is not stated anywhere, though there is an overall research question. The FE workshop assignments of all topics explicitly stated a research question.

The selected skills from the principle ‘being a researcher’ were generally addressed in the designs. Only the ‘presenting’ skill was selected by both designers, and is not addressed in the design. At least not in the way students’ presentations as final phase of research are meant; preparing a presentation, describing the research phases, and results, addressing the research questions and possibly discuss audience reactions (Spek, Rodenboog -Hamelink, SLO, 2011).

The following skills were agreed to be addressed in both HE and FE designs; observe, collecting data, interpreting, visualize, explain, generalize, apply, verbalize and evaluate. Predicting is only implicitly addressedin the HE workshop. In the FE workshop it is explicitly addressed for three of the five topics by assigning the students to formulate a hypothetical evolution line based on their first observations. It should be noted though, that in those worksheets that assign an hypothesis, this is not numbered as the true first assignment. The assignment indicated with number ‘1’ starts with the key determination. Modelling is only addressed in FE, as graphs and tables are used.

Some aspects of the ruler related to curiosity being provoked by objects (principle ‘experience the real things’ and ‘to raise wonder’), were not selected by the FE designer. However following the design criteria described for raising curiosity (don’t encounter them in daily life or at school, see Appendix III) they were selected for the agreed selection, just as the HE design. Indeed, the curiosity of adolescents might be difficult to trigger, and might differ between the objects used in subgroups of the FE workshop; perhaps giant elephant molars raise more curiosity compared to small mouse bones. It is hard to predict excitement; the questionnaire can only provide insights on this. Although the ‘motivational’ aspects of this principle were all selected for the designs, the ‘knowledge’ related aspect; ‘these questions initiate a searching process that leads to knowledge and insights’ was not selected as for both workshops the research questions are provided.

Reasoning skills and inquisitive attitudes.

During the verification interview it turned out that designers had interpreted some elements of aspects four (higher-order reasoning forms) and five (inquisitive attitudes) of principle one ‘being a researcher’ quite differently (compared to each other, and to literature).

Both designers indicated that they find the reasoning forms as listed in the Educational Ruler important. The HE designer even stated that it was the intention to address as many reasoning perspectives as possible, using them as different angles to perceive and get a grip on reality. It came to light however, that both designers had difficulties in distinguishing several reasoning forms. Besides this a comment should be made considering the development of reasoning skills as an educational learning objective.

Considering higher-order reasoning and argumentation skills in education, there is a distinction between general and domain specific reasoning. For general reasoning, Toolmin's model with several argument components is often referred to. According to this model main elements of an argument include claims, data, warrants, backings and rebuttals (Toulmin 1958, 2003 as cited by Venville and Dawson, 2010). Without going into depth much, this indicates that for the formulation of arguments students should learn to make several steps, considering evidence, justification and also counterarguments. Domain specific reasoning emphasizes the fact that there are different subject-specific ways of thinking; different types of questions are posed, using different types of reasoning to provide answers (Janssen & De Hullu). The higher-order reasoning forms from the Educational Ruler could indeed be used to addressing certain specific questions. However, as Boerwinkel (2003) states, a condition for learning and developing (higher-order) reasoning skills is that the reasoning itself should be central and addressed explicitly. Neither workshop seems to meet with this condition, nor do the scripts address this for supervisors.

For the sake of this study, as the designers stated that they find the higher-order reasoning forms important and interesting, it was indicated which workshop elements might implicitly relate or support the different forms. See Table 11 for an overview of the higher-order reasoning forms that were studied in literature, and what aspect in the designs could potentially support this kind of reasoning.

'Pattern recognition' might be the reasoning form supported implicitly by both workshops, whereas the other forms are not addressed, or even less apparent. For the HE workshop, the reasoning forms are mainly addressed through the group discussions, and thus depend highly on the discussion form and supervisor guidance. For FE besides pattern recognition, also cause-effect, form-function and product-area reasoning could be supported, and indeed the assignment touches upon these aspects. However, it should be noted that in the light of evolution biology addressing cause-effect and form-function should caution is needed, as misconceptions could arise if these two reasoning forms are confused. In case of the studied elephant topic, the form-function reasoning could address questions related to the molar chewing surface and what this form could be used for (processing certain food). The cause-effect reasoning could for example address the topic of natural selection as an explanation for the gradual change in molars; variation in characteristics, selection pressure and how the changing environment could cause the spread of more beneficial features within a population. It is doubtful though, if the assignments elicit this domain specific reasoning without explicitly guiding this.

Table 11

Higher-order reasoning forms with literature definitions ([a]Oorschot et al., in press, [b] Van Graft & Kemmers, SLO, 2007, [c] Boerwinkel, 2003), and parts of design that might (implicitly) support these reasoning forms

	Literature	In HE design	In FE design (elephant evolution)
Pattern recognition	Patterns can be organized and classified and graphs can be used to recognize patters [a] Regularity in space [b]	Patterns in gradual changes of characteristics, when all skulls are investigated and discussed	Pattern in gradual changes, becomes apparent trough table and graphs (assignment 3). And similarities in development other animals; during museum tour when all results are compared and linked
Cause - effect	Causal relations; what caused it and what is the consequence [c]	During discussion, arguments for causal relations in evolution (e.g. changing environment, brain size - tool use theories)	Environmental changes influencing available food, causing gradual changes in features over time. Assignment 4, what factors could be responsible for changes in molars
Form-function	Functional relations; what does it look like and what is it for [c]	During discussion; function of skull characteristic	Assignment 4, when animals' diet is related to the form of the molars
Unity-diversity	Comparative perspective; what is different, what is the same [c]		
Part-whole	What does it consist of, and of what is it a part [c]		
Change-continuity	What has changed, what has remained the same Related to evolution, change through time and to stages of development [c]		
Product-area	'Organism and environment': how organism is influenced by the environment and how organisms affect their environment [c]	During discussion; how environmental changes work as selection pressure for animals	Assignment 4; addressing environmental changes as selection pressure for animals.

The selections on the 'inquisitive attitudes' of the principle 'being a researcher' were quite similar for the two workshops. Cultivation of scientific attitudes are often named as educational aims, it is a controversial subject in education research though, due to its intangible character (Van der Rijst, 2009). It was hard to check the designs on their potential to cultivate certain attitudes, in agreement with the designers this was attempted though, and almost all attitudes could potentially be cultivated. Table 12 provides an overview of the inquisitive attitudes from the Educational Ruler, with definitions from literature, whether it was selected by the designers, underpinning for the designs' potential to cultivate this attitude, and indicators for student behavior as stated in literature or named by the designers.

The workshops' potential as discussed were similar to the intentional selections by the designers; so as far as the designs could indicate, it seems that the designers intentions to cultivate certain attitudes were achieved. During the interview it came to light that the attitude 'openness' was interpreted in different ways; the HE designer linked openness to intrinsic motivation, whereas the FE designer could not come to one understanding at all, and literature refers to it as ‘an open attitude for other interpretations and conclusions’(Van der Rijst, 2009).

According to the HE designer creativity is not used as the assignment is too directed, indeed this seems to be the case when looking at the design. The FE designer states that creativity is needed as the assignments are quite difficult, and first obvious outcomes need reconsiderations and therefore creative reasoning. When looking at the FE design, and the literature’ definition of a creative attitude, originality and innovative skill use or methods do not seem to be enhanced; the assignments direct the way and the information sources contain a lot of information for correct solutions. The final attitude listed in the Educational Ruler is ‘accuracy’, and although this aspect was not stated separately in Van Rijsts dissertation (2009) it is an aspect of a critical attitude. Critical reasoning was named in Van Keulens Basic Document (2011), but is not particularly stated as an inquisitive attitude, while it has been indicated by Van Rijst (2009) that being critical is at the core of scientific practice. It was indeed named by the designers as well; therefore a critical attitude is included in this study.

Table 12

I*nquisitive attitudes with definitions from literature ([a] Van der Rijst, 2009, [b] Van Graft, & Kemmers, SLO, 2007), whether it was selected by the designers (intended selection), whether the workshop has the potential to cultivate the attitude as discussed during the interview, and behavioral indicators for these attitudes based on the same literature and as stated by the designers.*

Inquisitive attitudes	Literature	Intended selection	Potential in design	Indicator during observation
Curiosity	To know, knowledge thirsty [a] become intrigued by unknown things, to perceive the environment with attention, looking for new challenges [b]	HE FE	HE/FE enhanced due to special objects	Observe with attention, handling the objects, posing questions, wondering
Openness	To share, openness towards other ideas and conclusions [a]	HE	HE/FE In group work, open for other students’ ideas	No negative reaction of students for other conclusions, opinions during group discussion,
Creativity	To be innovative [a] surprising solutions, act resourceful, use acquired skills and knowledge in an original way [b]	FE		
Tenacity	To achieve, not giving up until you are satisfied [a]	HE FE	HE during group discussion trying to convince others	Not giving up (distracted), activity of students throughout the workshop
Accuracy	Conscientiously, strive for perfection, tidy work [b] To be critical [a]	HE FE	HE/FE precise observations, needed	Handle objects with respect, look and describe, draw in detail/ work with tables and graphs with precision Question own ideas and data, honest management of data

b. Hypothetical Learning Trajectories

As a result on the extensive design evaluation HLTs are composed based on the agreed principle selections form the Educational Ruler (as stated in Table 9), interviews, the materials and scripts. It was also indicated where possibly certain higher-order reasoning forms could be supported, and where certain inquisitive attitudes could be cultivated, as discussed above. Tables 13 and 14 on the following pages are the HLTs for HE and FE. For HE also supervisors activities were included.

Table 13
*HLT of the Human Evolution workshop, with time indications for assignments as indicated in the script (T), assignment number (N) and texts from the worksheets (elements indicated with * refer to group activities not addressed in the worksheet), supervisor and student activities, and hypothetical learning.*

T	N	Texts from assignment sheets	Supervisor activity	Observable student activity	Hypothesized learning result and skill/ higher order reasoning forms from ruler/ inquisitive attitude
5			Give brief introduction about activities	Listening to supervisor	Understanding the goals for the workshop Motivated to start with the assignment
10	1	Observe the skulls and compare them	Throughout group work supervisor passes by; encourage reasoning, pose questions, build on students' curiosity/questions/ storytelling	Skull examination, expression of wonder, posing questions, comparing the skulls, discussing differences	Motivated, raise wonder (about skulls, differences, strange shapes) First image of diversity and change Skill: observe Attitude: curiosity
		Take the skull of the most primitive characteristics		Naming characteristics of old/recent, posing possible arguments , deciding with group which one is oldest	Being able to identify differences in primitive vs modern skull characteristics, recognizing and arguing which are more primitive First reasoning about origin of diversity and change Skill: observe, verbalize, explain Higher order reasoning: form-function Attitude: openness
		Draw the form of the skull on the shape of modern man on the sheet (modern man)		Observe closely, handling the skull, drawing (on the right scale)	Become more aware of detailed characteristics, get an image of primitive characteristics. Recognize obvious differences compared to modern man Skill: observe, visualize Attitude: accuracy
		Specify in your drawing the differences between the primitive and the modern skull		Examine, evaluate drawing. Verbalize, discuss the differences between own drawing and modern skull outline, emphasize in drawing	Become more aware of specific differences, being able to describe more specific image of primitive characteristics Skill: observe, evaluate, verbalize Attitude: accuracy
10	2	Examine your drawing and the skulls		Drawing is compared to the other skulls on table, other groups	Corresponding primitive characteristics are recognized Skill: observe, evaluate Higher order reasoning; recognizing patterns Attitude: curiosity
		Name 6 striking differences between the various skulls. What characteristics would you use to describe the main differences		Name differences explicitly, compare these differences, indicate the most important characteristics that differ, argue why	Being able to argue which are important to identify primitiveness Skill: collecting data, explain
	*	*Central discussion* *Three non-human skulls are excluded*	Explain procedure; each group shows skulls and names one primitive characteristic Pose guiding questions, addressing form-function Encourages to check for differences named by other groups in own skulls React and build on students' curiosity/questions/ storytelling	Naming primitive characteristic Listening, compare named characteristic to own skull Pose form-function relations	Realize that there are similar 'primitive' and 'modern' characteristics, in various degrees Learn to determine which characteristics can be used for tree reconstruction Skill: verbalize, interpreting, explain Higher order reasoning: form-function, pattern recognition, change-continuity
			Asks students to decide which skulls do not belong in the tree Encourage students to pose arguments	Discuss, apply knowledge/ pose arguments,	Being able to distinguish the monkey skulls based on arguments Skills: interpreting, apply, explain

5	3	Check all skulls with a number and place the numbers on the spot where you think they belong in the family tree	Hand out extra sheet with tree	Argue, compare all skulls, compare characteristics/degree of, discuss	Knowing the appropriate characteristics for tree reconstruction, reason the place of skulls in the tree, through the gradual changes in these properties Skill: predicting, apply, verbalize, explain Higher-order reasoning: pattern recognition Attitude: critical
20	4	Combine with the whole group all trees into the big tree on the poster by placing the skulls on the spots where you think they belong		Participate in group discussion, collaboration, apply knowledge/ arguments, verbalize, convincing others with arguments	Understanding and visualization of general evolutionary development of humans, general processes; reasons for changes through time Skills: visualize, generalize, explain Higher-order reasoning: pattern recognition Attitude: openness, tenacity, critical
	*	*Discussing the tree reconstruction*	Discuss mismatches, pose questions, encourage students' reasoning Showing reconstruction drawings (Lucy eg) Show tool React and build on students' curiosity/ questions/ storytelling Discuss the following topics: 'species' definition, island evolution, environmental changes, evolutionary processes other animals due to similar environmental changes (other than island), how new findings change ideas (uncertainties science), dead-ends in tree (general tree reconstruction), non-linearity tree, other data used for human evolution research	Reflect, evaluate, express doubts, listen, link to known concepts, tell personal stories	Understanding evolutionary tree reconstruction process, understanding origin and meaning of diversity, influence environment Become aware of science as a way of knowing, work of scientists Skills: evaluate, apply, generalize, explain Higher order reasoning: cause-effect, product-area Attitude: critical

Table 14.
*HLT of the Fossils and Evolution workshop, with time indications for assignments as indicated in the script (T), the assignment number (N) and texts from the worksheets (elements indicated with * refer to group activities not addressed in the worksheet), supervisor and student activities, and hypothetical learning.*

T	N	Texts from assignment sheets	Observable student activity	Hypothesized learning result, skill/attitude/higher order reasoning forms from ruler
10			Listening to supervisor presentation	Understanding the goals for the workshop Motivated to start with the assignment
50			Reading assignment, listening to group member	Understanding, more specifically the goal of the particular assignment; Understands why objects are used
		Take from the box the two most recent molars. Watch them closely. Try to make a hypothetical evolution line with all molars from the box based on the characteristics you see.	Observe and handle the molars. Choose the two recent molars, compare those with other molars, describe differences/similarities, argue about possible development, compose order in molars	Motivated, raise wonder about objects Recognize differences in molars Being able to hypothesize about possible development Skill: observe, predict, visualize Higher-order reasoning: pattern-recognition Attitude: curiosity
	1. determine	Identify the fossil molars using the determination key. Put the correct name tags to the fossils	Examine the molars closely, name observable characteristics, use key to identify each molar, assign label	Learn to work with a key Learn that certain characteristics are used to identify molars Skill: observe, collecting data Attitude: accuracy
	1. check	Check the name tags with supervisor	Ask supervisor to check	Skill: explain
	2. organize data	Write names in the blanks of sheet . Put them in order of age (oldest fossil at the bottom).	Filling in the table	To get an overview of the molars through time, recognize gradual changes Skill: modelling
	2. table	Fill in the table, write down for each species in what period they lived and give a brief description of the molars. Use the identification key and if the information from the file	Filling in the given table, determining characteristics, discussing in group, use/read extra available information, use determination table	Recognizing which characteristics are valuable for determining the species; Being able to describe valuable characteristics Skill: observe, collecting data Attitude: accuracy
	2. measure, graph	Measure crown height of all molars and put the results in the table and fill in the chart on page 2.	Determine what is crown height, determine difference in height and decide where to measure, measure with tapeline, fill in given table and graph	Learning what should be measured and how Skill: collecting data, visualize, modelling Attitude: accuracy, critical
	3. examine data	Look at the data in the table and graph. What characteristics change over time?	Use data from table and graph, describe changes, discuss in group	Becoming aware of gradual changes through time Skill: interpreting, generalize Higher-order reasoning: pattern recognition Attitude: critical
	4. evaluate data	What factors could be responsible for the changes in the elephant molars? Give arguments.	Posing possible arguments, listening to group member, open for other ideas	Become aware of possible reasons for diversity and change through time; influence of environment Understanding and visualization of general evolutionary development of elephant Skill: apply, explain, interpreting Higher-order reasoning: cause –effect, form-function, product-area Attitude: openness, curiosity
	4. compare	Compare the tree from the file to the hypothetical tree, was it correct? Explain any differences.	Comparing two trees, name differences, pose explanations for differences discuss arguments	Realize that different features change over time, though not all can be used for tree reconstruction Skill: apply, explain Attitude: critical
	4. extinctions	The tree has quite a few dead-end lines. Why did these animals extinct? Use the table and file when answering this question	Discuss and pose possible arguments within the group. Use information from table/graph and background information	Student relate information from different sources to explain a phenomenon, learn about general tree line interpretations Skill: explain, interpret, apply Higher-order reasoning: cause –effect , product-area Attitude: openness
10	*	*Group discussion. Each groups gives brief results, supervisor links different topics and addresses environmental influences*	Verbalizing main findings, listening to other groups	Verbalizing results of assignment Understanding and visualization of general evolutionary development; link own findings to results from other groups, recognizing similarities. Understanding origin of diversity, environmental influence Skill: verbalization, generalize, evaluate Higher-order reasoning: pattern recognition, cause–effect , product-area Attitude: curiosity

c. IB goal, features and variation in the designs

Now that discrepancies between intentions and the designs were indicated based on the selections in the Educational Ruler, the designs were evaluated more closely in the light of IB education. Based on the HLTs the second research question could be answered while describing the verified IB goals, checking for the essential IB features, and identifying the forms of these features.

IB goals for Human Evolution. The IB goals from designers intentions were; motivation and interest (M2), skills and abilities for inquiry (E1), learning about scientific investigation and the uncertainties (E2) and learning about human evolution and topic related concepts (M1).

The first IB goal, M2 motivation and interest, is addressed mainly by the topic of the workshop; human evolution is expected to get students' excited and interested even before the workshop has started. Then the objects, the skull casts, should enhance further curiosity as students probably never worked with these objects in school. These related to selected principles from the Educational Ruler; 'experiencing real things', 'to raise wonder' and 'heart head, hands'. Skills, in first instance close observations, are also used to as a means to motivate active engagement. Assigning the students to draw the skull is used to ensure close examinations and this should then provoke more questions and curiosity.

It was stated that logical reasoning, as an ability to acquire insights, is of great importance. This E1 goal seems to be highly dependent on the supervisor and students' participation during group discussions. The assignments seem to address mainly the differences between several skulls. Students are asked to find six differences, these should be filled in. Reasoning is appealed with the following question 'what characteristics would you use to describe the main differences', aiming to elicit a discussion about these six differences and which of these are more useful than others.It is doubtful however, if posing this question in the assignment elicits this logical reasoning, whether the students start to discuss and pose possible arguments and explanations. During the following central discussion pattern recognition is intended to be elicited, as they hear and see similarities in 'modern' and 'primitive' characteristics. Here it should also be discussed which characteristics cannot be used in the tree reconstruction, and why. The way this occurs, whether students are encouraged to reason by themselves, or whether the answer given, depends on the supervisor. The higher-order reasoning forms as discussed previously are highly domain specific and relate to M1, as they could be used to learn about evolution processes. Whether these reasoning forms are supported relies on the supervisor though.

The same accounts for the E2 goal learning about science and the uncertainties of knowledge construction. The worksheet assignment do not address this or leave space for uncertainties. During the group discussion at the end of the workshop, this is supposed to be emphasized by the supervisor.

The M1 goal to develop content knowledge, indicating which characteristics of skulls can be used to construct the human evolution lines is addressed in the worksheet assignment as stated above, but should be ensured during the discussion. As named previously, the assignment focusses on the differences. Other concepts about evolution processes, tree reconstruction, and the principle aspects from the Educational Ruler concerning 'learning about biodiversity and change' are also only addressed in the group discussion. In the design, 'origin and meaning of diversity' and 'the dynamics of natural environments' seem only addressed to a certain extent. The diversity in skulls is evident, but whether actual 'meaning of diversity' (on which selection takes place) and the changing environments (causing selection pressure) is addresses depends on the supervisor. This indeed relates to the previously mentioned higher-order reasoning forms. When this reasoning is not supported by the supervisors by posing the right questions, only the greater effects of these evolutionary processes are shown, and students might miss the aspects of variation and natural selection. The 'image of diversity and change', a selected principle aspect related to M1 is addressed throughout the workshops as the students are confronted with a variety of forms; two skulls to start with, then by drawing, differences are emphasized. Then students broaden the image by examining the other skulls, and during the tree reconstruction all different forms and change over time is made evident.

IB goals for Fossils and Evolution. The main IB goal of the FE workshop was indicated as E1, doing science, to enhance a scientific attitude and to make students aware of relations that can be discovered through investigation of objects. This is visible in the design as the assignments resemble scientific investigations more closely compared to the HE design. Most basic elements of the inquiry cycle are followed; several groups have to form a hypothesis, and through observation and measurements with tools (compared to only observations in HE), and data processing in tables and graphs, students seek for answers. It is also made explicit in the worksheet that students will investigate, and follow these inquiry steps as to find answers to the research questions. During all these activities a variety of both inquiry skills and attitudes seem to be supported. Reasoning about findings is domain specific, as the arguments aimed for relate to evolutionary processed, and therefore relate more to M1. Developing general reasoning seems not supported by the designs.

After relationships are discovered curiosity about these relationships is intended to be provoked. As well as motivation about inquiry; that investigations can reveal information otherwise remaining invisible, causing wonder that students might take home. This is how the M2 goal of motivation and interest is being addressed.

Both workshop confront students with objects they probably have not worked with at school, this could attribute to curiosity as well. However, it is expected that students are more excited about human skulls compared to the objects of the FE workshops, and there might be difference between excitement experiences between the subgroups of the FE workshop as well.

The conceptual knowledge, the M1 goals, related to the aspects form principle 'learning about diversity and change' are addressed throughout the assignments, sometimes supported by domain specific higher-order reasoning.

Looking at the elephant group, step by step the students become aware of the features that changed over time. This is initiated by the formulation of the hypothetical evolution line. In subsequent steps they will be guided through measurements and evaluation of their data, and learn how features changed over time. Hereby the students will develop an image of diversity and change, and pattern recognition might be stimulated.

The assignment thus focusses on change over time, and students should come up with the environmental influences on this development. As with the HE workshop, the 'meaning of diversity', in terms of variation and selection processes is not directly addressed through the assignment. The assignment does seem to steer upon form-function reasoning, when asking student to come up with reasons for the changed features. And this could also lead to cause-effect reasoning, in the sense of natural selection as an explanation for the gradual change in molars. If students independently come to this, linking to what they know about variation and selection pressure and the processes of evolution to their findings by this doubtful

The learning goal, for students to realize that not all characteristics can be used for tree reconstruction, related to both general reasoning skills (E1), and conceptual knowledge(M1). The way the assignment addresses this point, it seems to focus on the M1 aspect, as general reasoning seems not strongly supported. The fact that not all information is appropriate for tree reconstruction, could become apparent in the final assignment when students evaluate their hypothetical evolution line to the final outcome. Students based their hypothesis on a certain probably obvious feature, and then during the following assignments they are guided to collect certain data, and use this to reconstruct the elephant evolution line. The students are assigned to compare the trees and it is posed if their hypothesis was correct, and that differences should be explained. If their hypothetical line was different, this was probably due the fact that they based their line on the wrong feature, and they realize that apparently not all features can be used. So the assignment steers upon a certain outcome, thereby addressing this M1 goal.

The E2 goal for inquiry, which was mainly to show how paleontologists work, how their investigations lead to knowledge about evolution, is reached by the resemblance of such investigations. It is also made explicit in the worksheet that students are investigating, following certain steps to find answers to the research questions. Hereby they get the experience themselves, of how research provides insights. When looking at the uncertainties of science, it is stated above that the assignments literally asks the students if their hypothesis was 'correct'. Indeed the assignments directs for specific data collection, using the 'right' features for the tree reconstruction. So when meeting with the M1 goal, realizing that only certain features can be used, the workshop outcomes seem quite certain and leave less room for uncertainties.

e. IB Features and Variations on the intended level

Table 15 gives an overview of the five essential features as provided by the 'Inquiry and the National Science Education Standards' (Olson & Loucks-Horsley, NRC, 2000), as presented in the theoretical framework and it is indicated whether the features are apparent in the workshop designs and in what variation.

Table 15
The five essential features and variations for IB education abstracted from 'Inquiry and the National Science Education Standards' (Olson & Loucks-Horsley, NRC, 2000, p. 29). The circles indicate the features' variation based on the workshop design evaluation (dotted line; HE, solid line; FE)

Essential Feature	**Variations from left to right: more to less learner self-direction, and less to more direction from teacher or material**			
	A	B	C	D
1. Learner engages in scientifically oriented questions	Learner poses a question	Learner selects among questions, poses new questions	Learner sharpens or clarifies question provided by teacher, materials, or other source	Learner engages in question provided by teacher, materials, or other source
2. Learner gives priority to evidence in responding to questions	Learner determines what constitutes evidence and collects it	Learner directed to collect certain data	Learner given data and asked to analyze	Learner given data and told how to analyze
3. Learner formulate explanations from evidence	Learner formulates explanation after summarizing evidence	Learner guided in process of formulating explanations from evidence	Learner given possible ways to use evidence to formulate explanation	Learner provided with evidence and how to use evidence to formulate explanation
4. Learner connects explanations to scientific knowledge	Learner independently examines other resources and forms the links to explanations	Learner directed toward areas and sources of scientific knowledge	Learner given possible connections	*Teacher links to current knowledge*
5. Learner communicates and justifies explanations	Learner forms reasonable and logical argument to communicate explanations	Learner coached in development of communication	Learner provided broad guidelines to use sharpen communication	Learner given steps and procedures for communication

1. It was indicated previously that the HE workshop only implicitly addresses a research question which would be "how did the evolution of humans proceed" as stated in the interview. Besides the fact that the student might not even be aware that their activities are addressing this question, it could not be answered by these activities either.

   The assignments for the FE workshop provide a question which centers on objects that are investigated leading to gathering and using data to develop explanations, this feature is indicated as D.

2. During the HE workshop, students collect data by close observations and descriptions, focusing on the two skulls. The complete 'data set'; all useful characteristics and their gradual appearances, is revealed during the group discussion by other students and the supervisor. Therefore this feature was indicated as C; the greatest part of the data is given. Based on this evidence, the students are asked to use this, to reconstruct the human evolution tree.

   The assignments for the FE workshop direct the students to collect certain data; first close observations and characteristic descriptions, than measurements and translation into graphs, B . The evidence collected will be used to find explanations.

3. Explaining is addressed in the HE workshop, indeed the 'path from evidence to explanation' (Olson and Loucks-Horsley, NRC, 2000, p. 29) is emphasized in this workshop as the designer stated that it's all about reasoning. Although the students do not truly gather evidence, as this feature emphasizes the explaining phase, and the materials do not provide guidelines on how these should be formulated, the HE workshop is indicated as A. A great part of the explaining in this workshop occurs during the group discussions which are directed by the supervisor. The level of directedness will be indicated during the observations.

   In the FE workshop both evidence and reasoning should be used to establish relationships and formulate explanations. The students are provided with ways to analyze their gathered data, it is stated how table and graph should be used and links to explain relationships are provided as well; the FE workshop is indicted as C.

4. After the group composed the human evolution tree, it will be evaluated by the supervisor. Here the students could make the connection between their result and the current scientific knowledge, however students will not examine other sources themselves; the supervisor will have to make connections to the existing knowledge. The workshop is indicated as D, a variation not in the original NRC guide 'Inquiry and the National Science Education Standards' (Olson and Loucks-Horsley, NRC, 2000), but useful here as to indicate that not the students but the teacher makes the connections.

   The FE workshop also provides connections, but then the students need to examine these themselves; this is variation C. Students have to compare their results to provided results and extra information.

5. Neither workshop properly address this feature that emphasizes student communicating and justifying their explanations in a way scientist do; in a way that it can be reproduced, articulating the question, the procedures, evidence, proposed explanations and review of alternatives.

Based on the HLTs, the HE workshop meets with three essential features, and the variations were indicated in and between the two extremes. The FE workshop design seems to meet more with IB education as more essential features are addressed. As far as the material evaluation can indicate, the variations on these features are on the right side of the scale being less self-directed and more directed by the materials.

4.2.4. Conclusion on the second research question; 'To what extent are Naturalis' intentions, the IB goals and essential features addressed in the workshop designs, and through what design characteristics?'

For both workshops, it was indicated that the selected intentions from the Educational Ruler, were mostly addressed in the designs. For the FE workshop there are differences for the subgroups though, working on different topics with different assignments. The formulation of hypothesis is the most important difference; a step in the inquiry cycle that should enhance first awareness about the differences in objects, which is not assigned for all subgroups.

The main IB goal differed for the two workshops; HE using inquiry mainly as a means, whereas for FE the inquiry itself is an important outcome. Indeed the FE design seems to resemble scientific inquiry more closely compared to the HE design and more essential features are being addressed.

For the HE workshop many intentions highly depend on the implementation as the worksheet assignments do not address many goals from the designer, and these aspects should be discussed by the workshop supervisor during group discussions. Considering both the assignments and the planned discussions as the design, all intended IB goals (M1, M2, E1, E2) seem addressed more or less.

There is one notable discrepancy between the intended IB goals and addressed IB goals in the FE design, namely that the formulation of an hypothesis is not included in all designs, and if it is included it is not as being a distinctive assignment. The FE design addresses more essential features and learning goals through the worksheet assignments, therefore outcomes for the intentions depend less on the supervisor. At the same time, the directions from the assignments make the variation of these features mainly on the 'less self-directed side of the scale. The self-directedness for the HE workshop seems to vary highly, as features were indicated between both extremes.

5. Implemented Level

5.1. Method

From the intended goals that were addressed in the designs through certain assignments or activities, only those actually performed can potentially lead to the obtained outcomes. Therefore the most important goal of the implemented evaluation is answering the third research question; '*To what extent are the designs conducted in practice as planned and expected?'*

After piloting, the produced HLTs as described in the previous section were transformed into a data collection instrument. It was checked whether observable student behavior or supervisor activities and the procedural aspects were implemented as stated in the HLTs. Duration of certain parts of the workshops was timed, this could be compared to the time indications from the script. The number of questions posed by the supervisor during group discussion, and number questions posed by students were counted during the group discussion (HE) or museum tour (FE). These numbers are insignificant data itself, but illustrate the different forms these discussions had. Also guiding questions, students' questions and personal storytelling by students were counted, as these were aspects selected from the educational principle ruler (see Table 9 in the intended level section).

Results will discuss the operation of the HLT activities, mainly addressing whether elements were observed or skipped by the students and supervisors. Here the consequences will not be discussed yet, as this relates more to the attained level.

5.2. Results Implemented Level

Regarding the third research question, ‘To what extent are the designs conducted in practice as planned and expected?’

Human Evolution

Table 16 provides the HLT for the HE workshop with checks for the four observed groups. The observable activities indicated with an arrow (>) regard procedural or supervisor aspects, the others relate to student activities. Groups 1 and 2 were fifth grade pre-university education classes from the same school, groups 3 and 4 were fifth grade senior general secondary education classes from the same school.

The first notable result is that there is quite a difference in observed student and supervisor behavior between the four groups. The second is that several aspects from the HLT were not observed for any group.

The most important missed aspects, or observed differences will be discussed in order of the HLT elements for each workshops.

Table 16

Checked HLT for the four observed groups of the HE workshop, ‘>’ regards procedural aspects or supervisor activities, ‘!’ indicates that elements of assignment were executed in different order

Texts from assignment		Observable student /supervisor activity	1	2	3	4
		listening	X	X	X	X
1	Observe the skulls and compare them	skulls are examined	X	X	X	X
		expression wonder				
		questions are posed (within group)		X		
		skulls are compared	X	X	X	X
		differences are named, discussed within group		X	X	X
	Take the skull of the most primitive characteristics	characteristics of old/recent are named		X	X	X
		possible arguments for old/recent are named		X		
		group decides which one is most primitive	X	X	X	X
	Draw the form of the skull on the shape of modern man on the sheet (modern man)	observing the skull	X	X !	X	X !
		handling the skull	X	X !	X	X !
		drawing	X	X !	X	X !
	Specify in your drawing the differences between the primitive and the modern skull	drawing is evaluated				
		skull is examined again				
		differences between drawing of primitive skull and modern skull outline are named				
		make changes in drawing, emphasize differences in drawing				
2	Examine your drawing and the skulls	other skulls on table are observed				
	Name 6 striking differences between the various skulls. What characteristics would you use to describe the main differences	differences are named explicitly (filled in on assignment sheet)	X	X !	X	X !
		these differences are compared				
		it is indicated which characteristics that differ are most important				
		arguments are posed why certain characteristics are more important				

*	*Central discussion.*	> each group names 1 characteristic			X	
		listening	X	X	X	X
		> supervisor provides help (guiding questions/ answers)	X	X	X	X
		> supervisor encourages to check for differences named by other groups in own skulls		X	X	
		> Supervisor poses form- function questions	X	X	X	X
		> supervisor uses poster with list of all characteristics			X	
	Three non-human skulls are excluded	> supervisor asks students to decide which three skulls do not belong in the tree (stud = students, sup = supervisor)	All stud	2 by stud	All stud	All sup
		students pose arguments for exclusion of three non-human skulls				
		> time spend on discussion (in minutes)	14	16	16	19
3	Check all skulls with a number and place the numbers on the spot where you think they belong in the family tree	> supervisor hands out the sheet with tree	X			
		observe and handle other skulls on table				
		discuss within group for each skull where it should be in the tree				
		characteristics of skulls are compared				
		degree of difference in characteristic is compared				
4	Combine with the whole group all trees into the big tree on the poster by placing the skulls on the spots where you think they belong	posing argument for placing skull on certain spot in tree	X	X	X	X
		discussion about positions, disagreements, arguments	X	X	X	X
		> time whole-group tree reconstruction (in minutes)	5	3	5	6
		> times supervisor offers help /intervenes	5	1	2	4
*	*Discussing tree reconstruction*	> mismatches are discussed	X	X	X	X
		> supervisor shows reconstruction drawings (Lucy e.g.)	X		X	
		> supervisor shows tool	X		X	
		> supervisor discusses 'species' definition	X	X	X	
		> supervisor discusses island evolution	X	X	X	X
		> supervisor discusses environmental changes			X	X
		> supervisor links to evolutionary processes other animals due to similar environmental changes (other than island)				
		> supervisor discusses how new findings change ideas		X	X	X
		> supervisor discusses dead-ends	X	X	X	X
		> supervisor discusses non-linearity tree		X	X	
		> supervisor discusses other data used for human evolution research	X		X	X
		> supervisor poses questions	8	2	5	14
		> time spend on tree discussion (in minutes)	16	14	31	23
		students pose questions	2	2	2	0
		students tell personal stories	1	0	0	0

All groups did the drawing assignment. It should be noted though, that one set of worksheets it handed out per group, so one student does the drawing. In some cases it was observed that other group members were very involved; watching closely, helping, indicating on the skulls what should be drawn. Whereas in other cases one dedicated student did the drawing alone, while the others get distracted or go ahead with the next assignment, which is consequently missed by the student occupied with drawing.

Groups 2 and 4 started with naming the six differences before they did the drawing, this was indicated with a (!) in Table 16. Both did the drawing later, there was however not enough time to finish this part completely.

The subsequent evaluation of the drawings, specifying differences and comparing the drawing to other skulls (start of assignment two) was not performed by any of the observed groups. It seems that the students focus on their two skulls, missing the opportunity to link their findings to the other objects and recognize similarities. Now students did the drawing and stated the most striking differences (or the other way around). This was the first moment where students could become aware of gradual changes in skulls, and recognize patterns; all students missed this part of the HLT. These goals now rely on the whole-group assignment.

An important aspect that seems to be missed as well, is comparing the different characteristics that have been observed and listed (assignment two), and indicate which are important to identify primitiveness. It seems that students in most cases focus on only listing the differences. According to the HLT the question from the worksheet 'what characteristics would you use to describe the main differences' related to the M1 goal to learn that not all skull characteristics can be used to construct the human evolution line. And also to elicit logical reasoning skills. The design evaluation already indicated that it was doubtful whether this question would elicit the intended discussion. Indeed, the question seemed to be over read by some groups, or not interpreted in the sense it was meant. Whether the related M1 learning goal might still be reached totally relies on the group discussion.

Only one of the observed first central discussions (after second assignment) occurred as planned by the designer; each group naming one apparent characteristic. Often times the groups named all or several characteristics they had listed. The form of this discussion differed, in some cases first all groups named their findings before the supervisor raised questions, whereas other supervisors did this in between, asking for clarifications. All supervisors posed guiding questions related to the function of certain characteristics at a certain moment. Encouragement to check certain features in the skulls, was observed in some cases, though this often occurred only when discussing the foramen magnum. One supervisor used the poster with the enumeration of all characteristics, repeating what was named by the students.

Then, it should be indicated which skulls do not belong in the human evolution tree (three ape skulls) and why. The latter, argumentation for exclusion was not observed for any group, and one supervisor did not let the students point out the non-human skulls but gave indicated them herself.

This HLT part related to both E1 (reasoning) and M1(content) goals; posing questions related to content learning and eliciting students' reasoning by supervisors should support these goals. Whether learning to reason is supported depends on the way the supervisor implements this part of the workshop, and apparently supervisors have diverse styles, and thus at least not for all groups this learning objective will be reached..

Assignment three was skipped; one supervisor did hand out the extra worksheet, but the students only had to decide where their own skulls should be placed, missing still the HLT aspects for this assignment. Here students should work within their groups on a smaller version of the tree reconstruction; using all skulls, predicting the place of the skulls in the tree and arguing why, discussing which characteristics should be used. This would be a preparation for the following group assignment, where similar goals are addressed. Now these learning goals (quite essential E1 reasoning goals and skills, and M1 content learning goals) now rely on the whole-group assignment, and might only be reached by those very participative students.

One supervisor did an extra activity, not stated in the script; students had to place all skulls in one line, from primitive to recent. After discussing this, the group started with the tree reconstruction with the poster.

The whole-group assignment with the poster tree reconstruction took three to six minutes. The number of interruptions (guiding questions or instructions) by the supervisor differed between the observed groups. The level of participation between students seemed to differ highly within and between observed groups.

The observed final group discussions were quite different, both in form and the topics addressed. This affects learning outcomes related to M1 (content), E1 (reasoning skills), E2 (learning about science) and it might influence students motivation, M2.

The use of supplemental attributes (from principle two 'experience real things'), such as the hatchet and the reconstruction drawings (which also relates to aspect of image forming from 'learning about diversity and change') differed between supervisors. The same accounts for addressing certain topics during the discussion such as general statements about tree reconstruction (from principle 'exemplary work'), and concepts from the principle 'learning about diversity and change'. Links to evolutionary processes of other animals due to similar environmental changes, was not addressed by any of the observed supervisors, missing the principle aspect 'dynamics in environment' and 'general statements based on exemplary work'. This might have been addressed before the workshop started though, during the guided museum tour which was not included in the study.

It seems that supervisors have a personal way of dealing with this part of the workshop. As shown in the results table, one supervisor addressed almost all topics and used all attributes. This is reflected in the time spent on this part of the workshop, being longest with 31 minutes, much longer than indicated in the script. Others took half the time, but consequently addressed fewer topics. The number of questions posed by the supervisors differed highly as well, this illustrates the different discussion forms observed; from interactive to more top-down style. Besides the number of questions, it was observed that the type of questions differed as well, from open to closed ones. The occurrence of questions raised by students, or storytelling's by students, was quite low for all observed groups.

Fossils and Evolution

Two 5-gymnasium groups were observed, Table 17 provides the HLT for the FE workshop with checks for the four observed groups. Asterisks (*) in the table mark where students asked the supervisor for help, double asterisks (**) indicates that supervisor was highly involved during the assignment, helping with guiding questions. The equal sign ($^=$) indicates that those activities were executed simultaneously instead of sequentially.

Table 17
Checked HLT for the two observed groups of the FE workshop., '>' regards procedural aspects or supervisor activities, '' the moment students asked for help, '**' when supervisor was highly involved with assignment, '=' elements were executed simultaneously, '!' element of assignment was executed later*

Texts from assignment		(observable) student activity	5	6
Introduction		listening to presentation	X	X
read intro assignment		reading	X	X
		handling the molars	X	X
		pose questions about objects(within group)	X	X
		listening to group member	X	X
hypothesize	Take from the box the two most recent molars. Watch them closely. Try to make a hypothetical evolution line with all molars from the box based on the characteristics you see.	observe molars	X	
		handle molars	X	
		choose two recent molars	X*	
		compare with other molars,	X	
		describe differences/similarities	X	
		argue about possible development	X	
		compose order in molars	X	
collecting data: 1. determine	Identify the fossil molars using the determination key. Put the correct name tags to the fossils	examine the molars	X	X
		name observable characteristics	X	X
		identify each molar using the key	X	X
		assign the labels	X	X
		> time occupied with determining (in minutes)	17.30	22
1. check	Check the name tags with supervisor	check	X	X
2. organize data	Write names in the blanks of sheet. Put them in order of age (oldest fossil at the bottom).	fill in given table	X$^=$	X $^=$
2. table	Fill in the table, write down for each species in what period they lived and give a brief description of the molars. Use the identification key and if the information from the file	fill in given table	X$^=$	X $^=$
		determine characteristics	X	X
		discuss in group	X	
		use/read extra information		
		use determination key	X	X

2. measure, graph	Measure crown height of all molars and put the results in the table and fill in the chart on page 2.	determine what is crown height	X*$^{=}$	X*$^{=}$
		determine difference in height	X$^{=}$	
		decide where to measure	X$^{=}$	
		measure with tapeline	X$^{=}$	X$^{=}$
		fill in given table	X$^{=}$	X$^{=}$
		fill in given graph	X$^{=}$	X !
linking 3. examine data	Look at the data in the table and graph. What characteristics change over time?	describing changes	X$^{=}$	X$^{=}$
		use data from table		
		use data from graph	X	X
4. evaluate data	What factors could be responsible for the changes in the elephant molars? Give arguments	posing possible arguments	X**	X
		listen to group member	X	X
		open reaction to group member	X	X
4. compare	Compare the tree from the file to the hypothetical tree, was it correct? Explain any differences	compare two trees		
		name differences		
		pose arguments		
		discuss arguments		
4. extinctions	The tree has quite a few dead-end lines. Why did these animals extinct? Use the table and file when answering this question	pose possible arguments		
		relate to information from table/graph.		
		use background information		
presenting	*Museum tour, passing all topics from subgroups. Each group gives brief results, links between different topics, links between environmental changes*	verbalize findings and arguments	X	X
		> similarities in topics is discussed	X	X
		> environmental influences is discussed	X	X
		> supervisor poses questions	16	
		students pose questions	4	
		students tell personal stories	0	

Compared to the HE workshop, there is less difference in implementation between these two groups; a greater share of the workshop is following steps as indicated in assignment sheets.

However, one group skipped the first assignments completely; the two most recent molars were not selected, and no hypothesis was formulated, thereby an important step in the inquiry cycle was missed. Besides this, close observations and starting to become aware of the variety of the objects was missed as well. This group started straight away with identifying the elephant molars, and this group also did not fill in the table in order of age. It was indicated in the design evaluation that this might not be a clear assignment for the students as it is not numbered. Indeed one of the groups started with assignment '1'. The group who did the hypothesis assignment, discussed within the group and posed arguments for their prediction. The questionnaire verified the implementation of the hypothesis assignment for the groups that were not observed and were also supposed to formulate a hypothesis. Only a small majority (58 %, n=19) had formulated a hypothetical evolution line before the determination assignment. This implies that indeed the assignment was not clearly enough.

Both groups had a lack of time, and did not make it to the final assignment (also the extra information file was not available). Here students should compare their hypothesis to a provided evolution tree and explain differences. Here they could learn to reason why certain characteristics are more or less useful in tree reconstruction. This related to the M1 goal to realize that objects contain a lot of information, but not all can be used for evolution tree reconstruction, and also to general tree interpretation addressing the topic of dead-ends. Also the skills to ‘apply’ previously gained knowledge and ‘explain’ are not used.

All other HLT elements and inquiry steps were implemented and executed; most planned skills seem to be used.

Interestingly, both groups needed help during the same task; when crown height needed to be measured. During the museum tour, supervisors of both groups addressed the same topics. Due to technical inconvenience, the tour of the second group was not recorded; there is no data to compare the two groups on the number of supervisor questions, students’ questions and student personal storytelling.

5.3. Conclusion Implemented Level

Regarding the third research question, ‘To what extent are the designs conducted in practice as planned and expected?’

The four observed HE groups differed highly, and none of them conducted the workshop design exactly as planned. Several expected students actions were not observed either; it appears that designs were often not eliciting the desired behavior. The results indicate that implementation highly affects the potential learning, related to all sorts of IB goals. This accounts for the FE workshop as well. Although differences between observed FE groups were less striking, the HLTs were not fully implemented as planned either.

Observed reasons for missed HLT aspects of the HE workshop related both to students skipping (parts of) assignments (due to over reading, misinterpreting or a lack of time) and to supervisors skipping whole parts of the planning.

For general reasoning (E1) it was indicated that the worksheet hardly addressed this, and now the implementation evaluation shows that those parts that should elicit reasoning, were not executed. This means that development of reasoning all depends on the group discussions, but it was the results showed that these took place in many different forms.

Besides the reasoning skills (E1), this will also affect learning outcomes related to conceptual knowledge about evolution, and goals related to science as a way of knowing (M1 and E2).

It was also observed that supervisors skipped parts of the HLT, or included a new element.

Indeed when looking at the curriculum levels as posed in the introduction (Thijs & van den Akker, 2009) the implemented level consists of ‘perceived’ and ‘operational’ sublevels. This study only addressed the latter, but the results indicate that perhaps the perceived level, particularly how supervisor perceive the workshop goals, is of great influence and importance as well. Especially if many learning goals rely on the supervisor activities, as for the HE workshop, their perception affects implementation and possible outcomes.

Overall, the implemented HLT activities seem to generally address the M1 goal related to image forming of diversity and change, being the result of evolution processes. Form-function seems to be addressed to a certain extent. Also E2 goals related to uncertainties, showing how findings change ideas on current knowledge seems to be implemented.

For the FE workshop, the implementation seems to follow the HLT more closely compared to the HE workshop. Only one of the two groups skipped the first assignment (formulating the hypothesis), indicating that a small revise in the design could probably fix this problem. Perhaps more problematic was the result that both groups could not perform the final assignments due to time restrictions. These latter HLT elements related to M1 and E1 goals. Reasoning skills, evaluating findings, connecting results to current knowledge are missed here. The M1 goals that are now missed, related to learning that not all characteristics can be used for tree reconstruction, and learning to interpret information from evolution trees (what dead-ends mean).

All other HLT elements addressing reasoning skills, and conceptual knowledge were executed. Therefore many skills and the M1 goals relating to the general evolution processes could be attained.

6. Attained Level

6.1 Method

Both qualitative and quantitative approaches were used to address the fourth research question *'To what extent are Naturalis' intentions contributing to IB education attained?*'

These intentions as determined previously related to the following goals for IB education: M2 goals (motivation, excitement, interest, challenge), E1 goals (inquiry abilities, mainly general reasoning and skills), M1 goals (concept knowledge about evolution, domain specific reasoning), E2 goals (learning about science as a way of knowing).

Observational data and the questionnaire data were used to provide a complete picture of attainment of these intentions, relating to students' experiences and learning outcomes

were used to gain insights on.

As stated previously, measuring actual, attained learning outcomes will not be feasible for this study. However, a so-called implicit assessment will be used to obtain insights on what students might have learned during the workshops. The composed HLTs as provided in the intended section will be used for this aim. The results of the observations on implemented level are the starting point as these indicate which activities, related to certain intentions, were performed. These activities were described in terms of observable student behavior that could be used as indicator for potential learning. This way of implicit assessment can be used for conceptual knowledge, skills and attitudes as stated by Kemmers, Klein Tank and Van Graft (SLO, 2007).

The evaluation on the implemented level already indicated certain aspects of the HLT not being executed as planned or expected, Consequently, the related learning goals can hardly be attained, or depend on other parts of the HLT with corresponding goals. This section will focus mainly on those HLT elements indeed implemented. For those elements, the hypothetical learning might have occurred if desired student behavior is observed. Situations will be described, using quotes from the video transcripts, to illustrate typical or striking observations that influence the attained outcomes.

Besides the descriptive analysis of hypothetical learning, a questionnaire will be used to gain insights on attainment of the intentions. As the observations mainly provided information on the followed groups of three students, the questionnaire data provided data about all students who participated in the workshops, increasing the validity of this study.

One questionnaire for each workshop was developed, considering their specific objectives, and they also consisted of a set of the same questions providing data to compare the two workshops. Both open and closed questions were included, and scaling (0 'strongly disagree' – 3 'strongly agree') was used for statement questions.

Questions related to; the selected principles from the Educational Ruler addressing intangible aspects (indicated with a ‘q’ in Table 9), intangible aspects of the IB goals (in relation to motivation, attitude and science as a way of knowing), student opinion on the workshop, asking the student to indicate which part of the workshop they enjoyed most, and which part they enjoyed the least. This provided information on student experience in relation to certain IB related features of the workshops. Appendix IV provides an overview of questionnaire and to what goals the questions related.

The final question asked students to indicate which parts of the workshop they liked most, and which parts they likes the least. Some students scored more than one element of the workshop. For analysis, the scores were divided; if a student selected three options, these were all counted for one-third.

6.2. Results Attained Level

Regarding the fourth research question, 'To what extent are Naturalis' intentions contributing to IB education attained?'

Results will be discussed for the main intentions as previously indicated; M2goals (motivation, excitement, interest), E1 goals (inquiry abilities; general reasoning and skills), M1 goals (concept knowledge about evolution), E2 goals (learning about science as a way of knowing).

M2 goals; motivation and interest

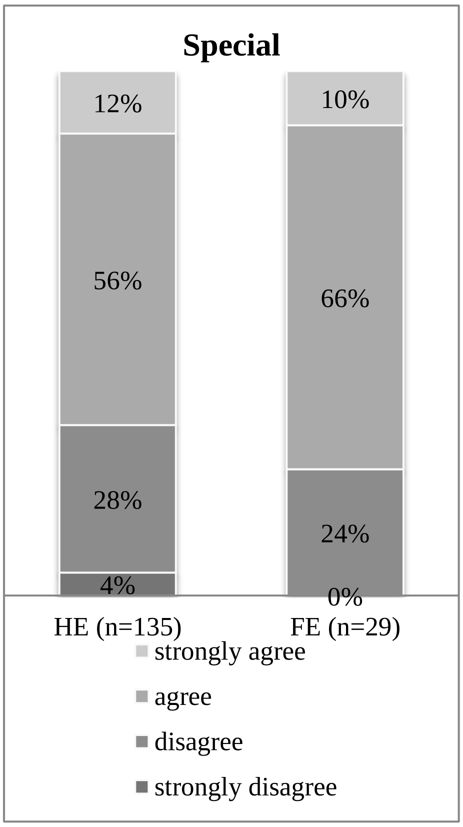
Implicit assessment. Indicators for wonder, excitement and curiosity were included in the HLT as 'students posing questions when first confronted with the objects'. This related also to principle aspect 'the objects raise questions and can focus attention'. Only one of the four HE groups was observed posing questions within their group at this moment. Students from both the FE groups did pose questions about the elephant molars.

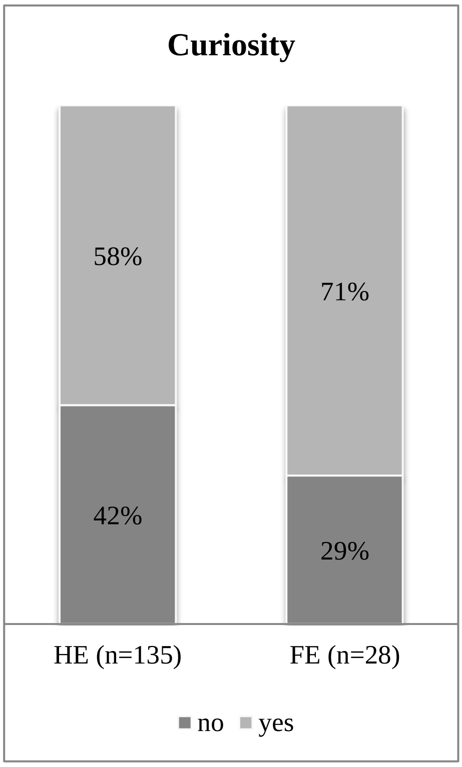
As this highly intangible aspect is hard to observe, or students might not express their wonder and excitement, the questionnaire addressed this topic through several angles of incidence. These questions also related to certain principle aspects from the Educational Ruler selected by both designers (indicated with a 'q' in Table 9 on the intended level section).

Questionnaire addressing M2 goals.

Wonder about the objects: related to principle 'raising wonder', aspect 'children are confronted with objects that could raise wonder'.

For both workshops, the majority of the students stated they had not worked with such objects before. As expected, this percentage is higher for the skulls than for the bones/molars (*HE: 93%, n=136, FE: 65%, n=29).* The goal to use objects that students are not confronted with daily, seems to be reached. Interestingly, the percentage of students agreeing on the statement that it was special to work with these objects is lower for the HE workshop. 68% of the HE students (n=139) (strongly) agreed on the statement '*I thought it was special that I could work with skulls'*. Whereas 76% of the FE students (n=29) (strongly) agreed on the statement '*I thought it was special that I could work with bones or molars'.*



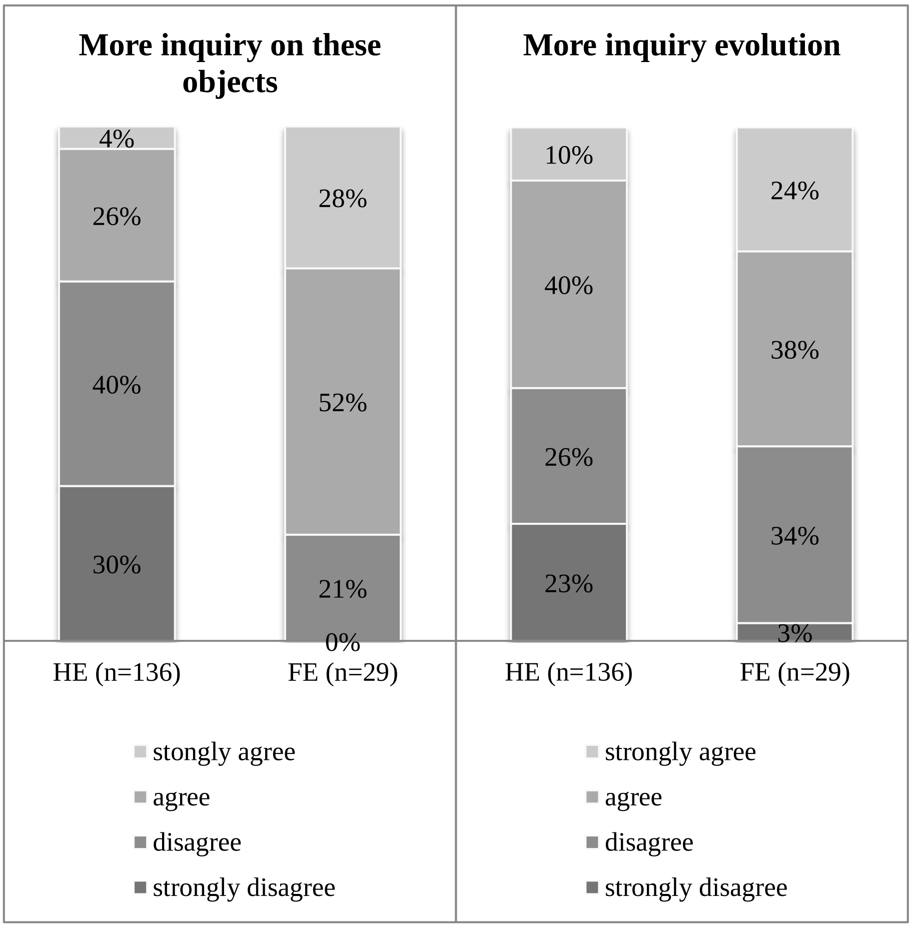


Curiosity: Was indicated as an inquisitive attitude, and also as an aspect to make students wonder. It was selected in the principle aspect 'students' attention is focused by aspects that arouse curiosity and can be used to build upon'. On the statement ' *did this assignment make you curious',* 58% of the HE students answered 'yes' (*n=135),* and *71%* of the FE students answered 'yes' *(n=28).*

Then it was asked; '*If yes, about what would you like to know more? (multiple options possible)'.*The three most selected options of the 78 HE students who answered yes were: 'more about the evolutionary development of the current topic' (44 students), 'more about the evolutionary development of other subjects' (31 students) and 'how evolutionary research is done'(29 students). The most selected option of the 20 FE students who answered yes was 'the evolutionary development of other subjects'(12 students). So, the majority of the students who did human evolution would like to know more about human evolution, whereas the students who did fossils and evolution would like to know more about other subjects.

For those who disagreed it was asked how come curiosity was not raised. The two most selected options of the 57 HE students to answered no were; 'I don't like evolution'(26 students) and 'the assignment was boring' (15 students). From the 11 students to selected the option 'other', the majority referred to evolution as not being a topic of their interest. Five of the eight FE students who answered no, selected were 'I don't like evolution' and two 'the assignment was boring'.

Motivation about evolution and research: The majority of the HE students (strongly) disagree with the statement '*I would like to do more inquiry on skulls'*; 30% strongly disagreed, 40% disagreed, (n=136), whereas the FE students do seem to be motivated in doing more inquiry on fossils; 52% agreed, 28% strongly agreed (n=29)Slightly more FE students would like to do more research on evolution compared to the HE students. The HE students however seem to disagree more strongly (23%, n=136), whereas the FE students agree more strongly (24%, n=29).



It was asked if the workshops had changed their opinion about evolution, and about doing research. Then, if their opinion had changed, they were asked to indicate what changed. 20% of the HE students (n=135) of the students said their opinion had changed; 10% got a more positive opinion about evolution, 9% stated what they had learned in relation to the topic related or about research. From the 38% (n=29) FE students with a changed opinion, 17% was more positive, and 17% stated what they had learned, either about evolution or research about it. None of the students indicated their opinion about evolution had become more negative.

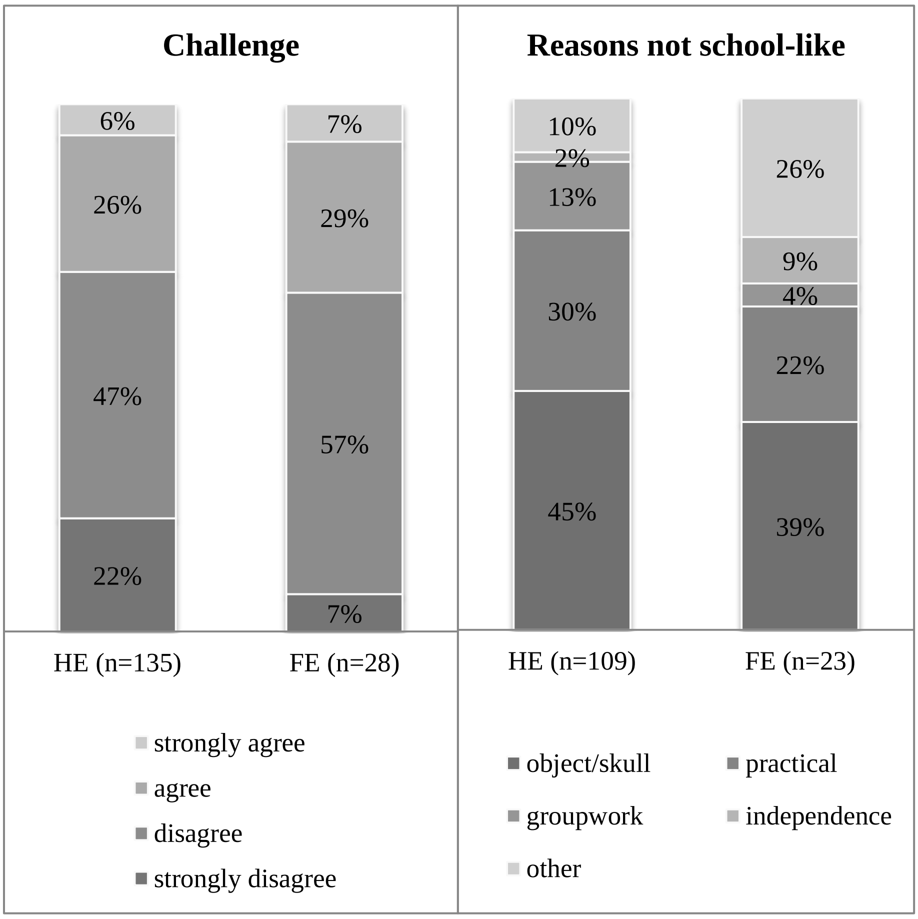
The same questions was posed, concerning their opinion about research. 11% of the HE students (n=131) remained positive, 7% remained negative, 22% stated what they knew already, 43% did not state their opinion. For 10% their opinion was changed, it had become more positive, 2% was more negative. 76% of the FE students (n=29) answered that their opinion remained the same; 21% remained positive, 17% remained negative, 14 % stated what they knew already, 24% did not state their opinion. From the 24% who had changed opinion, 10% was more positive, and no one stated that their opinion had become more negative.

<u>Challenge:</u> related to principle aspects 'the assignments are challenging and uncommon for school education' and 'students are challenged to think'

For both workshops, the majority thought the workshop was not challenging, (total of 69% of HE (strongly) disagrees, n=135, 64% of FE (strongly) disagrees, n=28). The division within the 'disagreement' is interesting, as 22% of the HE students strongly disagrees, whereas only 7% of the FE students strongly disagrees. It seems that students participating in HE feel more strongly in 'not being challenged' compared to students participating in the FE workshop.

For both workshops, the students strongly agreed on the statement that the assignment was different from assignments you get in school (HE:83%, n=133, FE:79%, n=29).

If students agreed it was asked what they find different. Answers were categorized and for HE 46% related to skulls, 30% to the practical form (n=109). For FE 39% related to the bones/molars, 22% to practical form and 26% to 'other' (n=23). Other referred to 'working nicely', 'the environment', 'independence', and 'it's cozier'.



Conclusion M2 goals HE and FE

Both workshops seem to raise curiosity but the Fossils and Evolution workshop seems to address this point slightly stronger. The topic of human evolution was expected to raise more interest, as students could relate to this. This was not supported by the questionnaire results; working with human skulls seemed to provoke interest less than the fossils. Curiosity could have been raised more during the FE workshop, as students were introduced to other subjects, and it seemed that they would like know more about these topics.

Also Fossils and Evolution seems to positively influence students' ideas on research as the majority would like to do more inquiry with objects, and the majority would like to do more research on evolution.

Both workshops were not experienced as challenging. Indeed the reason given for curiosity not being triggered was the fact that students said the workshop was boring.

E1 goals: general reasoning and skills

Implicit assessment E1 for Human Evolution. As stated previously the development of general, logical reasoning depends almost completely on the guided group discussions, as the worksheet assignments addressing reasoning were not implemented. For general reasoning then to be enhanced, supportive, challenging questions are necessary. These discussions were already indicated as taking place in all kinds of forms, from interactive to top-down, so it seems that supporting real reasoning will hardly occur. Illustrations of interesting observations will be given here, and the other E1 goals related to the use of skills and attitude will be addressed.

As discussed on the implemented level, all groups did the drawing assignment, however only one student did this activity. Dependent on their involvement, 'observation' and 'visualization' skills could be developed to a lesser extent for those who are not drawing.

After the drawing, students are intended to evaluate their drawing, and compare also other skulls on the table. This activity was not observed for any of the groups, Hereby the related skills; observe, evaluate, verbalize and visualize, were not used.

None of the observed groups addressed the question 'what characteristics would you use to describe the main differences', so no reasoning could develop. Searching for differences was observed, so the skill 'collecting data' is addressed, but again the skill to 'explain' is missing.

During the first group discussion, the supervisor should support logical reasoning by raising questions, encouraging students to compare, and argue about what they see. This occurred only partially, and only in some groups. Indeed the skills 'verbalization', 'interpreting', 'applying' and 'explaining' seem to be supported only in those groups were supervisor had a highly interactive style, posing many open questions, eliciting student to argue.

When ape skulls are excluded at the end of the discussion (mostly by the students), argumentation was not observed for any group, thereby missing the reasoning aspect during this HLT activity, as well as the possibility to verbalize and apply what has been discussed earlier. Now the skulls are only indicated as being different, students were not elicited to underpin their choice.

Then, during the whole-group assignment were several skills and reasoning should be used, it was observed that certainly not all students equally participate. And it is hard to imagine that a group assignment with 18-20 students, which takes about five minutes, arguing and reasoning occurs for all or even most students. In some cases it was observed that students gave a reason for placing a skull on a certain spot in the tree, however, oftentimes students just started placing the skulls until one would disagree but again arguments were hardly provided.

As with the previous discussion part of the HLT, the final discussion guided by the supervisor took place in many variations. Raising questions and building on students' questions could enhance reasoning but both the number, and the type of questions raised by the supervisors differed a lot. Whether students actually learn to reason during these guided discussions is doubtful as this should be a conscious learning activity. This was not truly assessed here, but at least it could be stated that due to the different styles supervisor have dealing with this part of the workshop, the learning goal seems not assured for all students.

The following situations illustrate this; during the final discussion the supervisor was explaining about every branch in the tree, told about new findings, showed tools and pictures, highlighted skull characteristics and their function, all very enthusiastically. All this, took half an hour, and the group was asked to think along explicitly only five times, the rest was all story telling. In a situation like this, learning to reason in certainly not supported.

Implicit assessment E1 for Fossils and Evolution. It seems that as the inquiry steps are followed, students are doing science while using the appropriate skills. Especially the close examinations, using the determination key, working with tables and graphs, analyze the gathered data will support many of the intended skills: observe, interpret, explain, visualize, modelling and applying. Learning to formulate a hypothesis will only be attained by a small part of the students though.

In relation to another E1 goal; 'supporting scientific attitudes', the following observation is worth noting. After measuring crown heights, data was processed in a graph. One group though had a hard time understanding their data. They measured repeatedly but their data seemed incorrect. Finally, they decided to change the data, to make a more or less continuous line.

> *'Make this one 5 cm, than it suits', 'ok, but it's still weird', 'all right, so make the other one 6".*

This shows that a fundamental aspect of a scientific attitude being accurate and critical is counterworked. And although students feel the urge to get a satisfactory answer, this form of tenacity was of course not aimed for, and when the designer stated that the 'creativity is needed as the assignments are quite difficult', this seems not to work out very well. What it means for conceptual knowledge will be discussed later.

Conclusion E1 goals HE and FE

For both workshops, it seems that not executing certain assignments strongly decreases the possibility for students to develop general reasoning skills. During the HE workshop, student argumentations were hardly observed during the assignments where they were supposed to, even if these assignments were executed. Besides this, reasoning largely depends on the group discussions, which were executed very differently and reasoning was certainly not assured for a large part of the students. The designer stated that students should be rewarded for logical reasoning, this intention seems not attained as reasoning seems to occur in the discussions hardly at all. Inquiry skills explain, interpret and generalize oftentimes not used. The attitude accuracy seemed to be supported by the close observations. Observe, collecting data, visualize and apply were skills generally used by the observed students.

For FE, the intention to perform inquiry seems attained; obtaining information from objects by observations, measurements, data analysis through table and graph use and interpretation were all executed. Only the formulation of a hypothesis, quite an important intention, and important step for inquiry seems attained for a minority of the students. Skills generally used were observe, collecting data, interpreting, visualize, explain, generalize, verbalize and modelling. Apply and evaluate to a smaller extend, depending on the museum tour and predicting only for those who did the hypothesis assignment. Whether scientific attitudes were supported is hard to state based on only the two observations, also the opposite effect was seen once.

M1 goals: concept learning goals

Implicit assessment M1 for Human Evolution. As stated in the intended section on IB goals in the design, many content learning goals rely on central discussions. This accounts for the domain specific reasoning, and the topics that should be addressed by the supervisor. And as shown in the implemented section, these discussions took place in many different forms. This already implies that learning outcomes are probably not assured for all students. Besides this, some notable observations affecting learning outcomes will be discussed briefly.

The first group discussion then should provoke form-function reasoning and pattern recognition. All supervisors did raise form-function questions, some more than others. For the realization that there are similar ‘primitive’ and ‘modern’ characteristics students should compare the characteristics named, check their own skulls again and realize there are degrees in these similar characteristics. All supervisors encouraged the students to check the foramen magnum in their skulls. Only a few did this for other features as well and it is doubtful whether indicating students once or twice, elicits reasoning about certain patterns. The pattern might become more clear then later, when the students reconstruct the tree with all the skulls. The change and continuity in skulls features could become apparent to the students when they see and hear about all features and how these changed over time. A distinction between what remained, and what developed was not addressed during the observed discussion however, characteristics that differ seems to be the focus. When ape skulls are excluded, mostly by the students, argumentation was not observed for any group, thereby missing the reasoning aspect during this HLT activity, as well as the possibility to verbalize and apply what has been discussed earlier. Now the skulls are only indicated as being different, students were not elicited to underpin their choice.

As indicated previously, learning that not all skull characteristics can be used to construct the human evolution line relied on the group discussion as the worksheet question was skipped by the students. However, for all observed groups, during the whole-group tree reconstruction assignment, the skull of Australopithecus boisei was placed at the root of the human line. This illustrates that the learning goal was not reached at this point, as apparently it had not become clear that boisei’s characteristics seem primitive, but are not useful in the human line reconstruction. All supervisors had to point out that their placement was incorrect. And they emphasized at this moment how a characteristic that is not apparent in the descendent lines should probably not be at the root, but in a different branch of the tree. Only know, this important learning goal seems to be addressed, and apparently it is not reasoned by the students but given by the supervisors.

The image of diversity and change does seem to be attained as indeed the students are confronted with a variety of forms throughout the workshops; starting with the two skulls, then by drawing, difference emphasized, then students broaden the image by examining the other skulls, and finally the image of the overall development is portrayed in the big poster. One note should be made though, as only one worksheet is handed out, only one student does the drawing. Although in some cases the other students were involved too, the hypothetical learning from the drawing, by the close observations and actual drawing, might be missed out by the group members.

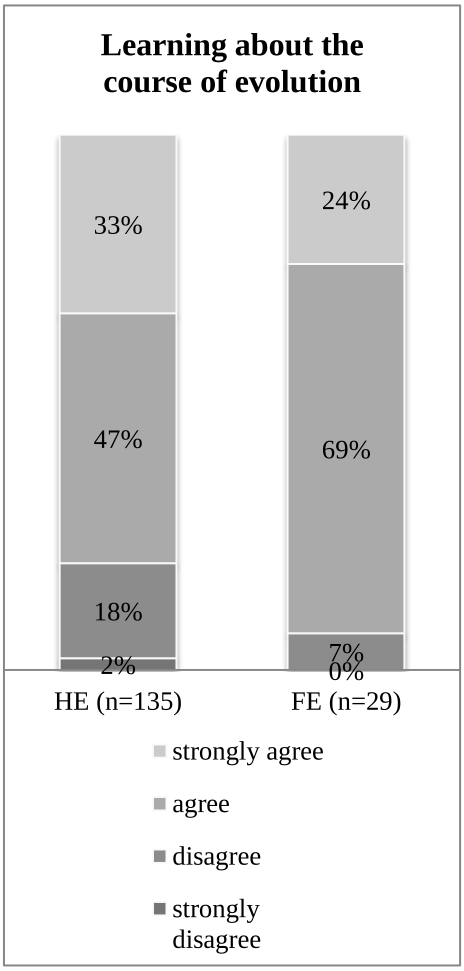
Only one student was observed telling a personal story, during all four observed group discussions. This related to the aspect ‘students are challenged to verbalize their stories, experiences and opinions’ and the intention was that supervisors then could built on it, linking these personal stories or experiences to the subject. It was indicated in the design evaluation that students are not elicited to tell personal stories through the assignments, and now it seems not to occur during the group discussion either. The same accounts for students’ questions that should be supported and used. As counted during the group discussion, students hardly posed any questions.

Implicit assessment M1 for Fossils and Evolution. As stated in the intended section on IB goals in the designs, the FE assignments guide students to gradually become aware of the characteristics that changed over time. When data are examined during the third assignment, changes over time were named by both groups. This could indeed stimulate the pattern recognition. Change continuity, the other higher-order reasoning addressed here, seems to be addressed to a lesser extent. As expected based on the design evaluation students seem to focus on what has changed over time. Then students posed these explanations about these changes. One group named both diet and the changing environment, but these students were highly supported by the supervisor during this assignment. By posing questions the students were guided to link a change in diet, to a change in the environment. The other group, working independently, came up with the explanation of diet as well. But the link to a changing environment was not made. This link was made later, during the museum tour when guiding questions were posed by the supervisor. Based on the observation it seems that learning goals related to ‘dynamics in the environment’ from principle ‘learning about diversity and change’ is hardly reached by the students independently through the worksheet assignments. Both groups needed guiding questions to link the observed variety in the molars to the changing environment, either during the workshop or during the museum tour. This shows that form-function, cause-effect and product-area were touched upon, with guidance. As stated in the intended section tough, not in the sense of real higher-order reasoning, but at least the assignment elicited thinking about form and function and environment to a certain extent.

During the museum tour, both supervisors addressed similarities between the topics. Hereby the learning about general evolution processes, related to principle aspect ‘general statements based on exemplary work’ seems to be supported.
Both supervisors addressed the topics of environmental influences and the outcomes of the different topics were linked; this could support pattern recognition, cause-effect and product-area thinking

The described situation from above, referring to the scientific attitude, is also illustrative for the M1 goal relating to general evolution concepts. It was observed that both groups started with connecting the dots in their graph. One student came to the conclusion that this could not be done; there was not one line but several branches. The other group had a hard time explaining what they saw in the graph and instead of reasoning how their findings could be explained, they started to cheat with the data. Their graph was correct though, but it seems that for this group the fact that evolution is not linear seems to be missed. This aspect was not addressed by the supervisor, nor explicitly through the assignment.

Questionnaire addressing M1 goals. Although student perception on content learning might depend on several factors, it is interesting to note that for both workshops the students indicated that through the assignments they learned about the course of evolution. 80% (n=135) of the HE students agreed, of which 33% even strongly agreed on the statement : By this assignment I learned more about the course of evolution. For FE, even 95% (n=29) agreed, of which 24% strongly agreed.



Conclusion M1 HE and FE

As with the previous goals, for the HE, attained outcomes on specific content knowledge seem also not assured for all students, as addressing certain topics, and learning to think about evolutional phenomena differed so much.

For FE students the domain specific reasoning was elicited to a greater extent compared to the HE workshop. Indeed if the assignments were executed the students were observed posing some arguments, related to form-function. However, it was also noted that for students to really come to reason about the phenomena behind their observations, terms of cause-effect, supervisor support was needed. The step to variation and selection was not made.

What does seem to be attained for students of both workshops, is to develop an image about diversity and change generally. Recognition in patterns, of gradual changes in the development of features seems to be assured as well. The actual meaning of diversity, and what processes and forces caused these characteristics to change over time, seems hardly addressed, but a global picture seems to be attained.

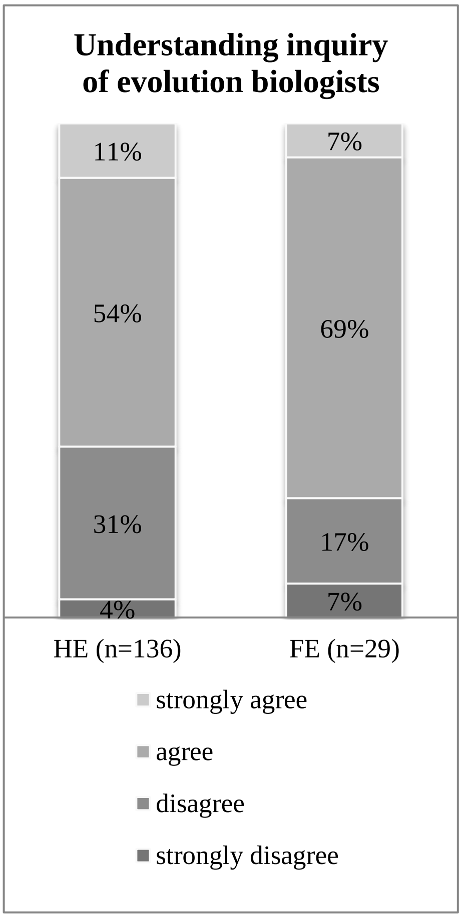
E2 goals: learning about science

Implicit assessment E2 for HE and FE. The worksheet assignments of both workshops do not address the nature of science, or seem to leave space for uncertainties.

During the group discussion at the end of the HE workshop, this is supposed to be emphasized by the supervisor. This was addressed in different ways; some named this aspect several times, some explicitly while others touched upon this aspects only superficially. Three of the four supervisors were observed addressing the topic of a recent discovery, influencing the current state of knowledge on human evolution.

For the FE the uncertainties were less important and the intended E2 goal referred more to students' realization that scientific investigations by paleontologist contribute to knowledge development, that knowledge is constructed by people. By resembling such investigation it was meant to show this, while at the same time students might experience to be a researcher. This intangible aspect can hardly be assessed. The questionnaire aimed to get an insight on this aspect.

Questionnaire addressing E2 goals. The majority (65% of the HE, n=136 and 76% of the FE, n=29) of both workshops (strongly) agreed with the following statement; '*By this assignment, I now understand more about the research carried out by evolutionary biologists'.*



Only for FE, as this workshop mimics research more closely the following question was included:

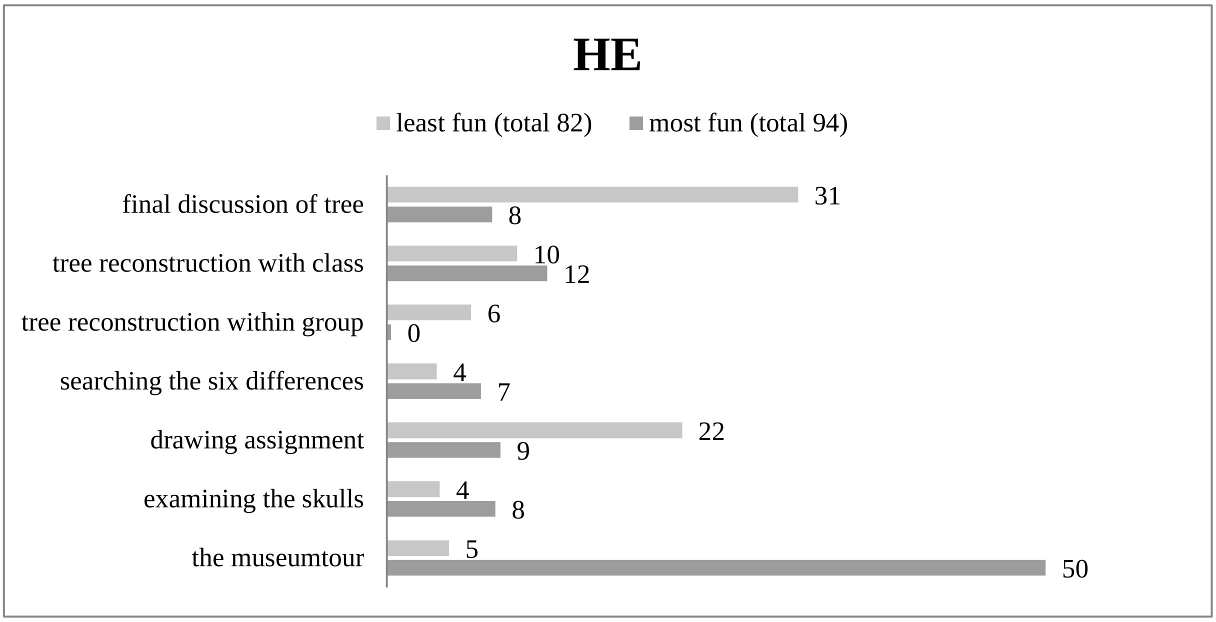
Q: I thought it was challenging to work as a researcher.

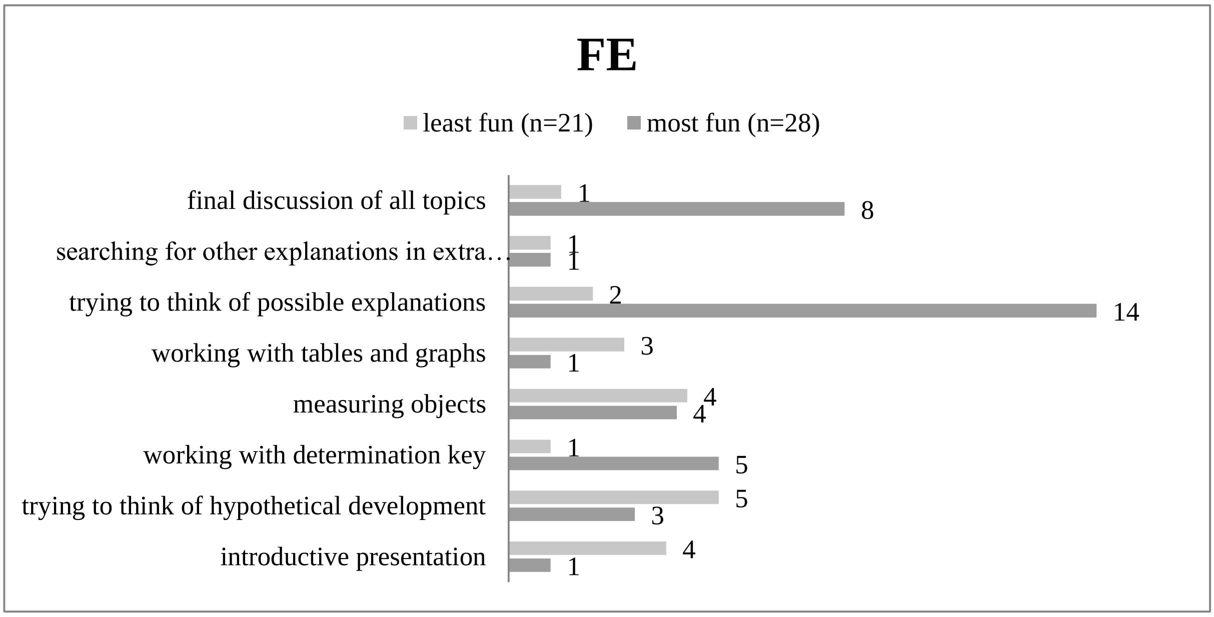
FE: Strongly disagree: 11%, disagree: 39%, agree 43%, strongly agree 7% (n=28)

This seems to correspond to the previous results; slightly more FE students would like to do more research on evolution compared to the HE students. And he FE students would like to do more inquiry on objects, whereas the majority of the HE students don't.

Finally, the students were asked to indicate which parts of the workshop they liked most, and which parts they likes the least. Interestingly, for HE the museum tour at the beginning of the workshop was selected as being the most fun part. For the FE, the museum tour (final discussion of all topics) was rated as second most fun part. Here students indicated the part where they had to come up with explanations for themselves, was rated as most fun.

The two HE parts indicated as the least fun were the final discussion of the tree, and the drawing assignment. For the FE the outcomes were a bit more divided.





6.3. Conclusion attained level

Regarding the fourth research question, 'To what extent are Naturalis' intentions contributing to IB education attained?'

Evaluation of these intentions, related to possible IB goals, considered students' experiences and implicit assessment of learning goals.

For the HE workshop, the questionnaire outcomes on students' experiences gave somewhat striking outcomes. It seems that although the skulls were indicated as special, and the assignment was indicated as not being school-like, motivation, interest and especially challenge scored quite low. The FE was also rated as not challenging, though on several other aspects students' experience seemed more positive compared to the HE workshop. Unexpectedly, the fossils seem to raise more interest than the skulls. And both motivation about doing inquiry on evolution and with fossils as study object gave positive results.

The fact that the HE museum tour was indicated as most fun, is a notable outcome that might be explained in the light of the learning outcomes of the rest of the workshop and students' experience of challenge and motivation. Many learning objectives depended on the group discussions, and it was shown that students were hardly activated to participate, and many learning goals are missed. Indeed generally the students indicated that the group discussions were the least fun part of the workshop. Not being activated could also explain the rating on the drawing assignment, which scored second on least fun part of the workshop, as only one student in each group actually does the drawing.

The part of the FE workshop that was indicated as the most fun, related to the assignment where they had to come up with explanations. This implies that students like to think and reason for themselves, but they need to be incited for this by an explicit question or assignment. The final conclusion and discussion will elaborate on this.

Conceptual learning relating to the image forming about diversity and change seems attained for both workshops. During the FE workshop, learning about the processes behind these phenomena was slightly touched upon, addressing domain specific reasoning to a small extent. However, supervisor support was needed to reason about environmental influences.

Doing science was important for the FE workshop and this seemed generally attained, as all kinds of skills were used and most steps were followed. However, both first and final elements of the inquiry cycle were not implemented. Therefore learning to formulate an hypothesis could not attained for the majority of the students. The same accounts for independently evaluating research findings.

The HE designer stated that students should be rewarded for logical reasoning; this intention seems not attained as reasoning seems to occur hardly at all during the observed HE workshops. Also for FE, the development of general reasoning skills seems hardly supported. Several skills were used by during HE. Only explain, interpret and generalize were oftentimes not used. The attitude accuracy seemed to be supported by the close observations. Other attitudes could hardly be evaluated by this study.

For learning about science, again the attained outcomes for HE highly depend on the supervisor. The majority did address topics that relate to the uncertainties of this field of science. For FE the intention to show how paleontologists work, how knowledge is constructed through inquiry seemed supported; indeed students did discoveries based on their observations and investigation. The majority of the students also stated that because of the workshop they now understand better what paleontologists do.

The following sections will give implications for the results on the three curricular levels. Besides the possible IB goals, contribution to IB education was also related to essential IB education features and the variations. To what extent these were attained will be discussed in the following section.

7. Final Conclusion

It has been indicated that school field trips to museums are often underused, while they could provide great learning possibilities if the unique opportunities of the destination would be used. Naturalis' uniqueness is two sided, it can offer students to work with objects otherwise inaccessible to them and besides this, Naturalis is not just a museum, it is also a research institute. Taking these ingredients together, the museum has great potential to offer valuable education that contributes to in-school science education.

Being strongly affiliated to research, Naturalis aims to use inquiry in the educational programs they offer schools. Inquiry Based education is a widely implemented strategy in formal science education that could be used to deal with a wide range of science learning objectives, though structure and scaffolding are of great importance. To serve this aim, several Inquiry Based workshops were developed. Hereby investigations with unique objects could enhance student motivation while ensuring structure and guidance to support learning. Physical and concrete examples of abstract, evolutionary phenomena can be provided while engaging the students in scientific inquiry.

As Naturalis aims for the IB approach to guide their educational designs, but this has not been structurally evaluated, this study addressed the following research question; '*to what extent do Naturalis workshops designed for school visits contribute to Inquiry Based education?'* Two IB workshops were investigated throughout the three curricular levels (intended, implemented and attained) and evaluated in the light of Inquiry Based learning theories. For the attained level only implicit assessment could be used to indicate what students possibly learned. Three aspects important for defining IB education were examined; possible *goals* for IB education, essential IB education *features* and the *variations* of these features.

IB goals, from intentions to outcomes

As described in the theoretical framework, IB education could be used to serve several goals. These were divided as follows; inquiry 'as a means' for concept learning (M1), 'as a means' to motivate (M2), or 'as a means' to stimulate engagement in sociopolitical action (M3). The inquiry itself could also be the aimed learning outcome, inquiry 'as ends' was divided in doing science (E1; abilities and skills etc.) and learning about science (E2; reflect on science as a way of knowing).

As the workshops intentions were not clearly defined, these were determined in retrospect, and then linked to these possible goals for IB education. It was shown that only 'engaging in socio-political action' was not indicated as a goal by either designer. All others were intended goals to at least a certain extent, with slightly different emphasis for the two workshops. Overall inquiry as a means to motivate was most important for the HE workshop, whereas for FE learning to do science comes first, then motivation and wonder are served by the inquiry as well. The evaluation of the designers intentions for the workshops, allocated to the different IB goals (M1 to E2) will be briefly summarized indicating to what extent the workshops contribute to IB education and how.

M.1. Content learning. Both designers intended to use inquiry as a means to teach content knowledge about evolution and especially to help students to form an image about diversity and change. Both workshop designs seemed to ensure this, and this goal was attained as indicated implicitly; the gradual changes over time and pattern recognition become evident through the sequence of assignments. For HE the potential to learn about the meaning of diversity, a learning goal from the Educational Ruler, highly depends on the supervisor. On implemented level, the actual evolutionary processes were hardly touched upon. It seems that students will mainly get a global idea of the effects of evolution, the different skull forms and features that existed throughout the human evolution line. Indeed the questionnaire data confirmed this, the majority of the students stated that they had learned about the course of human evolution. Form-function reasoning seemed supported by the supervisor in the discussions to a small extent.

For FE, form- function and cause-effect reasoning were elicited in the assignments, addressing the processes behind the diversity and change. Indeed students posed arguments for their observations relating forms to functions (molar shapes to diet). However, for the next step, to reason about the causes of these forms to change over time, relating to dynamic environments, support was needed. Concepts of variation and selection were not explicitly addressed. Almost all students stated that they had learned about the course of evolution.

Another indicated learning objective for both workshop, was for students to learn that not all object features can be used for tree reconstruction. This goal seemed not attained. For FE this was due to a lack of time, the related assignment was not implemented. For HE, it was already not clearly addressed in the design, and the supervisors did not explicitly address it either.

M 2: Motivation. For the HE designer enhancing interest and motivation was the most important goal, for the workshop overall and the inquiry in it. For the FE workshop, doing inquiry had priority, first abilities should be learned to investigate, then relationships can be discovered, then students might start to wonder about these relationships.

Indicators for wonder and curiosity about the objects were only observed during the FE workshop as students started to pose questions about the objects at first confrontation with the molars. Student experience was further investigated with the questionnaire and it this result was generally confirmed; though for both workshops students were not highly impressed and motivated by the objects, results for the FE students were slightly more positive.

The FE designer was skeptical about raising wonder about the objects, though for both workshops the majorities of the students thought it was special to work with the objects, and unexpectedly relatively more students had this opinion about fossils compared to the skulls. Besides this, it seems that FE motivated students for inquiry stronger compared to the HE workshop, while HE students were quite strong in their opinion not being interested in doing more inquiry with the skulls, or on evolution. For FE these questions gave opposite results, the majority of the FE students would like to do more inquiry with fossils, and on evolution.

These outcomes make sense in the light of the following IB goal discussed, the E1 'going science' as the HE has a less investigative character compared to the HE and this could explain the motivation and interest on doing more research.

But first one other important outcome on M2 needs to be pointed out, the results on student experience of challenge. For both workshops, the majorities of the students indicated that the workshops were not challenging. Again the HE students were more pronounced, feeling more strongly in their experience not being challenged. These outcomes could be explained in the light of the M1 goal as it was indicated that conceptual learning was quite superficial. Considering the fact that these upper level secondary education students are already familiar with at least the basics of evolutionary concepts from school, it does not seem surprising that this workshop is not very challenging. Only showing the diversity and change, not explicitly eliciting further thinking about these phenomena does not really level with the current knowledge of the students and therefore students might be easily bored.

E 1. Doing science. Participating in inquiry, doing science was the most important goal for FE. Students were aimed to develop inquiry abilities, skills, and a scientific way of thinking. The designs indeed seemed to resemble scientific research, as most elements of the inquiry cycle are addressed in the assignments and skills need to be used. However, the FE workshop consists of different group assignments, and not all of these included an hypothesis assignment. And during implementation it was also shown that even half of the students with an hypothesis assignment skipped this part. It seemed that the design could have pointed this assignment out more clearly. The final assignments could not be implemented due to a lack of time, also missing an important part of inquiry. In short, it seems that the FE has strong potential for doing inquiry but now some essential aspects are missing.
For HE doing inquiry was less important in terms of skills or following the inquiry cycle. A research question was not even included in the design, although this was indicated as an intention.

For both workshops the development of general, logical reasoning skills was an important intention. It was however indicated that this is not truly supported by either workshop design because the condition for the development of reasoning skills was not met. To support reasoning, this should be explicitly addressed and considered consciously. The designs nor the supervisors did this.

Students of FE were observed posing possible explanations for their observations, relating to form-function thinking. And HE supervisor addressed form-function aspects in several cased as well. Because this is highly domain specific this was related to M1, but more importantly, this is not considered as developing true reasoning skills. Supervisors were generally not actively supporting reasoning by questioning students, posing counterarguments, make students argue and justify their statements. This learning objective seems not attained.

E.2. Learning about science. Using inquiry activities to learn about science and knowledge construction was an intention for both workshops, with slightly different emphasis. During the HE workshop uncertainties about science and current knowledge about human evolution should come to light, not through the assignments but during the discussions. Indeed most supervisor addressed the topic of the changing ideas about the course of evolution.

The worksheet assignments of the FE did not address the nature of science, or seem to leave much room for uncertainties about knowledge construction as the steps and data guided data collection steers upon a certain set outcome. The designers intention was more to show students how paleontologists work, and although some important parts of the inquiry were not implemented, the fact that information about evolution can be gained through investigations on fossils seems generally attained. This was confirmed by the questionnaire data and the majority of the students stated that they now understand more about the work of paleontologists.

IB Features and Variations

Table 18 shows the essential features for IB education from the 'Inquiry and the National Science Education Standards' (Olson & Loucks-Horsley, NRC, 2000). On the intended level the variations of these features were indicated based on the design evaluations. Now it is shown what features are implemented and attained, again by encircling the variations of the features. A cross indicates that the intended feature was not attained.

Table 18

The five essential features and variations for IB education abstracted from 'Inquiry and the National Science Education Standards' (Olson & Loucks-Horsley, NRC, 2000, p. 29). The circles indicate the variation as being attained (dotted line; HE, solid line FE), a cross indicates that the intended feature was not attained.

Essential Feature	**Variations from left to right: more to less learner self-direction, and less to more direction from teacher or material** A	B	C	D
1. Learner engages in scientifically oriented questions	Learner poses a question	Learner selects among questions, poses new questions	Learner sharpens or clarifies question provided by teacher, materials, or other source	Learner engages in question provided by teacher, materials, or other source
2. Learner gives priority to evidence in responding to questions	Learner determines what constitutes evidence and collects it	Learner directed to collect certain data	Learner given data and asked to analyze	Learner given data and told how to analyze
3. Learner formulate explanations from evidence	Learner formulates explanation after summarizing evidence	Learner guided in process of formulating explanations from evidence	Learner given possible ways to use evidence to formulate explanation	Learner provided with evidence and how to use evidence to formulate explanation
4. Learner connects explanations to scientific knowledge	Learner independently examines other resources and forms the links to explanations	Learner directed toward areas and sources of scientific knowledge	Learner given possible connections	*Teacher links to current knowledge*
5. Learner communicates and justifies explanations	Learner forms reasonable and logical argument to communicate explanations	Learner coached in development of communication	Learner provided broad guidelines to use sharpen communication	Learner given steps and procedures for communication

Human Evolution

For the HE workshop, features two, three and four were indicated as present in the designs and the variations were indicated as ranging between the two extreme sides of self-directedness scale.

On the implemented level, indeed the second feature was observed, in the variation as indicated on the intended level. The assignments and actions related to this feature were executed as planned; first students gathered data by describing characteristics of their own two skulls, then the rest of the evidence was provided during the first discussion.

It seems however that feature three, related to formulating explanations and reasoning, hardly occurs on the implemented level. On the intended level, this feature was indicated on the far most self-directed side of the scale, since workshop materials do not guide students in how they should formulate explanations. As described in the attained level section, parts from the HLT where this should occur, where students should discuss, argue and underpin their thoughts (about primitiveness, differences in characteristics and position of skulls in the tree) were oftentimes not executed, or the reasoning was not expressed. Reasoning was also supposed to be addressed and included in the HLT during the group discussions. After the implemented level, it even turned out that most reasoning relied on these discussions. The forms of these discussion differed highly as extensively described in the previous sections. Since student argumentations were hardly observed during the related assignments, if they were even executed, and potential reasoning during group discussions is doubtful and certainly not assured for all groups, this feature was indicated as not implemented or attained.

The fourth feature was indicated on the right side of the scale, as the supervisor instead of the students links the activities and findings to current knowledge. Although supervisors address and emphasize different topics during the final discussion, indeed connections to current knowledge were made in most observed groups.

Fossils and Evolution

Features one to four were indicated as being addressed in the workshop designs. For the first three, the implementation occurred in the same variation as indicated on the intended level.

The ways to gather and analyze the data are provided, and indeed these steps were executed on the implemented level. And indeed students proposed possible explanations for the changes in characteristics they had investigated and organized.

The fourth feature was not observed on the implemented level, as the final assignments were not executed. During these assignments, the students were intended to evaluate their findings, to compare their own results to provided results and extra information. Since there was not enough time, and the extra information files were not available, this feature was not addressed.

Overall this study indicates that FE contributes to IB education to a greater extent compared to the HE workshop. Already on the intended level the FE workshop design addressed more essential IB features. Also on attained level more features and slightly more IB goals are reached. The latter related mainly to doing inquiry, conceptual knowledge and students’ experience on motivation was more positive. Results showed that supporting logical reasoning is currently hardly supported by either workshop.

This study indicated how the potential to reach certain goals decreased when moving from the intended to implemented and attained curricular levels. Already through the first translation from intentions to design, but especially the implementation level showed to highly influence possible learning outcomes. Certainly for the HE workshop, where many learning objectives depend on the supervisor guided discussions. Supervisors seemed to have very different strategies, using different discussion forms. This indicated the implementation sublevel 'perception' as an important factor affecting outcomes for all kinds of intentions.

8. Discussion and Recommendations

Relating the findings of this study to the general aim, to identify the potential of IB workshops for valuable science education, it seems that even though attained outcomes did not reach all intentions, the workshops do seem to have great potential.

The literature study showed that learning in museums mainly addresses certain science curricular goals, namely the informal science learning strands 1, 3 and 5. These were related to inquiry as a means (to motivate), and to inquiry as ends (to doing science), see framework Table 3. It was suggested that the workshops could have the potential to also address the other strands and thereby providing programs that are highly educative and meet with teachers' agendas dealing with congested curricular programs. It seems that indeed, through the inquiry approach both evaluated workshops address content learning, science as a way of knowing, and science identity, related to strands 2, 4 and 6. This shows that the workshops could be of additional value for museum learning, covering learning goals otherwise left unaddressed.

Two important notes should be made here which will lead to several recommendations afterwards.

First, this study showed that intentions for both workshops did relate to all these IB goals and strands for science education, but not all were attained. It was shown here that implementation can highly impede the learning outcomes. On the one hand, the hypothetical learning based on the designs might have been too ambitious. When critically looking at the assignments, the desired student behavior and cognitive learning steps were perhaps unrealistic at certain points. On the other hand, this shows how rich and full of potential learning the workshops might be, if only students and supervisors would perform as planned. This points out where gains can be made, in first instance at the intended (design) and then the implementation levels. To ensure students' activities it seems that assignments could be more explicit, and small adjustments in the designs could already achieve this. Besides this, especially for HE, the supervisors' perception and implementation forms should be considered and needs attention to assure learning outcomes. For HE also time management was an important factor on the implemented level.

Second, and this relates to development of education generally, effective education starts with clearly defined learning objectives. This study attempted to indicate the goals for the workshops, both generally and related to inquiry, in retrospect. And as these were shown to be rich and diverse, they were not clearly defined.

This leads to the general recommendation to predetermine learning goals for the workshops and define these specifically. From the formulation of learning objectives, it should be clear what is expected from the students, whether it concerns cognitive, affective or conative goals. And when formulating goals in terms of conditions, action verbs and criteria, they could be easily assessed.

After this, structurally develop the designs and materials, keeping in mind these goals, and create a sequence of assignments that naturally succeed each other. Producing HLTs could be a useful tool to ensure this. Then, when moving from the intended level to the implemented level, it seems valuable to pay attention to the 'perceived' sublevel, as the designers are not the executive educators. The supervisors should be aware of the intentions of the workshops and how the assignments are supposed to support these learning goals. Updating the workshop scripts with the learning objectives and HLTs with indicated supervisor activities could be very useful.

The next recommendation would be, when delimiting the number of aims, to develop distinctive programs. Choosing for either inquiry 'as a means' or inquiry 'as ends' and support either mainly affective or cognitive learning. Although the two workshops already seem to address inquiry differently, they both currently address all kinds of goals. And it seems that hereby potential outcomes are overshoot. Even as the workshops might have great potential to address different goals, as indicated previously, it seems that when choosing a focus much more can be achieved.

Then this could all be communicated, both internally and externally. By presenting what main goals will be achieved through the workshops and how, teachers can choose which workshops suit best to their goals, resulting in desirable outcomes.

As teachers' agendas have been indicated as an important factor for effective field trips, and they are dealing with accrued planning, it would also be advisable to consult and stay close to the curricular goals for the target group. If of course, content learning is an important goal for the workshop. Then again, it would be useful to communicate clearly; naming specifically which curricular goals are addressed in the different available programs. By these means, teachers can plan connecting activities in school before and after the visit, and this could increase the overall learning outcomes. It would be interesting to investigate what teachers' goals of the currently visiting schools are. These insights could possibly help to choose a focus for the workshops.

When serving a certain target group, offering education on the suitable level does not only please the teacher, but the students certainly as well. This study showed that students' experience of motivation was not all too positive, and they generally did not feel challenged.

Working with the objects was indicated as special though, so there seems to be a good starting point and the rest of the workshop should build on this. One possibility would be, to use the objects and then relate to content about evolution on the curricular level. It was shown that now only a general image of diversity and change is supported by the workshops. While the objects could be a great starting point to dive deeper into evolutionary processes behind this diversity and change. The objects are currently serving to illustrate how features changed over time, whereas at the same time they could be used to show diversity at a certain moment. The combination of these two aspects provides the opportunity to address the evolution concepts variation and natural selection. By this means, these abstract phenomena from the biology book could be shown with examples, using unique objects. This would be a valuable learning moment and while raising the level of addressed content, suitable to current knowledge of the target group, it might positively influence students' experience of motivation and challenge.

Truly using a higher-order reasoning form could be another possibility to increase challenge, again starting with the objects. Indeed the workshop part where students should think and argue for themselves was experienced as the most fun part of the FE workshop. This shows that students like to be challenged to think and explain for themselves. However, they need to be steered and activated through the right questions, and this is what the HE workshop missed. When using reasoning forms as educational approach, the reasoning itself should be addressed explicitly to ensure a conscious process of building up an argument. Specific questions should be posed, and students should justify their ideas taking into account alternative explanations. To support this, supervisors would need to learn how to guide such discussions. Moreover, it should be considered that a large group of students would hardly learn to reason if it all depends on the group discussion with one supervisor, as observed during the HE workshops. Therefore, the worksheets should also explicitly activate the students and elicit reasoning steps.

Besides the special objects, Naturalis other unique quality is the fact that it is a research institute as well. This was indeed indicated as one of the motives to aim for IB education. It seems however, that current inquiry in the workshops in not fully supported. Moreover, it seems that the used level of inquiry does not suit the level of the target group very well. This could also attribute to the explanation of the results on students' challenge and motivation.

It was already indicated on the intended level that in the Educational Ruler, the aspects of 'being a researcher' address primary education pupils. In addition, this study indicated that only several essential inquiry features are attained. Even if doing inquiry itself is not the main goal, it was indicated in the theoretical framework that the use of inquiry could support all kinds of science learning goals as long as suitable variations are chosen. When contributing to IB education to a greater extent by addressing more essential features, or using somewhat more self-directed variations in accordance with the intentions, the inquiry level might suit the target group better and students might feel challenged more. At the same time, it would suit Naturalis' authentic environment and research identity, thereby offering students a unique learning experience.

Besides the innovational recommendations on the workshop designs and implementations, the final advise relates to future evaluations.

Educational practices should be evaluated continuously; the designs will improve through cycles of redesign, implementation and assessments. This study now performed the assessment of two workshops, implying what could be improved. As it was pointed out that many intentions were currently not attained, it seems advisable to investigate the other workshops as well.

After redesign, followed by evaluation on intended and implemented levels again, it would also be valuable for future studies to assess learning objectives on the attained level more extensively. Implicit assessment as currently executed could be a starting point. Then, explicit assessment on actual learning outcomes would probably shed new light on the workshops and how these could be improved.

As the results on students' experiences were somewhat different from what was aimed and hoped for, and raising wonder has such high priority, it would be interesting to keep on tracking this as well.

This study showed how structural evaluations provide insights on reaching the objectives; especially which intentions are not attained and why. When such critical evaluations are structurally performed and succeeded by indicated adjustments, it seems that Naturalis could indeed offer truly effective Inquiry Based- workshops.

References

Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., ... & Tuan, H. L. (2004). Inquiry in science education: International perspectives. *Science Education*, 88(3), 397-419.

Abell, S. K. & Lederman, N. G. (Eds.). (2007). *Handbook of research on science education.* Psychology Press.

Alfieri, L., Brooks, P. J., Aldrich, N. J., &Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, 103(1), 1.

Bakker, A., & Van Eerde, H. A. A. (in press). An introduction to designbased research with an example from statistics education. In A. Bikner- Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Doing qualitative research: methodology and methods in mathematics education.* New York: Springer.

Bell, P., Lewenstein, B., Shouse, A. W., &Feder, M. A. (Eds.). (2009).*Learning science in informal environments: People, places, and pursuits.* National Academies Press.

Boerwinkel, D.J., 2003. *Het vormfunctieperspectief als leerdoel van natuuronderwijs: leren kijken door de ontwerpersbril.* Academisch Proefschrift: Universiteit Utrecht.

Braund, M., &Reiss, M. J. (Eds.). (2004*). Learning science outside the classroom*. Psychology Press.

Bybee, R. W. (2004). Scientific inquiry and science teaching. In: Flick, L. B., & Lederman, N. G. (Eds.). (2004). *Scientific inquiry and nature of science: Implications for teaching, learning, and teacher education* (Vol. 25). Springer.

Coburn, W.W.,Schuster, D., Adams, B., Skjold, B., Applegate, B., Undreiu, A., Loving, C.C., Gobert., J.D., (2009) Inquiry instruction or direct? Experiment-based answers to what practices best promote conceptual development of significant science content. In M.F. Taşar& G. Çakmakcı (Eds.), *Contemporary science education research: teaching* (pp. 283-291). Ankara, Turkey: PegemAkademi.

Colburn, A. (2000). An inquiry primer. *Science Scope*, 23(6), 42-44.

Cox-Petersen, A. M., Marsh, D., Kisiel, J., Melber, L.M. (2003)Investigation of guided school tours, student learning, and science reform recommendations at a museum of natural history, *Journal of research in science teaching,*40-2, 200

Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of research in science teaching*, 44(4), 613-642.

DeWitt, J., & Storksdieck, M. (2008). A short review of school field trips: Key findings from the past and implications for the future. *Visitor Studies*, 11(2), 181-197.

Duschl, R. A. (2004). International prespectives on inquiry in science education: a commentary. In: Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok- Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., ... & Tuan, H. L. (2004). *Inquiry in science education: International perspectives. Science education*, 88(3), 397-419..

Falk, J. H., & Dierking, L. D. (2012). *Museum Experience Revisited*. Left Coast Press

Falk, J. J. H. (1992). *The museum experience.* Howells House.

Flick, L. B., &Lederman, N. G. (Eds.). (2004). *Scientific inquiry and nature of science: Implications for teaching, learning, and teacher education* (Vol. 25). Springer.

Furtak, E. M. (2006). The problem with answers: An exploration of guided scientific inquiry teaching. *Science Education*, 90(3), 453.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), 99-107.

Hodson, D. (2003). Time for action: Science education for an alternative future. *International Journal of Science Education*, 25(6), 645.

Janssen, F., & de Hullu, E. (2008). A toolkit for stimulating productive thinking. *Journal of Biological Education*, 43(1), 21-26.

Justice, C., Rice, J., Warry, W., Inglis, S., Miller, S., &Sammon, S. (2007). Inquiry in higher education: Reflections and directions on course design and teaching methods. *Innovative Higher Education,* 31(4), 201-214.

Justice, C., Warry, W., Cuneo, C., Inglis, S., Miller, S., Rice, J., & Sammon, S. (2001). *A grammar for inquiry: Linking goals and methods in a collaboratively taught social sciences inquiry course.* The Alan Blizzard Award.

Kemmers, P., Klein Tank, M., van Graft, M., (2007) *Onderzoekend en Ontwerpend Leren bij Natuur en Techniek Evalueren van brede ontwikkeling van leerlingen in open onderwijsvormen,* SLO, Enschede

Khasnabis, D. (2008). *Developing Scientific Literacy Through Classroom Instruction: Investigating Learning Opportunities Across Three Modes of Inquiry-based Science Instruction*. ProQuest.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, 41(2), 75-86.

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of research in science teaching*,29(4), 331-359.

Luft, J., Bell, R. L., & Gess-Newsome, J. (Eds.). (2008). *Science as inquiry in the secondary setting.* NSTA Press.

Lunetta, V.N., Hofstein, A. & Clough, M. (2007). Learning and teaching in the school science laboratory: An analysis of research, theory and practice. In: N. Lederman, & S. Abell (Eds.), *Handbook of research on science education* (pp. 93–441). Mahwah, NJ: Lawrence Erlbaum

Mark R. Wilson & Meryl W. Bertenthal (Eds.) (2005) *Systems for State Science Assessment. Committee on Test Design for K-12 Science Achievement, Board on Testing and Assessment, Board on Science Education, Division of Behavioral and Social Sciences and Education, National Research Council.* National Academies Press

Martin A. Simon & Ron Tzur (2004) Explicating the Role of Mathematical Tasks in Conceptual Learning: An Elaboration of the Hypothetical Learning Trajectory, *Mathematical Thinking and Learning*, 6:2, 91-104

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist*, 59(1), 14-19.

Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing research*, 40(2), 120-123.

Olson, S., &Loucks-Horsley, S. (Eds.) (2000). *Inquiry and the National Science Education Standards: A guide for teaching and learning.* National Academies Press.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: a review of the literature and its implications. *International journal of science education*, 25(9), 1049-1079.

Otto, K. N., & Wood, K. L. (2001). Product design: techniques in reverse engineering and new product development. 清华大学出版社.

Prince, M. J., & Felder, R. M. (2006). Inductive teaching and learning methods: Definitions, comparisons, and research bases. *Journal of Engineering Education*, 95(2), 123-138.

Quinn, H., Schweingruber, H., & Keller, T. (Eds.). (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. National Academies Press.*

Rennie, L. (2007) Learning Science Outside of School. In: Abell, S. K. & Lederman, N. G. (Eds.). (2007). *Handbook of research on science education*. Psychology Press.

Rennie, L., &McClafferty, T. (1995). Using visits to interactive science and technology centers, museums, aquaria, and zoos to promote learning in science. *Journal of Science Teacher Education,* 6(4), 175-185.

Rijst, R.M. van der (2009). *Exploring scientific research disposition from the perspective of academics in The research‐teaching nexus in the sciences: Scientific research dispositions and teaching practice. Published dissertation.* ICLON – Leiden University Graduate School of Teaching, The Netherlands.

Schmidt, H. G., Loyens, S. M., Van Gog, T., & Paas, F. (2007). Problem-based learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, 42(2), 91-97.

Schwartz, R. S., Lederman, N. G., & Crawford, B. A. (2004). Developing views of nature of science in an authentic context: An explicit approach to bridging the gap between nature of science and scientific inquiry. *Science education,* 88(4), 610-645.

Schwartz, R.S., & Crawford, B.A., 2004. Authentic Scientific Inquiry as Context for Teaching Nature of Science: Identifying Critical Elements for Success. In: Flick, L. B., & Lederman, N. G. (Eds.). (2004). *Scientific inquiry and nature of science: Implications for teaching, learning, and teacher education* (Vol. 25). Springer.

Spek, W. & Rodenboog-Hamelink, M., (2011). Natuurwetenschappelijke vaardigheden onderbouw havo-vwo, SLO, Enschede

Spronken-Smith, R., & Walker, R. (2010). Can inquiry-based learning strengthen the links between teaching and disciplinary research? *Studies in Higher Education*, 35(6), 723-740.

Spronken-Smith, R., Walker, R.., Batchelor, J., O'Steen, B. and Tom Angelo (2012) Evaluating student perceptions of learning processes and intended learning outcomes under inquiry approaches, *Assessment & Evaluation in Higher Education*, 37:1, 57-72,

The PRIMAS project: Promoting inquirybased learning (IBL) in mathematics and science education across Europe, Deliverable D.4.1 Version, March 31st, 2011

Thijs, A. M &Akker J. (2009) Curriculum in development, SLO - Netherlands Institute for curriculum development

Uitto, A., Juuti, K., Lavonen, J., &Meisalo, V. (2006). Students' interest in biology and their out-of-school experiences. *Journal of BiologicalEducation,* 40(3), 124-129.

Van der Rijst, R. M., Visser-Wijnveen, G. J., Verloop, N., & Van Driel, J. H. (2013). Undergraduate science coursework: teachers' goal statements and how students experience research. *Innovations in Educationand Teaching International,* 50(2), 178-190.

Van Graft,M., & Kemmers, P., (2007), Onderzoekend en Ontwerpend Leren bij Natuur en Techniek: Basisdocument over de didactiek voor onderzoekend en ontwerpend leren in het primair onderwijs. SLO, Den Haag: VTB.

Venville, G. J., & Dawson, V. M. (2010). The impact of a classroom intervention on grade 10 students' argumentation skills, informal reasoning, and conceptual understanding of science. *Journal of Research in Science Teaching*, 47(8), 952-977.

Appendix I: Educational Principle ruler by Van Keulen (2011)

Meetlat Educatieve Principes Naturalis

1. Zelf onderzoeker zijn

De natuurlijke werkelijkheid kennen, beschrijven en begrijpen is mensenwerk, gebaseerd op onderzoek. Kinderen ontwikkelen een beter begrip van de ‘feiten’, de zekerheden en de onzekerheden, wanneer ze zelf als onderzoeker aan de slag kunnen gaan en keuzes moeten maken en verantwoorden op basis van het materiaal dat voorhanden is.

1. Er is een (onderzoeks)vraag.
2. De onderzoekscyclus wordt gevolgd (van vraag naar zoekproces naar resultaat).
3. Een scala aan vaardigheden wordt ingezet: vraag formuleren; voorspellen; onderzoeksplan opstellen; waarnemen; ontwerpen; gegevens verzamelen; logboek bijhouden; interpreteren; modelleren; visualiseren; verklaren; generaliseren; toepassen; onder woorden brengen; evalueren; presenteren).
4. Er wordt een beroep gedaan op hogere-orde redeneervormen (overeenkomsten-verschillen; patroonherkenning; oorzaak-gevolg; vorm-functie; doel-middelen; eenheid-verscheidenheid; deel-geheel; verandering-continuïteit; product-omgeving; uniek-algemeen; toevallig-regelmatig).
5. Een onderzoekende houding wordt gecultiveerd (nieuwsgierigheid; openheid; creativiteit; vasthoudendheid; nauwkeurigheid).
6. Er is ruimte voor onzekerheid.

2. De echte dingen ervaren

Directe, rijke zintuiglijke ervaring van ‘de echte dingen’, in een betekenisvolle context, zonder tussenkomst van hulpmiddelen zoals meetinstrumenten, modellen, afbeeldingen, teksten of video, is een indringende ervaring die de aandacht van kinderen richt en vragen oproept.

1. Er wordt gewerkt met echte objecten.
2. De objecten roepen vragen op en kunnen de aandacht richten.
3. Er zijn rijke en gevarieerde mogelijkheden voor directe zintuiglijke ervaringen (zien, voelen, horen, ruiken).
4. Andere elementen en hulpmiddelen (replica’s; modelvoorstellingen; tekstbordjes; video’s; waarnemingsinstrumenten) helpen om de aandacht te richten.
5. De objecten staan in een authentieke context die kinderen helpt hun ervaringen te duiden en betekenis te geven.
6. De objecten hebben een verhaal dat ontdekt en verteld kan worden.

3. Verwondering wekken

Er zijn fenomenen en objecten die niet goed passen bij wat kinderen al weten of verwachten. Als kinderen zich daar bewust van worden is de basis voor een leerproces gelegd. Verwondering leidt tot nieuwsgierigheid en tot ‘willen weten’. Soms dringt een fenomeen zich gemakkelijk op, zoals vuurwerk; soms moet verwondering gewekt worden door de aandacht te richten.

Op verwondering moet voortgebouwd kunnen worden: onbegrijpelijke fenomenen, hoe mooi ook, zetten geen voortschrijdend proces van interpretatie en betekenisgeving in gang maar kunnen juist aanleiding geven het denken uit te schakelen: “Ik snap hier niets van; dit is dus niks voor mij”.

1. Kinderen worden geconfronteerd met objecten of fenomenen waarvoor ze zich kunnen verwonderen.
2. De aandacht wordt gericht op facetten die nieuwsgierig maken en waar op voortgebouwd kan worden.
3. De natuurlijke nieuwsgierigheid van kinderen wordt benut.
4. De confrontatie leidt tot nadenken en vragen van kinderen: Wat is dit? Hoe kan dit? Ik wil dit weten!
5. Deze vragen zetten een zoekproces in gang waardoor de kinderen tot meer kennis en inzicht komen.

4. Leren over diversiteit en verandering

Naturalis is een groot en toegankelijk archief met een inhoudelijke focus op de geologische processen die de aarde vormen en vormden, over de aanpassingen van soorten op de veranderende aarde en over hoe dieren en planten overleven in hun dagelijkse strijd om het bestaan. Door dit archief te ontsluiten ten behoeve van het onderwijs kan Naturalis scholen helpen leerdoelen te realiseren. Naturalis benut daarbij de mogelijkheden van informeel, buitenschools leren

kan de stamboom van het leven worden gereconstrueerd, gewaardeerd en begrepen.

1. Kinderen zijn er op voorbereid dat Naturalis zich richt op diversiteit en verandering en ervaren dat in Naturalis voortdurend.
2. Naturalis wordt door kinderen ervaren als een vorm van buitenschools, informeel leren.

3. De aandacht wordt gericht op verschillen en de oorsprong en betekenis van verschillen.
4. Er is aandacht voor de dynamiek in de natuurlijke omgeving van de objecten.
5. Kinderen vormen zich een beeld en een mening over diversiteit en verandering.

5. Exemplarisch werken

De aandacht vasthouden volgt op aandacht krijgen. Het gevaar bestaat dat leerlingen in Naturalis door de veelheid van objecten afgeleid worden voordat ze tot de kern zijn doorgedrongen. Het is verstandig om prototypische voorbeelden te selecteren die passen bij het inhoudelijke doel, het voorstellingsvermogen en de ervaringshorizon van de leerlingen, en deze voorbeelden grondig uit te werken. “Niet het vele is goed maar het goede is veel”.

1. De meeste tijd wordt besteed aan een beperkt aantal voorbeelden.
2. De voorbeelden sluiten aan bij de voorkennis en de belevingswereld van kinderen.
3. De voorbeelden zijn gemakkelijk te ontsluiten voor directe zintuiglijke ervaring.
4. De voorbeelden zijn prototypisch: ze staan voor een grotere groep van objecten of fenomenen.
5. Algemenere uitspraken of conclusies worden opgebouwd aan de hand van deze voorbeelden.

6. Hart, hoofd en handen

Als objecten de aandacht krijgen die ze verdienen, dan ontsluiten ze zich. Ze zetten kinderen niet alleen aan het denken maar ook aan het handelen. Deze denk- en doe-activiteiten leiden tot waardering. Het is belangrijk niet alleen de nadruk te leggen op de cognitieve elementen maar evenwichtig alle faciliteiten aan te spreken. Objecten of fenomenen zijn mooi, of spannend, of gek, of interessant, of nog iets anders. Je bouwt er een band mee op. Je kunt je identificeren en je verbonden weten. Dit kan onbewust gebeuren maar het kan ook onderwerp van gesprek zijn.

1. Kinderen worden uitgedaagd na te denken.
2. Er is ruime gelegenheid om te handelen.
3. Kinderen krijgen ruimte om de objecten of fenomenen te waarderen en dit te uiten.
4. Kinderen kunnen zich persoonlijk identificeren met de verhalen of objecten.

7. Rekening houden met verschillen

De ontwikkeling van kinderen is deels biologisch bepaald. Tegelijk is ieder kind anders. Zo kunnen er binnen een groep kinderen van dezelfde leeftijd grote verschillen zijn in aanleg, ontwikkelingstempo, interesse, persoonlijkheid en wat betreft ervaring en voorkennis. Leren is effectiever wanneer leerlingen worden aangesproken op hun eigen ontwikkelingsniveau. Kinderen die veel meer weten dan hun groepsgenoten kunnen je op het verkeerde been zetten en te hoge verwachtingen van de groep oproepen, maar ze kunnen ook de anderen op sleeptouw nemen. Ze stellen, gewapend met hun voorkennis, betere vragen en ze kunnen hun waarnemingen gemakkelijker interpreteren en onder woorden brengen. Tegelijk spreken zich nog de taal van hun leeftijdsgenoten.

1. Er is rekening gehouden met de interesse en leefwereld van zowel jongens als meisjes.
2. Het verschil in talent (handigheid, leesvaardigheid, redeneervaardigheid, zelfsturing) wordt productief gemaakt.
3. Er is rekening gehouden met verschil in interesse en ‘bètamentality’:
   - Concrete bèta’s zijn nieuwsgierig en willen graag hun handen gebruiken.
   - Carrièrebèta’s interesseren zich voor statusgevoelige aspecten (wat kun je er mee bereiken?).
   - Mensgerichte generalisten willen zich graag nuttig maken.
   - Non-bèta’s willen weten wat dit voor hun dagelijks leven betekent.

7.1 Leeftijdsgebonden aspecten: Basisschool onderbouw

Voor kinderen in de onderbouw van het basisonderwijs is nog niet helemaal duidelijk wat kan en wat niet kan in de natuurlijke wereld. Daarom zijn ze zintuiglijk en exploratief ingesteld. Ze gedragen zich als groep. Ze zijn niet met zichzelf bezig en volgen geen vooropgezet, individueel plan. Ze zijn impulsief: er staat nog geen rem op hun gedrag en ze denken niet na over de consequenties. Ze zijn makkelijk afgeleid door weer nieuwe indrukken. Ook zijn hun fijn-motorische vaardigheden nog niet volop ontwikkeld: complexe handelingen zijn lastig.

Ze houden van spel en toneel. Ze identificeren zich met rondleiders en workshopleiders. Ze kiezen niet voor het ‘waarnemer’ perspectief maar verplaatsen zich in het verhaal, in de hoofdrolspelers, en gebruiken hun fantasie om het verhaal aan te vullen. Ze willen ook graag hun eigen verhaal vertellen.

1. Begeleiders houden rekening met spanning (plan WC-bezoek).
2. Er zijn verrassende en wonderlijke objecten die de fantasie prikkelen.
3. Er worden verhalen verteld waarin de kinderen zich verplaatsen.
4. Er wordt rekening gehouden met een korte spanningsboog.
5. De kinderen kunnen volop bewegen.

6. Er is veel ruimte voor zintuiglijke exploratie. Kinderen mogen met hun handen aan voorwerpen zitten.
7. Kinderen kunnen hun verhaal doen en zelf ervaringen onder woorden brengen.
8. Opdrachten gaan over overeenkomsten en verschillen, zodat kinderen gaan generaliseren en abstraheren.
9. Kinderen kunnen worden geholpen met lezen en fijn-motorische opdrachten

7.2 Leeftijdsgebonden aspecten: Basisschool bovenbouw

Kinderen uit de bovenbouw van het basisonderwijs zijn geestelijk en lichamelijk goed in balans. Ze voelen zich groot en gedragen zich ongedwongen en zelfverzekerd. Als groep kennen ze elkaar en zijn goed op elkaar ingespeeld. Wie de ongeschreven regels van de groep niet kent of navolgt ligt er uit en loopt het risico te worden gepest. Deze kinderen staan erg open en vinden alles leuk. Ze zijn nieuwsgierig en stellen vragen. Ze gaan voor kennis, feitjes, weetjes en records. Ze houden van actie en wedstrijdelementen worden op prijs gesteld. Als het ‘echt’ is vinden ze het spannend. Ze kunnen zich concentreren op een opdracht en die ook afronden. Hun zelfvertrouwen vertaalt zich naar de wens om beloond te worden voor prestaties. Ze accepteren de spelregels en doen wat de bedoeling is. Ze kunnen helemaal in de flow komen en opgaan in waar ze mee bezig zijn.

1. Er is veel aandacht voor feitelijke kennis.
2. Opdrachten hebben elementen van spel en competitie.
3. Opdrachten leiden tot een produkt dat kinderen ‘af’ kunnen maken en waar ze inhoudelijke feedback op krijgen.
4. Kinderen worden gestimuleerd om vragen te stellen.
5. De kinderen worden als groep benaderd.
6. Er wordt een groter beroep gedaan op de zelfsturing van kinderen (luisteren, concentreren, niet rennen, opdracht afmaken).

7.3 *Leeftijdsgebonden aspecten: Voortgezet Onderwijs onderbouw*

Pubers zijn veel met zichzelf bezig en hebben zekerheid nodig. Ze zijn gevoelig voor wat anderen van hen vinden. Ze zijn geïnteresseerd in alles wat met sexualiteit en met het andere geslacht te maken heeft. Ze willen zelfstandigheid maar weten daar nog niet goed vorm aan te geven. Impulsen worden op deze leeftijd nog niet goed onderdrukt of aangepast aan de omstandigheden. Er is altijd gedoe maar het heeft meestal geen diepe betekenis. Ze willen duidelijkheid over of ze het goed doen. Ze willen haalbare opdrachten en bevestiging dat ze het goed gedaan hebben. Ze zijn te boeien maar tonen dit niet uitbundig. Ze zijn voorzichtig met het stellen van vragen. Vragen hebben soms het karakter van ‘uitproberen’ zodat ze zich in de groep kunnen profileren. Met name vmbo’ers waarderen directheid en spontaniteit.

1. Het is de leerlingen heel duidelijk wat van hen verwacht wordt.
2. De grenzen en spelregels worden duidelijk aangegeven.
3. De opdrachten zijn haalbaar en concreet.
4. De leerlingen worden individueel aangesproken.
5. Niet alles is met de hele groep; er is ruimte om je af te zonderen.
6. Er is aandacht voor ‘mannelijk’ en ‘vrouwelijk’ in de natuur.
7. Er is ruimte voor discussie.

7.4 Leeftijdsgebonden aspecten: Voortgezet Onderwijs bovenbouw

Bij oudere leerlingen spelen schoolsoort en profielkeuze mee. Ze zijn niet meer zo bezig met de groep en beginnen zich meer als individu te ontplooien. Een aantal VWO-ers wil echt graag leren en uitgedaagd worden. Voor de meeste kinderen geldt dat het bezoek aan Naturalis nuttig en belonend moet zijn.

In vergelijking met jongere kinderen zijn ze minder met de zintuigen bezig en meer met cognitieve vaardigheden zoals bedenken, redeneren en discussiëren. Ze maken mentale voorstellingen van de wereld door zich in te beelden hoe iets kan zijn. Ze hebben voorkennis en kunnen concepten als diversiteit in een kader plaatsen. Uitvoerende en controlefuncties zoals zelf doelen stellen en daarop evalueren zijn nog niet volledig ontwikkeld. Ook is het vermogen om open problemen op te lossen nog klein, zeker wanneer dit planning en aansturing van het eigen gedrag inhoudt.

1. De opdrachten zijn uitdagend en niet schools.
2. De opdrachten dragen bij aan schoolcijfers.
3. Het is nodig om te redeneren en mentale voorstellingen te ontwikkelen.
4. Er zijn ook opdrachten met minder structuur, zodat leerlingen kunnen oefenen met zelfsturing.

8. Sociale interactie

Kennis en inzicht zijn gegrond in confrontatie en ervaring met de natuurlijke wereld. Zonder de noodzaak ervaringen onder woorden te brengen blijven veel inzichten intuïtief en ‘tacit’. Sociale interactie lokt verwoording uit. De gesprekken die zo ontstaan dragen bij aan beter begrip en betere verankering. Ook draagt sociale interactie bij aan activiteiten die voorwaardelijk zijn voor ervaring en inzicht, zoals plannen, samenwerken, het verdelen en uitvoeren van taken, en het doen en delen van observaties.

1. Er wordt rekening gehouden met de groepsgrootte.
2. Er is gelegenheid of noodzaak tot samenwerken en overleggen.
3. Er is interactie met medewerkers / onderzoekers van Naturalis.
4. Kinderen worden uitgedaagd hun verhalen, ervaringen en meningen onder woorden te brengen.

9. Begeleiders een rol geven

Onderzoekend leren, en de nieuwsgierigheid van kinderen benutten, betekent niet dat kinderen zonder hulp of begeleiding met hun vraag aan de slag gaan. Begeleiding of coaching door een volwassene is belangrijk en meestal noodzakelijk. Volwassenen kunnen kinderen helpen hun bedoelingen of waarnemingen beter onder woorden te brengen. Ze kunnen vragen stellen en doorvragen en zo kinderen op ideeën brengen. Ze kunnen kinderen bevestigen en feedback geven. Ze kunnen kinderen indien nodig bijsturen.

1. De begeleiders vanuit de school hebben een duidelijke rol in het leerproces van de kinderen.
2. Er zijn tussenmomenten waarin kinderen gevraagd wordt wat ze tot nu toe gedaan en bereikt hebben.
3. Kinderen worden met vragen en doorvragen gehoplen hun ervaringen te verwoorden, hun resultaten te evalueren en op hun conclusies te reflecteren.

10. Afstemmen met de school

Een bezoek aan Naturalis is niet alleen een leuk uitje maar draagt bij aan het leren van kinderen. Het is een verantwoordelijkheid van de school de bijdrage van Naturalis aan het bereiken van leerresultaten in te passen in het programma, kinderen adequaat voor te bereiden en na het bezoek de ervaringen te laten beklijven. Hiervoor is overleg en afstemming nodig om wat Naturalis kan bieden optimaal te laten aansluiten bij wat de school wil bereiken. De leerdoelen van de school kunnen inhoudelijk zijn uit het domein Natuur, Biologie of NLT, of meer algemeen, zoals het ontwikkelen van een onderzoekende houding of leren kritisch redeneren.

1. Het bezoek aan Naturalis staat in het teken van het bereiken van leerresultaten.
2. Er zijn afspraken met de school over de rol en taken van de begeleiders (leraren, ouders) die meekomen.
3. De school bereidt het bezoek voor.
4. De school formuleert leerdoelen en evalueert na afloop in hoeverre deze bereikt zijn.

Appendix II: Educational Ruler related to IB goals – Underpinning

Underpinning for IB identification as stated in Table 3 and Table 9 with quotes from Van Keulens documents; the Educational Ruler itself and the 'Basic Document IB Naturalis' (2011)

Principle from Educational Ruler	Inquiry goal
1. Being a researcher	E1/E2/M1
1. There is a research question	E1
2. The inquiry cycle is followed	E1
3. The following skills are used ….	E1
4. Use higher order reasoning forms …	E1-M1
5. The cultivation of inquisitive attitudes' …	E1
6. There is room for uncertainty	E2
2. Experience real things	M2
3. To raise wonder	M1/M2
4. Learning about diversity and change	M1
5. Exemplary work'	x
6. Heart, head, hands'	M2
7. Take differences into account'	x
8. Social interaction'	M1, E1
9. Give supervisors a role	x
10. Agreements with the school	x

1. 'Being a researcher' and all six aspects in the Educational Ruler relate to the use of inquiry as ends. More specifically, the first five aspects relate to E1, to 'doing science'. The aspects from 1.1 to 1.5 that are named in the ruler address activities, skills, conditions, abilities, attitudes all relating to scientific inquiry as a learning outcome.

The sixth aspect, 'there is room for uncertainty' relates to E2, as this point refers to students' understandings about science as a way of knowing. Indeed in the Basic Document IB Naturalis' by Van Keulen (2011) it is stated that "knowing, describing, and understanding the natural world is based on research developed by people" (p. 6). It is not described in the Educational Ruler how this condition 'there is room for uncertainty' should or could be addressed besides participating in inquiry and experience how you should make choices and justify explanations based on materials.

Five out of six aspects in the Educational Ruler are formulated in a way that they contribute to the goal doing scientific research (E1), focusing on the appropriate conditions for inquiry. In the 'Basic Document IB Naturalis' by Van Keulen (2011) Inquiry Based learning is described as an educational approach using the natural curiosity of children. It is stated that "inquiry learning develops skills that are used by real scientists"… "also attitudes and other talents could be developed such as curiosity, accuracy, tenacity, openness, and creativity." (p. 4). These aspects relate to E1. The use of inquiry learning and the aspects in the ruler of principle correspond, and refer to inquiry as ends (E1).

However, the description on the first principle 'being a researcher' in the 'Basic Document IB Naturalis' by Van Keulen (2011) seems to refer mainly to the use of inquiry for other goals. It is stated that: "Students can develop a better understanding of various biological and geological concepts and the certainties and uncertainties, if they themselves work as a researcher." (p. 6). This refers to the use of inquiry as a means for content learning (M1) as well as to learning about science (E2). Indeed the rest of the description emphasizes understandings about science as a way of knowing, knowledge construction and recognizing that knowledge is constructed by people (E2); "Knowledge is constructed and is fundamentally uncertain and colored by the perspective of the researcher"… "When students can see scientists in action and participate in real research, a better understanding will develop on the status of the knowledge we have about the world."(p. 6)

2. 'Experience real things' is a principle that in its description seems to have a contradiction. It is stated that "Direct, rich sensory experience of 'real things' in a meaningful context, without tools such as measuring instruments, models, images, text or video, is an intense experience that focuses the attention of children and raises questions"(ruler). Then, the fourth aspect of this principle is: " Other elements and tools (replicas, model

representations, text panels, videos, sensors) help to focus attention" (ruler). The latter citation could be strongly supported by inquiry. The first citation could be impeded by inquiry if there is not enough time for the first exploration phase of the research cycle. If inquiry is used, while meeting with this principle, it seems that the focus lies on inquiry as means, to motivate active engagement for the subject, M2.

3. 'To raise wonder', is a principle that emphasizes stimulation of the natural curiosity to start the learning process: "Wonder leads to curiosity and 'want to know'"(ruler). Three aspects of this principle describe conditions to raise wonder. Inquiry could be used to meet with these aspects, the goal then relates to M2, motivation. The two other aspects of this principle relate to content learning, as the purpose of the raised wonder. It is stated that the need to know eventually should lead to increased knowledge and deeper insights, about these objects or phenomena (M1).

4. 'Learning about diversity and change', if inquiry is used to meet with this principle it would be for content learning, M1. It is stated that Naturalis could support school learning objectives related to geological processes, and evolution;" the tree of life can be reconstructed, appreciated and understood" (ruler). The fourth aspect relates to opinion forming about diversity and change, this could be a first step in feeling committed to take socio-political action, however the Van Keulens documents do not mention any relation to matters of social, economic, environmental and moral-ethical concern.

5. 'Exemplary work' relates to conditions to ensure students' focus. This is relevant for learning in general, and accounts for all kinds of learning objectives. The inquiry approach could support this principle, as during an investigation only a number of objects is investigated and a focus is ensured. Inquiry used for the purpose of focus, could still be used for learning science, motivation, learning skills, attitudes, and learning about science; no specific IB goal could be linked to this principle

6. 'Heart, head, hands' relates to the condition of appreciation for cognitive and affective learning aspects. This principle does not specifically relate to inquiry, but inquiry could be used as a way to meet with this principle related to appreciation. Then it would be used as means, to enhance excitement and motivation, this principle is indicated as M2.

7. 'Take differences into account' relates to conditions for effective activities in general, regardless the goal for inquiry

8. 'Social interaction', cooperation and verbalization are named as important for two aspects; critical reflection on experience, and better understanding and better anchoring of knowledge. The first relates to inquiry as ends, specifically to 'doing science' and the accompanied cooperation, using scientific language and discourse (E1). The second relates to inquiry as means for content learning, 'learning science', for real and anchored understanding of certain concepts (M1).

9. 'Give supervisors a role' related to conditions for effective activities, and does not relate to inquiry or its possible goal.

10. 'Agreements with the school' relates to the fact that the goal for the visit needs to be determined by the teachers, and communication about this goal is an important condition for desired outcomes. This principle is highly supported by this study, and perhaps should be the first principle on the list.

Appendix III: Design check for principle ruler aspects - Underpinning

The HE and FE designs were analyzed to determine whether the educational principles were being addressed. Here the criteria for the principle aspects are provided.
Conditional principles could be checked as being present or not. Aspects related to phases of research could be checked as being explicitly named, or implicitly addressed. Some principles' elements relate to highly intangible aspects. For some elements, the materials' potential to support or contribute to these goals was be determined.

Principle:
1. Being a researcher: most elements relate to conditions that could be checked as (explicitly/implicitly) present.
- 'There is a research question', this condition could be implicitly present, a research question could be given and explicitly named, or students could be asked to formulate a research question.
- 'The inquiry cycle is followed' refers to the representation of scientific work and the ruler states 'from question to search process, to results' and the design was evaluated on these aspects. If this path is followed without naming it specifically it was selected as being implicitly present
- The use of skills, could be explicitly named and addressed through certain assignments. Implicitly addressed skills are not named in the assignment, but they are indirectly used. Modelling Evaluating was interpreted as the reflection on the process and results.
- In relation to the 'higher-order reasoning forms' the material was analyzed as potentially supporting this form of reasoning. This does not proclaim in any way that these reasoning forms will actually be developed.
- 'The cultivation of inquisitive attitudes', was not related to the material, as they relate to behavioral aspects mainly. The investigation of environmental and material conditions to support such attitudes would be a study on its own. Data from the observations will be used to gain insight on students' research dispositions as described by Van Rijst (2009).
- 'There is room for uncertainty' is a condition relating to the way science knowledge is portrayed, it relate to the nature of science. The assignments could address nos by assigning the students to evaluate their results and conclusions in the light of alternative explanations. It could also be addressed during discussions, and possibly written in workshop script, for the supervisor.

2. 'Experience real things' strongly relates to principle three 'raising wonder', as the use of real object is used to raise this wonder and curiosity. Sensory experience of an abundance of objects could be impressive, focus students attention and raise questions (Van Keulen, 2011).
- 'Use of real objects' could be checked, are objects real
- 'The objects raise questions and can focus attention', relates to principle three. If conditions from principle three are apparent, this principle is selected as well, the objects then have to potential to raise question and thereby focusing attention.
- 'Rich and varied opportunities for direct sensory experiences', is checked based on the opportunity for multiple sensory perceptions. Van Keulen names that besides observing, feeling and hearing could be used, and the design is assigned as 'rich' if these are addressed.
- 'Use of supplemental materials', could be checked in the materials, if other devices are used
- It is not stated what is meant by an authentic context, and the interpretation determines whether the workshop meets this aspect. This authentic context should help students to interpret and give meaning to their experience (van Keulen 2011). In a study on instructional design for authentic learning, it is stated that an authentic context reflects the way knowledge will be used in real life, and the physical learning environment should reflect this. The design should preserve the complexity of the real-life setting, without fragmentation or simplification of the environment, and a large number of resources should be provided to enable examination from different perspectives (Herrington &Oliver, 2000). The context for the workshops, could be determined as authentic if the designs reflect the way the knowledge is ultimately used, in real life, or in a representative science or professional context.
- 'The objects have a story that can be discovered' is selected if the objects are not directly self-explanatory.

3. 'To raise wonder' contains conditions as well as goals for the attained level. These elements relate to highly intangible aspects and criteria that could be checked in the design were based on Van Keulens' document (2011) and Van Graft & Kemmers' SLO document (2007)
- 'Confrontation with surprising objects or phenomena that raise surprise': this relates to subjects that students' are not confronted with daily, that are just above their knowledge level though fit in their world of experience. The materials are assumed to be special, in a sense that probably most schools will not have a collection of fossil (replica's) and real bones for the students to work with. Whether these are objects of personal interest to students is debatable, and highly dependent on personal back-ground, gender and previous out-of school experiences (Paris, Yambor &Packard, 1998; Uitto, Juuti, Lavonen and Meisalo, 2006). The questionnaire will address this question.

- ‘Students’ attention is focused by aspects that arouse curiosity and can be used to build upon’. Curiosity could arise after surprise, therefore the same conditions account here. Curiosity could be used for knowledge construction, by asking open questions and steering in the form of examples is sometimes needed (Van Keulen, 2011).
- ‘The natural curiosity of children is being used’ is served by the other elements of this principle
- ‘The confrontation leads to reflection and questions from children: What's this? How can this be? I want to know!’. This goal describes what is meant by curiosity. Whether the curiosity leads to reflection and questions cannot be checked in the design.
- ‘These questions initiate a searching process that leads to knowledge / insight’ depends on the selection of previous aspects. If curiosity is enhanced, and questions are raised, it could be determined if the workshop gives opportunity for these questions to initiate the search process, or rather provides the questions to be investigated. To what extent actual insights and knowledge are gained, cannot be checked in the design, only on attained level, this is not thoroughly studied.

4. ‘Learning about diversity and change’ is a principle relating to content and conditions could be checked as (explicitly/implicitly) present. Several elements relate to student experience that are not of such value for this study to determine material criteria.

- ‘Students prepare on Naturalis focus on biodiversity’ is not evaluated
- ‘Students experience out-of school’ learning will be addressed in the questionnaire
- ‘Focus on diversity and origin and meaning of this diversity’ could be checked on the topic addressed by the design
- ‘Students get an image of, and form an opinion about diversity and change’. Naturalis portrays an enormous collection of objects as to impose biodiversity on the visitors. By bringing the students in contact with a variety of objects during the workshops, this aspect could be supported as well. The topic of the workshops, and the focus on change and diversity with the use of objects that display this, could support students imaging. This could be checked by the materials used. The opinion forming of students will be checked as present in the design when the assignment specifically asks students for their opinion, or encourages a discussion on diversity and change.

5. ‘Exemplary work’ principle emphasizes the need to focus students’ attention, and the abundance of objects and examples should be limited after the first impressions and wonder is raised. This principle should counterbalances the second principle that states to get students attention by impressing them with an abundance of rich sensory experiences.

- ‘Most of the time is spent on a limited number of examples’, is a description in order to prevent distraction, and to the focus of students but not overwhelming them (Van Keulen, 2011). The previous principle could conflict with this element, so a variety, though limited number of elements should be used.
- ‘Examples are in line with the prior knowledge and experiences of students’, is addressed already in principle three.
- ‘Examples are easily perceived by sensory experiences’, checked in materials if real, tangible objects are used this aspect is selected.
- ‘Examples are prototypical, represent a larger group of objects or phenomena’, is selected if the objects used for investigation resemble the kind of objects used in the real field of investigation on the topic
- ‘ General statements or conclusions are built based on the examples’, is a goal on attained level. Material can be checked as to pose generalization questions explicitly or implicitly, in relation to evolutional processes in a broader sense.

6. ‘Heart, head, hands’, a principle that reflects the importance of affective learning, by eliciting students’ appreciation for the objects. The conditions for appreciation as described by Van Keulen (2011) are very similar to those described previously for the aspects of wonder and curiosity and will not be considered again.

- The elements ‘opportunity to act’ and ‘space for student to appreciate and express this’ can only be addressed on implemented level, these aspects will however not be explicitly examined .
- ‘Students are challenged to think’, is a very interesting condition as ‘challenge’ has been indicated as an important motivational factor for science learning and attitudes towards science, for example stated in Osborne, Simon, & Collins, review (2003). At the same time the degree of challenge students experience depends on many factors, and is a complex topic for investigation. Activities that are too challenging impede effective content learning, whereas if activities are too easy higher level thinking skills will not be developed (Colburn, 2000). Since the degree of challenge is not particularly subject of this study, it will not be addressed in the design evaluation, the questionnaire will grasp on this element superficially.
- ‘Students can personally identify with stories and objects’, is very personal and is influenced by many factors. This aspect will not be checked in the design, the questionnaire addressed this aspect.

7. ‘Take differences into account’, relates to both gender, age and students interests. Without going into depth on what different interests’ and learning styles could be addressed for the different target groups, (this

would need a literature study), the elements could be checked for any diversity of subjects or methods, possibly contributing to the different target groups.

- The different beta-mentalities are described in the ruler, and it could be checked if the designs address more than one from the given criteria; ‘hands-on’, ‘daily life links’, ‘status-conscious aspects’, ‘useful’.
- The assignments are challenging and uncommon for school education, addresses ‘challenge’ which is discussed previously, and the fact that it is not school-like, is discussed in principle three. The selection will be based on the selection for principle three. Student opinion on this aspects will be obtained with the questionnaire.
- ‘It is necessary to reason and develop mental representations’, relates to what is described by Van Keulen (2011) as the cognitive abilities that students of upper school levels start to develop. Students are less fixed to sensational experience, therefore the assignment should not focus on sensory perceptions only. The designs are checked for the necessity to not only observe but to reason about what is seen. This aspects links to the higher-order thinking elements from principle one, being a researcher.
- The different kinds of structure could be checked in the designs, related to the assignments, and forms of group work.

8. ‘Social interaction’ is stated as an important aspect to elicit articulation and critical reflection on the experiences, and conversations contribute to better understanding and anchoring. Besides this, social interaction, is conditional for many activities that contribute to experiences and insight namely; planning, cooperation, dividing and executing tasks, and doing and sharing observations (Van Keulen, 2011).

- ‘It takes into account the group size’, the materials can be used in a flexible way and adjustments can be made on the group size
- ‘There is opportunity or necessity to collaborate and discuss’, selected as workshops are based on cooperation in small groups.
- ‘There is interaction with staff / researchers Naturalis’, this interaction is considered to be present when discussion with supervisor is present in the script. Besides this, the supervisors interact with the students while they work in small groups, the amount and type of interaction cannot be determined in the materials, and depend on the supervisor.
- ‘Children are challenged to verbalize their stories, experiences and opinions’, checked in material if assignments directly elicit personal story telling, or asks students to discuss experiences or opinions with each other. Challenge and support for students’ verbalization by the supervisors could be stated in script.

Appendix IV: Questionnaire Underpinning

- M2 Selected principles from the Educational Ruler addressing intangible aspects as motivation, curiosity and attitude

Principle aspect from ruler ('q' in Table 9)	Question
1.zelf onderzoeker zijn: Een scala aan vaardigheden wordt ingezet: voorspellen	Voordat jullie gingen determineren, werd in de opdracht gevraagd om een hypothetische ontwikkelingslijn te maken. Hebben jullie dat gedaan? Ja. Waarop baseerden jullie deze hypothese? Geef een korte beschrijving Nee. Waarom niet?
1. zelf onderzoeker zijn: Een onderzoekende houding wordt gecultiveerd: nieuwsgierigheid 3.verwondering wekken: De aandacht wordt gericht op facetten die nieuwsgierig maken en waar op voortgebouwd kan worden.	Heeft de opdracht je nieuwsgierig gemaakt? Ja, ik zou wel meer willen weten over… (kruis aan, meerdere mogelijk) Nee, omdat… (kruis aan, meerdere mogelijk)
3.verwondering wekken: Kinderen worden geconfronteerd met objecten of fenomenen waarvoor ze zich kunnen verwonderen.	Ik vond het bijzonder om met de objecten (kiezen, botten) / schedels te mogen werken Ik heb al vaker met zulke objecten/ schedels gewerkt
6. Hart, hoofd en handen: Kinderen worden uitgedaagd na te denken.	Ik vond de opdracht uitdagend
7.4: 1. De opdrachten zijn uitdagend en niet schools.	Vond je het soort opdracht anders dan opdrachten die je op school krijgt? Ja. Beschrijf kort wat je 'niet schools' vond aan de opdracht Nee. Waarin leek de opdracht op schoolse opdrachten Categories: object/skull (literally named), practical (active, practical, not from a book, really doing something with the theory), group work (cooperation, group discussions, large groups, working together), independency (literally named or 'working individually') and other.

- M2 Intangible aspects of the IB goals as identified previously (in relation to motivation, attitude and science as a way of knowing).
  - Ik zou het leuk vinden om meer onderzoek aan echte objecten/schedels te doen
  - Ik zou het leuk vinden om meer onderzoek naar evolutie te doen
  - Ik vond het een uitdaging om als een onderzoeker te werk te gaan
  - Is je mening over evolutiebiologie door de opdracht veranderd?
    - Ja. Dit is er veranderd:
    - Nee deze is nog hetzelfde gebleven, namelijk
  - Is je mening over onderzoek uitvoeren door de opdracht veranderd?
    - Ja. Dit is er veranderd:
    - Nee deze is nog hetzelfde gebleven, namelijk

Categories changed opinion about evolution:	Categories changed opinion about doing research
No, still negative (e.g. boring)	No, still negative (e.g.boring)
No, still positive (e.g. interesting, like subject)	No, still positive (e.g. interesting, fun)
No, [empty] or "just the same"	No, [empty] "just the same
No, learned nothing (knew already…)	No, learned nothing (e.g. knew already process, activities, still don't understand (1x)
Yes, more positive (e.g. interesting, challenge, fun)	Yes, more positive (e.g. interesting, can be fun)
Yes, less positive (e.g. it turned out more boring…)	Yes, less positive (e.g. even more boring)
Yes, [empty]	Yes, [empty]
Yes, learned that… [concept related evolution]	Yes, learned that research (e.g. is important, how it works)
Yes, learned that.. [related to research about evolution]	

- E2 Science as a way of knowing:
  Door de opdracht begrijp ik nu meer over het onderzoek dat evolutiebiologen uitvoeren

- M1 content learning
  - Door de opdracht ben ik meer te weten gekomen over het verloop van de evolutie (van de mens)
  - Door de opdracht ben ik meer te weten gekomen over het construeren van evolutionaire stambomen

- Overall opinion on the workshop, asking the student to indicate which part of the workshop they enjoyed most, and which part they enjoyed the least. This provided information on student experience in relation to certain IB related features of the workshops.

Master Thesis:
Science Education and Communication
Sofia J. van Santen
sofiavansanten@gmail.com
Utrecht University 10-03-2014