

To know or not to know: uncertainty is the answer. Synthesis of six different science exhibition contexts

Helena Thuneberg and Hannu Salmi

Abstract

This meta-article aims to explore the role of uncertainty in knowing in informal science learning contexts. Subjects (N=2591) were sixth-graders from four countries. In addition to the correct and incorrect questionnaire alternatives, there was a “don’t know” option to choose if uncertain of the answer. The unique path-analysis finding showed that the role of motivation was uniformly positive on correct and negative on uncertainty of answers. In all contexts the number of correct answers increased, incorrect and uncertain answers decreased. Interestingly, although there was no more difference in knowledge pro boys after the intervention, the girls were still more uncertain.

Keywords

Science centres and museums; Science communication: theory and models; Science education

DOI

<https://doi.org/10.22323/2.17020201>

Introduction

Uncertainty is a phenomenon everyone recognizes from their own experience of everyday situations or of more demanding tasks — including, of course, formal test situations. As Lindley [2013, p. 2] puts it: “Uncertainty is everywhere and you cannot escape it”. However, uncertainty of knowing is a rather un-explored area in informal education. Science exhibitions serve as informal learning interventions, they are not related to grading and assessment, but instead they aim to motivate pupils to explore, do hands-on experimentation, and to gain both content and procedural science knowledge and deeper understanding. Because the aspect of assessment is missing, unlike in studying at school [Nasir et al., 2006], the pupils do not have to be afraid of failure or mistakes [cf. Oppenheimer, 1968]. This is especially important, because the traditional school science instruction has been seen as failing to motivate student’s interest in science and to further elaborate that interest [Goldschmidt and Bogner, 2015; Metz, 2008].

Learning from informal sources and in out of school environment has, in turn, been found to be effective and motivating [Osborne and Dillon, 2008; Fenichel and Schweingruber, 2010]. Interventions encouraging exploration and inquiry have resulted in significant learning and motivation in regard to science content and process knowledge, to science concepts and scientific inquiry [see Cotabish et al., 2013; Banilower, Fulp and Warren, 2010; Baldassari, 2008].

The informal science learning environment has also been shown to be beneficial for sparking excitement and interest and in creating an atmosphere in which pupils have choice and sense of self-determination, because they have a say on what, when, and how to learn [Rennie et al., 2003; Renninger, 2007]. Motivation is enhanced by the factors of an ideal informal learning experience such as a science exhibition, which stimulates motivational factors [Perry, 1994; Tan and Subramaniam, 2003; Rennie, 2014]: curiosity, confidence, challenge, control (self-determined), play, and communication.

The positive effects of an informal environment for learning gains are, however, not necessarily self-evident: they vary according to the pupils' prior interest in science and their readiness to take responsibility for setting goals for themselves [Renninger, 2000]. The organization of the experience and environment may focus on fun and enjoyment, but lack the scaffolding of reasoning and have a lack of support for deepening the situational interest into a real desire to find out more [Hidi and Renninger, 2006]. If the organization is successful, pupils can more readily set goals that, in turn, develop and increase their science knowledge and knowing.

The purpose of this research was to study knowing, learning, and especially the role of uncertainty in gaining knowledge in formal (school) and informal (science exhibition) science learning contexts.

The analysis was based on pre- and post-knowledge tests of 12 to 13 years old pupils (N=2591) across six informal science learning contexts from Sweden, Latvia, Estonia, and Finland. As a meta-study it comprises those six prior studies, but the aim was to find out something more than just the sum of the earlier results. In the six earlier reports, we focused particularly on the exhibition contents and only on cognitive knowledge learning results. In the present article we wanted to complement and explain those results by especially exploring the role of *uncertainty of knowing*. The idea was to strengthen both the validity and reliability of the earlier findings of the single studies by creating a synthesis of the magnitude and direction of the effects of the variables [cf. Lipsey and Wilson, 2001] and characteristics common to the six different informal science learning contexts.

We were keen to find out the portions of correct, incorrect, and uncertain ("don't know") answers before and after the exhibitions and how they would vary within and between the contexts and boys and girls. The learning contexts were STEM (Science, Engineering, Technology and Mathematics education) exhibitions: Mars and Space, Dinosaurs and Evolution, Augmented Reality, Hands-on Science, 4-D Math and Discover the Natural Phenomena. The same research design was applied in every context, and in this article the aim is also to evaluate whether our theoretical presumptions about the intertwining of the motivational and cognitive components with knowing and learning would hold. Five of the papers have been published in peer-reviewed journals, and the sixth has been submitted.

Next we will explore aspects of *knowledge* and *knowing* in science education in order to understand the role of uncertainty of knowing.

"Science is not only a body of knowledge, but also a way of knowing"
[Harris, 2002, p. 168]

In the science education context it is useful to make a distinction between the substantive knowledge (*the subject matter knowledge*) and the syntactic or epistemic knowledge (*the nature of science*) [Anderson and Clark, 2011]. The latter supports understanding science as being humanly generated and being apt to testing [Rogers and McClelland, 2004], and prevents thinking of knowledge as frozen, final facts that cannot be revised [Harris, 2002]. Epistemic knowledge relates the central aims of science education, namely metacognitive awareness [Michalsky, Mevarech and Haibi, 2009; Harris, 2002], thinking skills and reasoning [see Adey, 2006; Adey et al., 2007; Demetriou, Spanoudis and Mouyi, 2011] and more broadly, learning-to-learn approaches in several European countries [Csapó, 2007; Crick, 2007; Demetriou, Spanoudis and Mouyi, 2011; Hautamäki and Kupiainen, 2014; Hoskins and Fredriksson, 2008].

Content is, however, crucial; and the depth of content knowledge has been found to be an essential factor either in supporting or limiting the learning of scientific reasoning [Brewer and Samarapungavan, 1991]. Teaching scientific reasoning only in an abstract framework is highly questionable [Duschl, Schweingruber and Shouse, 2007], and it should instead happen in the knowledge domains in which students are at their strongest [Lehrer et al., 2001; Stewart, Cartier and Passmore, 2005; Passmore, Stewart and Cartier, 2009; Wisner and Amin, 2001]. Thus, as Hernandez et al. state [2013], it follows that the content coverage forms one of the central quality indicators of STEM education and schooling.

The main component of cognitive ability is the capacity to learn, embrace, and remember knowledge once learned [Raven, Raven and Court, 2000]. The contemporary view of science teaching emphasizes the awareness of pupils' prior knowledge, both content and meta-cognitive, as the central factor of learning. These existing concepts and ideas can be more or less developed and either beneficial or harmful for integrating new contents. [Duschl, Schweingruber and Shouse, 2007; Harris, 2002; Schwarz and White, 2005]. Moreover, according to Duschl, Schweingruber and Shouse [2007], it is essential to realize that if pupils are not personally interested in scientific problems, they are less competent in solving them.

Research results have shown that if differences have been found between boys and girls in the cognitive domain, they have been small [cf., for review in Thuneberg, Hautamäki and Hotulainen, 2014]. However, in mental rotation [Hyde, 2007], combinatorial and proportional reasoning and propositional logic tasks [Meehan, 1984; Guðbjörnsdóttir, 1995], boys have been found to slightly outperform girls. But there are some signs that girls are starting to catch up with boys, and even outperform them in that are traditionally boys', e.g., in mathematical literacy [Vainikainen, 2014; OECD, 2014; Kenney-Benson et al., 2006]. Nevertheless, boys still tend to be better in science [Kenney-Benson et al., 2006].

Knowing, not knowing and uncertainty of knowing

If a student knows something conclusively, guessing is unnecessary. But the non-knowers have to guess, if the "don't know" alternative is missing. They might be tempted to guess even when that "don't know" alternative is given. Based on a review of Mondak and Davis [2002], there have been recommendations to avoid of

a “don’t know” alternative in questionnaires, especially in the educational field. If that alternative has been included, the respondents have been guided only to choose this possibility if they were very sure of not knowing the answer.

There are two opposing views of applying the “don’t know” alternative. Those, who advocate the so-called *number-scoring* method, i.e. offering only correct and incorrect alternatives [Muijtens et al., 1999], argue that a “don’t know” alternative would cause cognitive *noise*: they think it is clearer just to calculate the correct answers. Others, those favoring the *formula scoring* method think that adding the “don’t know” alternative reduces guessing and measurement error and increases reliability [cf. Muijtens et al., 1999] and offers more accurate estimates of knowledge [Sherman, 1976].

In the present study we included the “don’t know” answer option in the knowledge tests and were particularly interested in its results. The hypothesis was that this alternative would give some added information about science knowing and learning, which only correct scores would not provide.

According to Mondak and Davis [2002] there are several interpretation possibilities when the “don’t know” alternative is chosen. In our science study contexts the possibilities would be:

1. Pupils actually knew the answer and were fully informed of the matter in concern. They would have answered correctly, if there had not been this “don’t know” option, which for some reason they nonetheless chose.
2. Pupils were less than fully certain; only partially informed. In the absence of the “don’t know” alternative they would have taken the correct option and being either correct or incorrect.
3. Pupils wrongly believed that they know the answer (i.e., are misinformed), and would have answered incorrectly in the absence of the “don’t know” alternative.
4. Pupils were uninformed, didn’t know, and chose the right or wrong answer by guessing and chance.

These elements clearly appear in the informal learning environments. A huge amount of information and knowledge, especially about modern phenomena, is obtained during the lifespan from informal learning sources, like science exhibitions [Braund and Reiss, 2007; Osborne and Dillon, 2008; Salmi, Thuneberg and Vainikainen, 2016a]. These sources of certainty and uncertainty are, thus, plenty and differ according to the individual. Moreover, the degrees of certainty and uncertainty vary [Lindley, 2013].

Is an uncertainty of knowing a sign of a lack of confidence? Lundeberg [1992] claims that “knowing what one knows and what one doesn’t know” has a key role in learning. Calibration of confidence is part of that and implies a balance between over- and under-confidence, and finding an appropriate state relying on realism. Lundeberg [1992] observed that: males were overconfident and inappropriately so when in fact they were incorrect; females did not in general lack confidence, but this was dependent on the context.

In addition to those differences, cross-cultural factors are involved in uncertainty of answering Atkins [cf. 2000]. In her early paper Sherman [1976], who thoroughly analyzed “don’t know” answers in multiple choice science exercises also mentioned demographic factors, such as parents’ education. She also found that response style differences related, for example, to timidity, shyness, or lack of motivation. According to her study, impulsive pupils would give more wrong answers, and the self-confident, less uncertain, and anxious pupils more “don’t know” answers. All in all, then, uncertainty of knowing varies as a function of personality traits, such as self-confidence, risk taking and competitiveness [cf. review of Mondak and Davis, 2002].

Is uncertainty of knowing a sign of *an epistemic doubt* [Chandler, Boyes and Ball, 1990]? This concept implies to an epistemological crisis in pupils when they grow, drop their “absolutist” view of knowledge [Perry, 1970/1999], and figure out that knowing and knowledge are relative and based on human interpretation. As the National Research Council’s explanation [p. 174 2007] describes: “In this state, they struggle with the erosion of their certainty and may lose confidence altogether that it is possible to be certain about anything.” Results of a study from the late elementary level to the middle level found that students changed from a single right answer state to a dualistic view of science depending on interpretation and took on the uncertainty which follows that view [Perry, 1970/1999; Driver et al., 1996]. One goal of science education has been considered by the National Research Council [2007] to be that pupils would learn both an appropriate sense of trust and reasonable skepticism. They should be given opportunity to develop their personal understanding and, in addition, to take a critical stance in evaluating scientific information [National Research Council, 2007; Duschl, Schweingruber and Shouse, 2007].

The conflict between knowing and not knowing will at its best lead to intrinsic motivation. Motivation in science education is essential both for the individuals and for society. In the present study the aim is to base our approach of motivation on a solid theoretical basis, which is explained below.

Motivation as a key-factor for informal learning

According to the literature [Falk and Dierking, 1992; Tan and Subramaniam, 2003; Osborne and Dillon, 2008; Fenichel and Schweingruber, 2010; Salmi, Thuneberg and Vainikainen, 2016b], the main outcomes and results related to informal learning and especially about science center education are related to the effects of motivation on learning. It is important to be able to predict how the pupils will engage in science later in life [Woolnough, 1994].

The self-determination theory (SDT) applied in this study, provides a theoretically validated and practically reliable measure of motivation [see Deci and Ryan, 1985]. It offers a dialectical framework for understanding how pupils’ inner resources and the learning environment factors are interconnected. Learning environment can either enhance or thwart intrinsic motivation and the integration of external motives by means of autonomy supportive or controlling motivating style. [Reeve et al., 2009]. These are essential factors in our meta-study, in which learning context and especially informal environment form the basis of the analysis.

The SDT theory defines motivation as a continuum [Deci and Ryan, 2002]: The gradual move from *amotivation* (not motivated at all) to the *external* motivation style, means that concrete incentives and avoidance of punishments act as motivators. The next stage is *introjected* motivation, in which those incentives or punishments are symbolic, and motivation is based on experienced pressure. In turn, in *identified* motivation pupils accept external goals because they believe that such goals are beneficial for learning. The most autonomous form of motivation, in the end of the continuum, is *intrinsic* motivation. Here the task is interesting as such and pupils engage because they like it.

When intrinsically motivated, no other person persuades the learner to learn. It leads to deeper learning, creativity, higher achievement and more volitional and greater persistence, especially on tasks, which require conceptual understanding [Jang, Kim and Reeve, 2012; Reeve, 2002; Deci and Ryan, 2002; Niemiec and Ryan, 2009]. As Görlitz [1987] points out, play, exploration, and curiosity enhance a child's cognitive development. Externally, instrumentally motivated pupils, in turn, have been found to learn ineffectively in informal learning settings [Oppenheimer, 1968; Falk and Dierking, 2002; Holmes, 2011].

Educational research indicates that novelty is one of the principal factors in encouraging learning [Berlyne, 1960; Braund and Reiss, 2004; Rennie, 2014]. A new environment generates *situation motivation* [Braund and Reiss, 2004; Zoldosova and Prokop, 2005]. It happens through curiosity, and it involves both external and intrinsic factors. Situation motivation is short-lasting, attention tends to be orientated to irrelevant subjects, and learning can easily lead to superficial results [McClelland, 1951; Atkinson, 1964]. However, it also enhances active observation behavior and the use of the five senses. Moreover, situation motivation is connected with *attractiveness*, and it is found to be one of the keys explaining visitor behavior and learning in an exhibition context. It is the first step into deeper learning described as *holding power* [Screven, 1992]. One concrete indicator of the holding power is how much time pupils intensively spend in the hands-on demonstration in an interactive exhibition [Fenyvesi, Koskimaa and Lavicza, 2014]. However, in order to achieve the goal of transforming external regulations into internal engagement and further into self-endorsed engagement, the crucial factor is experience of autonomy, as the SDT-theory posits [Reeve et al., 2009].

The six science exhibition contexts of this meta-study that represent different hands-on-demonstrations are explained next.

The six contexts of knowledge tests

The research contexts were six modern, interactive science center exhibitions with the topics of Mars and Space, Dinosaurs and Evolution, Augmented Reality, 4-D Math, Hands-on Science and Discover the Natural Phenomena. The exhibitions were different and had different main topics. However, they had a firm common ground and were based on the same approach of science center pedagogy: hands-on opportunities, interactive exhibitions, and applying modern technology to concrete objects [Oppenheimer, 1968; Falk and Dierking, 1992; Rennie, 2014; Salmi, Thuneberg and Vainikainen, 2016b].

The exhibitions were partly touring in various sites and partly in one institutional location. The main idea of this article is to reveal the elements and phenomena that are the same or similar across the exhibition contexts and not dependent on topic and subject.

Research questions

The research questions were as follows:

1. How much does knowing and uncertainty vary within and between the six different science exhibition contexts?
2. Do the pupils learn: did the portion of correct, incorrect and uncertainty of answers change between the pre- and post-tests?
3. How well do the cognitive and motivational variables and gender predict change of correct and uncertainty of answers within and between the exhibition context SEM-models?

The main focus was on knowing, knowledge, and growth of knowledge, but particularly on uncertainty aspect of the knowledge. The purpose was to evaluate whether our theoretical presumptions of the effects of the motivational and cognitive components on knowing and learning would be confirmed. This was a realistic goal because similar design and instruments were applied within all six science and math contexts before. The hypothesis was that some learning would happen, even though the visits were short. In order to confirm the hypothesis, results of different exhibition contexts should not deviate greatly from each other. Deviation would, in turn, show that the roles of motivation, cognitive reasoning, and school achievement on learning were more dependent on content and context than generalizable to science learning overall. Based on social research literature and previous research, the results were expected to be limited and the effects small [Lipsey et al., 2012; Rosnow and Rosenthal, 2003].

We regard autonomous motivation as long-lasting and a relatively stable construct and expected it to have a connection with school achievement and learning results. The literature and in the individual six studies of this study showed this to be the case [cf. Deci and Ryan, 2002; Vainikainen, Salmi and Thuneberg, 2015; Salmi, Thuneberg and Vainikainen, 2016a; Salmi, Thuneberg and Vainikainen, 2016b; Salmi, Thuneberg and Vainikainen, 2016c]. A further hypothesis was that autonomous motivation will predict transient situation motivation. Pupils who had lower autonomous motivation were expected to perform less well in the knowledge tests because they focused on a 'fun' experience rather than on acquiring new information or skills [cf. Holmes, 2011]. They were also expected to show more uncertainty, by having more "don't know" answers than others [cf. Mondak and Davis, 2002]. Along with motivation, reasoning, and past achievement were assumed to predict knowing [cf. Michalsky, Mevarech and Haibi, 2009; Raven, Raven and Court, 2000].

Method

Participants

The participants in 4D-Math exhibiton were from Sweden (n=542) and Latvia (n=408), in the Discover the Natural Phenomena exhibition from Estonia (n=324) and Latvia (n=272). In the rest of the exhibtions all the participants were from Finland: in the Hands-on Science (n=432), in Dinosaur and Evolution (n=322), in Mars and Space (n=144), and in Augmented Reality exhibition (n= 147). In total there were 2591 participants in the sample (n=1278 boys and n=1313 girls).

The pupils were 12 to 13 years old and were chosen mainly because the exhibition planners saw this as being the main target group of the exhibition's educational purposes. In addition, the formal school curriculum of this age group fit best for bridging the gap between formal education and informal learning.

Overall, in these four countries the national curriculum offers degrees of freedom for the teachers to organise learning experiences in informal settings and in the open learning environments. The schools were selected by a random sample from the schools which had preregistered for the mobile exhibition. According to the long-term, big data surveys from the Scandinavian countries, no major social and demographic differences were expected. (As a matter of fact, in Finland the differences between the schools are smaller than even within the schools; cf. Thuneberg, Salmi and Vainikainen [2014] and Vainikainen, Salmi and Thuneberg [2015]).

The schools agreed to take part in the research, and the permissions were received from the parents and schools according to the local laws and common ethical research principles. The pupils were told that the results were confidential and would not have an effect on their school grades.

The idea of the research project was that the teachers would not prepare their pupils for the exhibition visit in order to avoid different types of intervention in the process. This was underlined for the teachers. In addition, the pre- and post-tests were administrated by an external reasearch assistants and not by the classroom teachers.

Measures

Equal instruments were administered in the five science learning contexts.

Pre- and post-test for topic-specific knowledge

The knowledge test was developed for the present study based on the content areas of the science exhibition contexts:

1. *Hands-on Science. Examples of the statements:* Different objects are placed on a horizontal plane. When the plane is tilted, it is always the lightest objects that start gliding first. / An adult human being has approximately 1,5 litres of blood in his/her body.

2. *Dinosaurs and Evolution*. Examples of the statements: Human beings and dinosaurs have existed for a short period simultaneously in the history on the Earth.
3. *Mars and Space*. Examples of the statements: The gravity is weaker on the planet Mars than on the Moon.
4. *Augmented Reality*: Examples of the statements: "The molecules in the air are moving faster when heated."
5. *4-D Math*. Examples of the statements: "A pyramid floor has four corners."
6. *Discover the Natural Phenomena*. Examples of the statements: "Burning produces oxygen".

The topic-specific knowledge tests were piloted with samples between 25–50 pupils to ensure that they were valid; not too easy or too difficult. The tests were administrated to the subjects one week before the science center visit and again about 7–13 days after the visit. The pretests' duration was 60 minutes, and tests were administrated during two school lessons with a break to avoid a too heavy cognitive loading. The post-test lasted only 30 minutes.

The pupils' task was to judge whether the statements were correct or incorrect. The characteristic feature of the approach of present study was that the students also had the option to say that they do not know the answer. The answering options to the statements in the test were 1= true, 2 = untrue, 3 = I don't know. Pilot testing and their analyses using Item Response Theory revealed that some very difficult items had poor discrimination value, so they were omitted from further analyses. The final test scores for pre- and post-test were calculated by summarizing the remaining items.

The knowledge measures showed to be reliable. The reliabilities for the tests were::

Hands-on Science, pre-test: $\alpha = .92$, 59 items, posttest: $\alpha = .93$, 59 items;

Dinosaurs and Evolution, pre-test: $\alpha = .92$, 59 items, posttest: $\alpha = .93$, 59 items;

Augmented Reality, pre-test: $\alpha = .72$, post-test: $\alpha = .77$, 31 items;

4-D Math, pre-test: $\alpha = .82$, post-test: $\alpha = .89$, 27 items;

Discover the Natural Phenomena, pretest: $\alpha = .78$, post-test: $\alpha = .76$, 21 items;

Mars and Space, pretest: $\alpha = .87$, post-test: $\alpha = .89$, 33 items.

Deci-Ryan Motivation

In the six exhibition contexts, testing students' motivation was based on Self-determination theory (SDT). The Deci-Ryan scale was administrated as a pre-test because, theoretically, after a one-day intervention, there should be no major changes in overall motivation, which is related to the whole personality.

The Deci-Ryan Motivation (SRQ-A: Self-Regulation Quality — Academic) scale has 32 standardized items. Each of them has four answering options: 1 = not at all true, 2 = not nearly true, 3 = somewhat true, 4 = totally true. The summative variables locate themselves to the self-determination continuum in the following order: External, Introjected, Identified, and Intrinsic.

The SRQ-A test includes a formula created by the Deci-Ryan research group [Ryan and Connell, 1989], and based on this formula, the RAI Relative Autonomy Index, was calculated. The RAI describes the overall autonomy level experienced by the pupils. In this study, the RAI was used as the indicator of autonomous motivation. The positive sign in RAI indicates that the experience is rather autonomous, negative that one depends on others.

Situation Motivation test

Situation motivation was measured with a questionnaire consisting of 13 Likert scale items (scale 1–5, totally agree — totally disagree). The questionnaire was administered as a post-test only. This test provided information about how attractive the exhibition was to the students.

Raven test

The cognitive measure was a visual reasoning and learning capacity test: Raven Standard Progressive Matrices [Raven, Raven and Court, 2003]. This test has been widely utilized both in practice and theoretical research [Greenfield, 2009]. According to the test theory [Raven, Raven and Court, 2000, pp. 1–2], the main elements in the common cognitive ability are the capacity to learn and the capacity to embrace and remember the knowledge once learned. The Raven test measures non-verbal cognitive skills, the particular ways in which people apply their minds to solving problems. Many researchers suggest that thinking skills are essential to effective learning. One of the researchers Adey [2006] claims based on long-time Cognitive Acceleration project results that developing higher order thinking skills in science will improve general intellectual ability, and help pupils get more out of learning and life. For example Greenfield (in the thematic issue of the Science 1/2009) considers the Raven test a useful method in relation to thinking skills.

It provides a reliable standardized tool for comparing the individuals' learning abilities compared to the representative age group, irrespective of sex.

In each test item, the subject is asked to identify the missing element that completes a pattern. The test contains sixty items that have been divided into five sets (A, B, C, D, E). Each of these groups contains twelve different tasks.

School achievement

The school achievement variable was the summary of the four school grades (physics, chemistry, mathematics, and mother tongue) provided by the teachers. Mother tongue was included because of the shown relevance of reading comprehension in science learning [cf. Snow, 2010]. The pupils were classified into

three categories according to their school achievement level: A+ (above average school achievement, 25% of the pupils in each class); A (average achievement, 50% of the pupils), A– (below average achievement, 25% of the pupils). Boys showed to be overrepresented in the lowest, girls in the highest achieving group, $\chi^2(df2)=15.40, p < .001$, when all exhibition contexts were analyzed as a total.

Analysis methods

Because we could not enter all the knowledge variables — the correct, incorrect, and “don’t know” answers — simultaneously to the SEM-model (because the result would have been total linear combinations from each other), we studied the change between the pre- and post-test also by GLM repeated measures method and the *incorrect* answers only that way. As the measure of effect-size we used partial η^2 coefficient, which does not deviate from the recommended generalized coefficient in analysis when using only one grouping factor [Bakeman, 2005]. For illuminating the pre- and post-test levels by gender and exhibition context we obtained confidence-interval plots. Without them the interpretation of the learning gains would have been problematic [see Becker, 2000].

In order to answer the third research question we used the structural equation modeling SEM (AMOS 22). By SEM we wanted to find out how the observed data would confirm the theoretically based connections. The *autonomous motivation (RAI)*, *sex*, *cognitive reasoning* and *school achievement* were used as covariates to control their effects on measured knowledge variables (*correct* answers and *uncertainty* of knowing in the pre- and post-tests) and *situation motivation*, which was only measured by a post-test. We also wanted to see whether a same model would fit across all exhibition contexts, and to test that used exhibition context as a moderator. To obtain these goals, we used the parametric bootstrapping method and the Bollen-Stine method. The goodness of fit evaluation of the models was based on a χ^2 -test and several goodness of fit indexes. By SEM it is possible to observe the stability or the level of change of knowledge and find out whether this is similar for the different sexes. The invariance was tested also by using sex as a moderator.

The overview in Table 1 presents the paths from 1) the research questions to, 2) the operationalized measures and to, 3) the analysis methods for obtaining the results and to, 4) the effects-sizes to evaluate statistically their meaning and value.

Missing values

In the data from the six exhibition contexts there were on average 5% missing values (most in the 4-D math knowledge tests: 16%). [In the 4-D math exhibition context school achievement was not measured]. The list-wise method was used to remove the cases with missing values in the SEM path analysis, because the bootstrapping method requires it.

Table 1. Research questions and related measures, analysis methods and effect-sizes.

Research questions	Measures	Analysis methods	Effect-size
How much does knowing and uncertainty vary within and between the six science exhibition contexts?	Pre- and post-knowledge tests (three alternatives in each question: <i>correct, incorrect, don't know</i>)	Comparison of portions of correct, incorrect and <i>don't know</i> answers by MANOVA using exhibition context and sex and their 2-way interactions as the fixed factor terms. For graphic presentation of pre- and post-test means: 95% confidence interval plots.	Partial $\eta^2 > .01$ small, $> .06$ middle, $> .14$ large
Do the pupils learn: does the portion of correct, incorrect and uncertainty of answers change between the pre- and post-tests?	Pre- and post-knowledge tests (three alternatives in each question: <i>correct, incorrect, don't know</i>)	Comparison of the change between pre- and post-tests by GLM Repeated measures.	Partial $\eta^2 > .01$ small, $> .06$ middle, $> .14$ large
How well do the cognitive and motivational variables and gender predict the change of correct and uncertainty of answers within and between the exhibition context SEMmodels?	1. Pre- and post-knowledge tests (three alternatives in each question: <i>correct, don't know</i>). 2. Relative autonomy experience test (RAI, relative autonomy index) 3. Visual reasoning test (Raven) 4. School achievement score based on science, math and mother tongue	SEM, path-analysis (by AMOS 22) of direct and indirect effects. Sex, reasoning, relative autonomy and school achievement were used as covariates and predictors on correct and uncertain answers in pre- and post-knowledge tests and on situation motivation. The prediction of the pre-knowledge test results on the post-tests and situation motivation on the post-knowledge tests. Application of parametric bootstrapping and the Bollen-Stine method, which produces corrected <i>p</i> -values also for indirect effects [Bollen and Stine, 1992]. The goodness of fit evaluation of the models based on a χ^2 -test ($p > .05$) and several indices: NFI and CFI (good fit $> .90$, or better $> .95$), RMSEA reasonable fit $> .08$, good fit $> .05$ [see Byrne, 2009]. For testing the invariance of the models across a) exhibition contexts and b) boys/girls, comparison of the unconstrained and fully constrained overall model by χ^2 -test and in case of non-invariance pair-wise comparisons of the regression weights (paths) between the models: z-test.	Standardized β coefficients, R^2 -multiple correlations for the explained total portion on the knowledge and situation motivation variables.

Results

The confidence interval plots of knowledge variables (*correct, incorrect* and *uncertainty* of answers) are presented by exhibition context and sex, and the significance indicated by stars.

Changes in knowledge from pre-test to post-test

The main result was that in all studies the change of the *correct* answers was positive and significant, the correct answers increased: Mars and Space ($p = .000$, $\eta^2 = .12$); Discover Natural Phenomena ($p = .005$, $\eta^2 = .013$), Dinosaurs and Evolution ($p = .002$, $\eta^2 = .03$), Augmented Reality exhibition ($p = .000$, $\eta^2 = .107$), 4-D Math ($p = .000$, $\eta^2 = .02$), and Hands-on in Science ($p = .000$, $\eta^2 = .411$). However, in the Hands-on in Science the interaction effect ($p = .001$, $\eta^2 = .03$) complicated the interpretation, and the analysis showed that the change was less powerful in the boys' group: ($p = .000$, $\eta^2 = .31$), than for the girls, ($p = .000$, $\eta^2 = .51$).

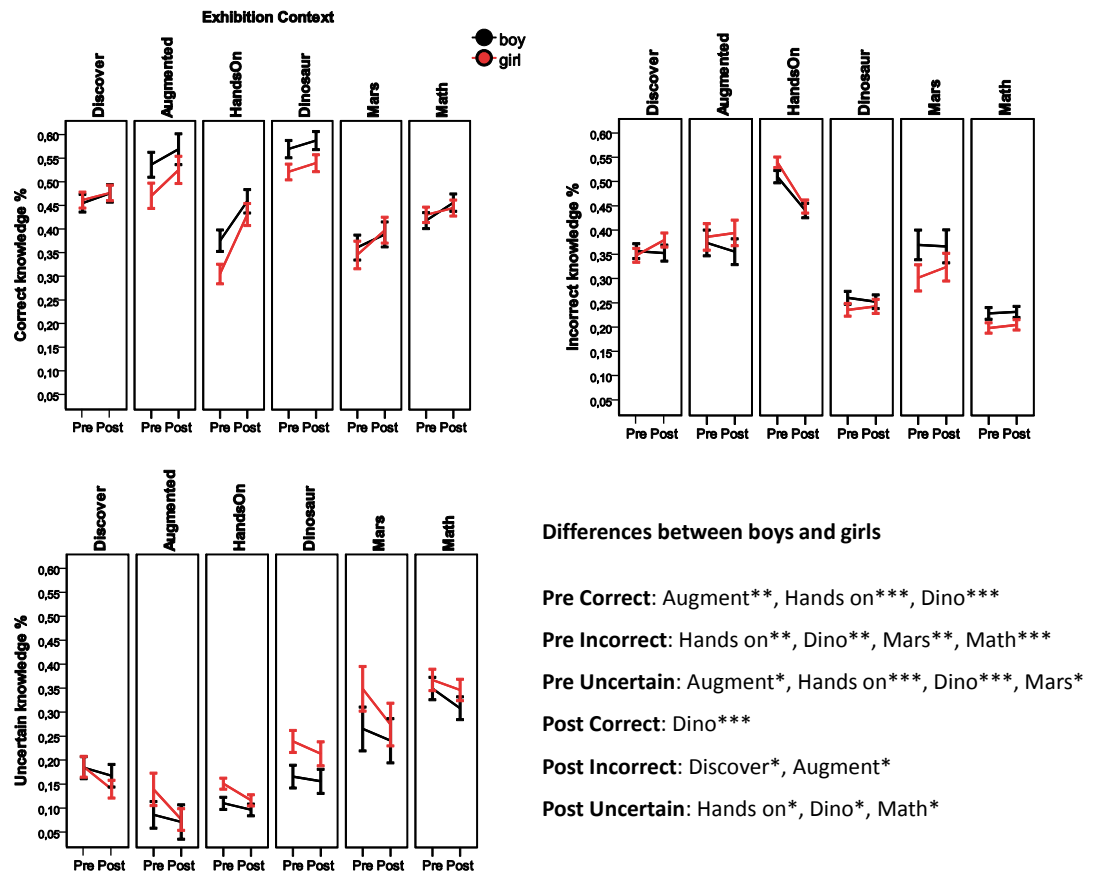


Figure 1. Correct (above left), incorrect (above right) and uncertain (below left) knowledge results of pre- and post-tests by exhibition context and gender. (** $p < .001$, ** $p < .01$, * $p < .05$).

In the case of *uncertainty* the main result was that it decreased in all exhibition contexts. The change was simple in three of the contexts: Dinosaurs and Evolution ($p = .017$, $\eta^2 = .017$), Mars and Space ($p = .000$, $\eta^2 = .082$), and 4-D Math, ($p = .000$, $\eta^2 = .02$). However, there were significant interaction effects: in Discovery of Natural Phenomena ($p = .028$, $\eta^2 = .008$) the change was non-significant for boys ($p = .134$), but highly significant for girls ($p = .000$, $\eta^2 = .100$); in Hands-on in Science ($p = .001$, $\eta^2 = .027$) the effect was much smaller for boys' ($p = .001$, $\eta^2 = .047$), than for girls' ($p = .000$, $\eta^2 = .236$); and in Augmented Reality ($p = .000$, $\eta^2 = .042$) the effect was non-significant in boys ($p = .35$), but for girls significant ($p = .000$, $\eta^2 = .260$).

The change of *incorrect* answers was not uniform. The change was nonsignificant in the Mars and Space, Dinosaurs and Evolution, 4-D Math and Space and Augmented Reality exhibitions. In two cases there were interaction effects due to time and sex. In the Hands-on Science ($p = .006$, $\eta^2 = .018$) the number of incorrect answers dropped significantly, but the effect was larger within the girls' group ($p = .000$, $\eta^2 = .59$) than for boys ($p = .000$, $\eta^2 = .42$). In the Discovery of Natural Phenomena exhibition ($p = .002$, $\eta^2 = .016$) the change was non-significant within the boys' group ($p = .644$), but significant within the girls' group ($p = .000$, $\eta^2 = .05$), in which the amount of incorrect answers increased.

SEM Path analysis

Path modeling was conducted in Amos 22 in order to find out whether the observed data would confirm the theoretically based connections in relation to the research questions. Autonomous motivation (RAI), sex, cognitive reasoning (Raven), and school achievement were used as covariates to control their effects on measured knowledge variables (Correct answers in timepoint 1 and 2, Uncertainty of knowing time-point 1 and 2) and Situation motivation, which only was measured as a post-test. The final model containing only significant effects showed to fit the data well: $\chi^2 = 39.612$, $df = 30$, $p = .113$; NFI=.993, CFI=.998; RMSEA=.014. [NB: The 4-D Math context was compared with other contexts by a model not containing school achievement]. The invariance test, exhibition context as the moderator, showed that the contexts were different at the model level ($p < .001$), and thus the path differences between the models were checked path by path (results, see Table 2). Similarly the gender groups were different at the model level ($p < .05$) and the paths were compared between the boys and girls.

Figure 2 presents the path model synthesis of the six science exhibition contexts. The figure illuminates which paths are most commonly significant within the individual models. Information of the significant effects are marked by arrows and the control variables and their correlations shown by the double arrows. The red arrows indicate effects which are significant at least in 4 to 6 exhibition contexts, the black arrows show significant effects in 3 contexts, and the dashed line arrows indicate significant effects in 2 contexts. The magnitude of the paths (the standardized beta-coefficients as the statistical indicators), are presented in Table 2.

The path differences between the 5 exhibition context models were obtained by pairwise comparisons (4-D Math not included). Although there were differences, the sizes of the effects in 70% of paths did not differ significantly. In those situations in which a predictor only in one to three exhibition context models was significant, the difference *between* the studies only occasionally reached significance. Furthermore, in only three instances there was an inconsistency between the sign of the coefficients of the predictors. Most of the significant differences appeared in the path from the relative autonomy experience (RAI) to Situation motivation, in which all predictors, indeed, were significant *within* the exhibition contexts; but the predictor in the case of Discover Natural Phenomena exhibition was smaller than the others. Furthermore, with the exception of one study, there was an effect from situation motivation on the correct answers in the posttest.

In the following sections, the relationships between the variables are explained in more detail.

Knowledge from pretest to posttest

The result of the SEM-modeling is that previous knowledge clearly predicted *correct* answers also later, in the posttest. In Hands-on Science the prediction was stronger than in other exhibition contexts, and in the 4D-Math and in Discover Natural Phenomena exhibitions weaker. Similarly *uncertainty* of knowing explained uncertainty in the later situation. Moreover, the less correct answers, the more often there were uncertain “don’t know” answers. However, the prediction

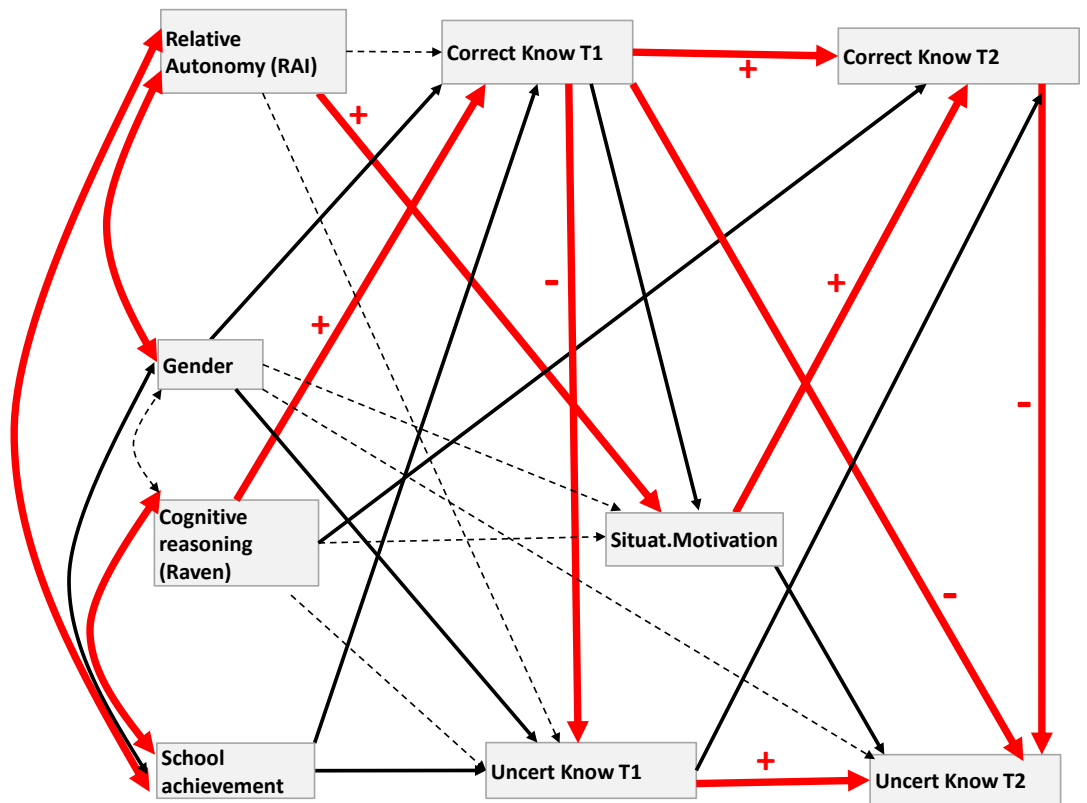


Figure 2. Path model synthesis of the six science exhibition contexts. *Note!* Red=significant effect in 4–6 exhibition contexts, Black=in 3 contexts, Dash line=2 contexts.

was lower in all exhibition contexts in the post-test than in the pre-test situation. The *correct* answers in the pretest somewhat predicted directly and negatively *uncertainty* in the posttest in all exhibition contexts.

Sex

Being a boy predicted *correct* answers in three out of six exhibition contexts studies (Augmented Reality, Hands-on Science and Dinosaurs and Evolution). In the posttest the effects were smaller. In the pre-test being a girl directly predicted *uncertainty* of knowing in three out of six exhibition contexts (Dinosaurs and Evolution, Mars and Space and 4-D Math).

In the post-test there was only one weak direct effect (4-D Math). In the case of Mars and Space, being a girl positively predicted situation motivation.

Because the invariance test showed that sex was a significant moderator, the regression weights (paths) were compared between the boys and girls. The path by path comparison revealed three significant differences: in the girls' group there were higher effects than in the boys' group: a positive effect of school achievement on correct pre-knowledge ($z=2.07, p < .05$), a negative effect of Raven on pre-uncertainty ($z=-3.175, p < .05$) and a negative effect of pre-uncertainty on Situation motivation ($z=-2.112, p < .05$).

Table 2. The significant standardized regression weights of direct effects and path differences within (marked by stars) and between the exhibition contexts (grey shading).

St. regression weights			Discovery	HandsOn	Augment	Mars	Dino	Math
			St β	St β	St β	St β	St β	St β
Correct T1	<-	sex		-.21**	-.30**		-.22**	
Correct T2	<-	sex					-.10*	
Uncertain T1	<-	sex				.15**	.07*	.07*
Uncertain T2	<-	sex						.03*
Correct T1	<-	RAI		.09*		.23**		
Uncertain T1	<-	RAI	-.05*		-.15*			
Correct T1	<-	Raven	.11*	.11*			.17**	.10**
Correct T2	<-	Raven	.19**	.11**			.13**	
Uncertain T1	<-	Raven		.07*		-.12*		
Situat motiv	<-	Uncertain T1			-.23*			
Correct T1	<-	Situat motiv	-.12*	.29*			.12*	
Correct T2	<-	Situat motiv		.12**	.18*	.19**	.17**	
Uncertain T2	<-	Situat motiv	-.05*	-.08*	-.11			
Correct T1	<-	School Ach	.16**	.19**	.32**			
Correct T2	<-	School ach		.12**				
Uncertain T1	<-	School Ach	.06*	.07*			-.09*	
Uncertain T2	<-	School Ach				-.12*		
Correct T2	<-	Correct T1	.34**	.62**	.47**	.58***	.41**	.29***
Uncertain T1	<-	Correct T1	-.76**	-.82**	-.60**	-.80***	-.83**	.88***
Uncertain T2	<-	Correct T1	.23**	.28**	.19**	.22*	.33**	.25***
Correct T2	<-	Uncertain T1	-.17**				-.19*	-.18**
Uncertain T2	<-	Uncertain T1	.54**	.65**	.51**	.53***	.51**	.39***
Uncertain T2	<-	Correct T2	-.56**	-.60**	-.51***	-.61***	-.71**	-.82**
Situat motiv	<-	Raven	.10*			-.29***		
Situat motiv	<-	RAI	.08*	.21**	.21**	.16*	.23**	.11***

Note! The significant direct effects *within* each exhibition context * $p < .05$, ** $p < .01$, *** $p < .001$. Math context has been compared with others by a model not containing school achievement.

Cognitive reasoning

Cognitive reasoning had a significant role in all exhibition contexts except for Augmented Reality and Mars and Space on *correct* answers. After controlling for the effects of the other variables in the pretest situation, cognitive reasoning still had a positive effect on the correct answers in the posttest in half of the exhibition contexts.

Cognitive reasoning was related to the *uncertainty* of answers, as well. In the pretest, the two weak effects, however, were contrary: one negative and one positive. The direct effect of the exhibition context (Hands-on Science) indicates that the higher the cognitive reasoning, the more uncertain answers there were; for the other effect (in Mars and Space), in turn, the higher the reasoning, the less uncertain the answers. In the post-test there were only indirect negative effects.

Cognitive reasoning had a negative effect on situation motivation in Mars and Space.

School achievement

Besides sex and cognitive reasoning, also school achievement predicted *correct* answers, the coefficients being higher than in the case of cognitive reasoning. School achievement predicted the *correct* answers in the post-test only in one exhibition context. In the case of the *uncertainty* of answers, there were small direct effects in the pretest, one negative (Dinosaurs and Evolution) and two positives (Discover Natural Phenomena exhibition, Hands-on Science). On the post-test *uncertainty* of answers, there was only one direct, negative effect (Mars and Space) [NB: School achievement was not included in the 4-D Math exhibition model].

RAI

All relative autonomy (RAI) predictions on the *correct* answers were positive. There were two direct effects in the pretest (in Hands-on Science and Mars and Space). The effects of RAI on *uncertainty* of answers were all negative. In the pretest there were two direct effects, but in the posttest only indirect effects in every exhibition context.

RAI predicted *situation motivation* directly and positively in all exhibition contexts.

Situation motivation

Situation motivation positively predicted the post-test *correct* knowledge results in all but the Discover Natural Phenomena exhibition. It also predicted negatively, directly and/or indirectly, posttest *uncertainty* of knowing in all exhibition contexts.

Discussion

Uncertainty as part of multiple-choice questionnaires has, indeed, a firm methodological tradition especially in large-scale studies, but that is not the case in the informal science learning context. The novel findings and the added value applying the meta-study concept relates to previously unexplored area of uncertainty of knowing and its change.

The first aim was to find out how much science knowing and uncertainty vary within and between the six exhibition contexts. *The six STEM contexts tell the same story despite different exhibition contents.*

The main discovery of this meta-study was that the results are quite uniform: most effects were in a relatively similar size range. There were, indeed, some single and inconsistent effects (negative vs. positive), but they seem not detract from the general picture. The cautious conclusion is that our research design, measures and model were cross-validated in a manner which can have theoretical and practical value for the planning of informal science education.

Pupils knew more

Our second question asked whether the pupils learned between the pre- and the post-tests. The results uniformly showed that one of the main goals of the science

exhibition interventions was realized: in every exhibition context the pupils learned based on the significant increase in the number of correct answers. There was a moderate to large effect of correct answers on the pretest to correct answers on the posttest depending on the exhibition context. This effect shows that the more the pupils knew beforehand, the more they knew also after, which supports the findings of the importance of prior knowledge on successful learning cited in the Introduction [Duschl, Schweingruber and Shouse, 2007; Harris, 2002; Schwarz and White, 2005].

A positive result was that the learning effects were larger than expected [cf. Lipsey et al., 2012; Rosnow and Rosenthal, 2003], even though there was some variation and some of the exhibitions were more effective than others.

Pupils were less uncertain of knowledge

Overall, the pupils were significantly less uncertain after the exhibition, despite the finding that the pre-test uncertainty rather often predicted the post-test situation. The amount of uncertainty variation in the pre-test was totally around 22% and the post-test 19%. So compared to the pre-test there were 14% less uncertain “don’t know”-answers after the exhibitions. This indicates that science exhibitions allow pupils to test by their own hands-on experimentation the basics of the knowledge, which then reduces uncertainty. In addition, they also answer more correctly after this experience.

Most uncertainty was gathered in 4-D Math and the least in Augmented Reality in both pre- and post-tests. In addition to the mentioned explanations for the uncertain answers, there might also be technical reasons relating to the test-instrument and the questionnaires. Muijtjens et al. [1999] pointed out that the relatively large portion of uncertainty answers might imply that some of the items would actually not belong to the domain of the test, or the formulation of the questions has not been successful. The advantage including the “don’t know” alternative on a pilot study questionnaire might, thus, further support the planning of teaching and assessment.

Exception of the rule

The results relating to our second question about the change the amount of correct, incorrect and uncertain answers were, however, somewhat mixed. Questions arise because, although in general the pupils learned and uncertainty decreased, an exception to the rule emerged. In one of the studies the decrease of uncertainty of the girls’ group changed partly and significantly into false learning (i.e. into incorrect answers), and not only to correct ones. This is interesting, because it leads one to consider and look at the learning process and its possible distraction.

Shulman [2005] claims that when uncertainty is present, one has to learn from experience. The science exhibition provides opportunities for exactly that. However, it might be possible that the hands-on experimentation has for some reason not succeeded, has been interrupted, or possibly there has not been sufficient scaffolding for interpretation of the experiment results. Then they might be more than others apt to believe they have learned, even though their answers

are incorrect. Their balancing of confidence [cf., Lundeberg, 1992], thus, seems not to have succeeded in relying on a realistic basis. Fortunately the main tendency according to the overall results of the exhibition contexts was that *uncertainty made space for correct knowing*.

Autonomy supported knowing and decreased uncertainty

The answer to the third research question of the role of motivation and cognition as predictors on knowing seems logical. Both motivation and cognitive reasoning were shown to support knowing, and they still had direct added value in the post-test, as well.

It was common to all exhibition contexts that relative autonomy predicted situation motivation. The more autonomous or motivated by the situation the pupils were, the higher were their scores of correct answers they received in the knowledge tests after the exhibition experience. This confirmed a considerable amount of previous evidence [Jalil et al., 2009; Lavigne, Vallerand and Miquelon, 2007]. Pupils having lower autonomous motivation correspondingly performed less well. This, maybe because, based on the literature, they are not able to set effective goals for themselves, and might focus rather on having fun than on learning new information or skills. They were also expected to show more uncertainty (i.e. “don’t know” answers), than others, which was only weakly true in the pretest, but indirectly met the expectations.

“Don’t know”-alternative was not harmful for high-achievers

Some earlier studies [cf. Bliss, 1980] relating to the elementary level have suggested that including the “don’t know” alternative would be more harmful for high-achieving students than for others, but at least the results of Muijtjens et al. [1999] from the college level did not support this prediction. We discovered only a weak evidence of effects of the cognitive and school achievement variables on uncertainty in general. Of the totally six effects, half were positive and half negative and, thus, contradictory.

Based on these results, we conclude that our hypotheses concerning the positive role of motivational and cognitive variables on correct knowing were confirmed, but they lack justification in the case of cognitive variables on uncertainty of knowing.

The exhibitions especially supported girls

We also asked, whether the sex of the pupils predict knowing, learning, and uncertainty in learning. Some differences in favor of boys have earlier been observed in mental rotation, combinatorial and proportional reasoning, and propositional logic tasks. In addition, a somewhat new trend has showed that girls are challenging the traditional areas of competence of boys. However, the overall result has been that there are no essential differences in cognitive reasoning and this was also confirmed in the present study, although a few weak correlations appeared to favor girls in Raven. So we can conclude that the outcomes of the

knowledge results are to be explained largely by the motivational factors, degree of autonomous self-regulation, interest and situation motivation (in case of *interest*, see Vainikainen, Salmi and Thuneberg [2015]), and in the ways pupils find meaning in the science subjects and experience the relevance of the informal environments for themselves, as Duschl, Schweingruber and Shouse [2007] suggested.

The results found that in relation to science knowing, boys were in a more advantageous position before the science exhibition in three out of the six exhibition contexts (in the rest of the studies there were no significant differences), so their knowledge level was higher than that of girls. The differences leveled in the post-test situation. There, the effects were mostly indirect, meaning that they were now dependent on previous knowledge, uncertainty, or situation motivation results; only in the case of Dinosaurs and Evolution did the results still directly favor boys.

Because the girls were more autonomous than the boys in all exhibition contexts, it could be that they were taking advantage of the informal learning environment and catching up with the boys.

Girls were still more uncertain

Based on the analysis of variance, the girls were clearly more uncertain in the pre-test than boys in four out of six exhibition contexts. But when the other factors — relative autonomy, cognitive reasoning and school achievement — were controlled by SEM-modeling, there were only two of that kind of significant direct effects. In the post-test there were, however, indirect effects in all but one case still indicating the greater uncertainty of girls.

Considerations and limitations of the study

The six science exhibitions can be described as pedagogical interventions. The learning gains cannot be hypothesized to be substantial, because there was only one visit. The studies were not controlled experiments with experiment and control groups and, thus, it is impossible to draw certain conclusions from the factors that led to learning gains. There were also four countries in which the exhibitions and tests took place, and cultural influence has to be taken into account when interpreting the results. Thus, the science exhibition effects could not be observed in isolation from those mentioned and other happenings at same time at school. It seems reasonable to conclude that the effects may be, at least to some degree, due to the common features of science exhibitions as six studies indicate similar results, and almost no contradictions in the results appeared.

As Visone [2010] comments about standardized tests, the individual test items can include specific characteristics which influence the performance of students. However, dealing with individual items was not possible in the present study of true, untrue and I don't know-answers due to the nature of the study design.

The use of alternatives 1 = true, 2 = untrue, and 3 = I don't know turned out to be a practical way of measuring the uncertainty in knowing. However, this approach has also several limitations, because there is always varying degrees of confidence in certainty of the answers of the pupils. It would have been possible to create a

questionnaire in which the pupils could have had an opportunity to express the degree of their uncertainty, but the weak side of this would have been the complexity and cognitive loading for young pupils. This report took into account the two opposing views of applying the “don’t know” alternative: number-scoring method vs. formula scoring method [Muijtjens et al., 1999; Mondak and Davis, 2002] as explained in detail earlier in this article. This dilemma is also a challenge for the future studies in this area.

Conclusions

Uncertainty of knowing is a rather unexplored area in informal science education. There are at least two main, rather contrary views, relating to the uncertainty, which are especially applicable in informal context. The first one causes us to consider whether the “don’t know” answers relate to rather negatively experienced personality traits, such as anxiety, insecurity in general, or timidity. In addition to demographic variables, sex, ethnicity or, for example, parents’ education may correlate with these phenomena. From the other perspective, it is possible to conclude that uncertainty might reveal confidence (“I’m bold enough to say, I don’t know”, or “I’m aware that there can be other possibilities and in science it can be a question of interpretation rather than fixed facts, as I formerly thought”). The latter view could also imply more developed metacognitive skills and knowledge about knowledge and epistemic knowledge.

In the science and math exhibition contexts uncertainty of knowing was thought to be a rather harmful sign of possible partial or incorrect knowledge. Uncertainty was, thus, to be merely experienced as a lack of confidence, lack of knowing, and lack of enough experience to decide. It was also a way to reduce guessing. The logical aim was then to increase the confidence of the pupils by letting them autonomously explore and observe by their own experience and in a hands-on way the phenomena and manipulate the situations in order to cause and analyze change and through that to understand the matters better.

Confidence was assumed, in addition, to be built on personal concrete experience, on wise scaffolding provided by the teachers/exhibition guides and, crucially, on peer-interaction and collaborative group-discussions.

“Relativity is out there waiting to be revealed” [Lindley, 2013, p. XIV]

But the other side of the coin is that uncertainty and doubt are inherent elements of scientific enquiry and, thus, related to metacognitive awareness, thinking skills and reasoning. If one then assesses uncertainty, is it possible that it might be — at least in some cases — a sign of *epistemic doubt* [Chandler, Boyes and Ball, 1990]: a matter of emerging critical thinking, self-reflection and weighing of whether one has sufficient evidence to decide, rather than just stating facts as being final. This latter view can offer some explanation about the rather baffling results: unexpectedly there was an effect in all six exhibition contexts that implied that to some degree the correct knowledge in the pre-test directly predicted more uncertainty in the post-test. As this was common to all exhibition contexts, it is unlikely to be a coincidence. Further, the weak but direct effects in some exhibition contexts indicated higher uncertainty when there was higher cognitive reasoning ability or

school achievement at stake. The real explanation remains open, but we suggest, that not all uncertainty is for ill in learning. Thus, provoking *epistemic doubt* can be considered an essential and plausible task for informal science learning.

References

- Adey, P. (2006). 'Thinking in science — thinking in general?' *Asia-Pacific forum on science learning and teaching* 7 (2), pp. 1–7.
- Adey, P., Csapó, B., Demetriou, A., Hautamäki, J. and Shayer, M. (2007). 'Can we be intelligent about intelligence?: Why education needs the concept of plastic general ability'. *Educational Research Review* 2 (2), pp. 75–97.
<https://doi.org/10.1016/j.edurev.2007.05.001>.
- Anderson, D. and Clark, M. (2011). 'Development of syntactic subject matter knowledge and pedagogical content knowledge for science by a generalist elementary teacher'. *Teachers and Teaching* 18 (3), pp. 315–330.
<https://doi.org/10.1080/13540602.2012.629838>.
- Atkins, A. (2000). *The Effects of Uncertainty Avoidance on Interaction in the Classroom*. URL: <https://www.birmingham.ac.uk/Documents/college-artslaw/cels/essays/language-teaching/Atkins1.pdf> (visited on 20th November 2015).
- Atkinson, J. (1964). *An introduction to motivation*. Princeton, U.S.A.: Van Nostrand.
- Bakeman, R. (2005). 'Recommended effect size statistics for repeated measures designs'. *Behavior Research Methods* 37 (3), pp. 379–384.
<https://doi.org/10.3758/bf03192707>.
- Baldassari, C. (2008). 'LabVenture: At the Cohen center for interactive learning'. In: ed. by S. evaluation report. Program evaluation and research group, Lesley University.
- Banilower, E. R., Fulp, S. L. and Warren, C. L. (2010). *Science: It's elementary. Year four evaluation report*. Horizon Research, Inc. (NJ1).
URL: <http://files.eric.ed.gov/fulltext/ED518455.pdf>.
- Becker, L. (2000). *Analysis of pretest and posttest scores with gain scores and repeated measures*. URL: www.uccs.edu/lbecker/gainscore.html (visited on 1st November 2015).
- Berlyne, D. (1960). *Conflict, arousal, and curiosity*. New York, U.S.A.: McGraw-Hill.
- Bliss, L. B. (1980). 'A test of Lord's assumption regarding examinee guessing behavior on multiple-choice tests using elementary school students'. *Journal of Educational Measurement* 17 (2), pp. 147–152.
<https://doi.org/10.1111/j.1745-3984.1980.tb00823.x>.
- Bollen, K. A. and Stine, R. A. (1992). 'Bootstrapping Goodness-of-Fit Measures in Structural Equation Models'. *Sociological Methods and Research* 21 (2), pp. 205–229. <https://doi.org/10.1177/0049124192021002004>.
- Braund, M. and Reiss, M. (2004). *Learning science outside the classroom*. London, U.K.: Routledge.
- (2007). 'What does out-of-school learning offer school science?' *The Science Education Review* 6, pp. 35–37.
- Brewer, W. and Samarapungavan, A. (1991). 'Child theories versus scientific theories: Differences in reasoning or differences in knowledge?' In: *Cognition and the symbolic processes: Applied and ecological perspectives*. Ed. by R. Hoffman and D. Palermo. Hillsdale, NJ, U.S.A.: Lawrence Erlbaum Associates, pp. 209–232.

- Byrne, B. M. (2009). *Structural Equation Modeling With AMOS. Basic concepts, applications, and programming*. 2nd ed. New York, U.S.A.: Routledge.
<https://doi.org/10.4324/9780203805534>.
- Chandler, M., Boyes, M. and Ball, L. (1990). 'Relativism and stations of epistemic doubt'. *Journal of Experimental Child Psychology* 50 (3), pp. 370–395.
[https://doi.org/10.1016/0022-0965\(90\)90076-k](https://doi.org/10.1016/0022-0965(90)90076-k).
- Cotabish, A., Dailey, D., Robinson, A. and Hughes, G. (2013). 'The Effects of a STEM Intervention on Elementary Students' Science Knowledge and Skills'. *School Science and Mathematics* 113 (5), pp. 215–226.
<https://doi.org/10.1111/ssm.12023>.
- Crick, R. D. (2007). 'Learning how to learn: the dynamic assessment of learning power'. *Curriculum Journal* 18 (2), pp. 135–153.
<https://doi.org/10.1080/09585170701445947>.
- Csapó, B. (2007). 'Research into learning to learn through the assessment of quality and organization of learning outcomes'. *Curriculum Journal* 18 (2), pp. 195–210.
<https://doi.org/10.1080/09585170701446044>.
- Deci, E. L. and Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York, U.S.A.: Plenum.
- (2002). 'Overview of Self-determination theory: an organismic dialectical perspective'. In: *Handbook of Self-Determination*. Ed. by E. L. Deci and R. M. Ryan. Rochester, NY, U.S.A.: The University of Rochester Press.
- Demetriou, A., Spanoudis, G. and Mouyi, A. (2011). 'Educating the Developing Mind: Towards an Overarching Paradigm'. *Educational Psychology Review* 23 (4), pp. 601–663. <https://doi.org/10.1007/s10648-011-9178-3>.
- Driver, R., Leach, J., Millar, P. and Scott, P. (1996). *Young people's images of science*. Buckingham, England: Open University press.
- Duschl, R., Schweingruber, H. and Shouse, A., eds. (2007). *Taking science to school: learning and teaching in grades K-8*. Washington, DC, U.S.A.: National Academic Press.
- Falk, J. and Dierking, L. (2002). *Lessons without limit*. Walnut Creek, CA, U.S.A.: AltaMira.
- Falk, J. H. and Dierking, L. D. (1992). *The museum experience*. Washington, D.C., U.S.A.: Whalesback Books.
- Fenichel, M. and Schweingruber, H. (2010). *Surrounded by science: Learning science in informal environments*. Board of Science education, Center of education, Division of behavioral and social sciences and education. Washington, D.C., U.S.A.: The National Academic Press.
- Fenyvesi, K., Koskimaa, R. and Lavicza, Z. (2014). 'Experiential Education of Mathematics: Art and Games for Digital Natives'. *Kasvatus and aika* 9 (1), pp. 107–134. URL: http://www.kasvatus-ja-aika.fi/dokumentit/fenyvesial__0804151248.pdf.
- Goldschmidt, M. and Bogner, F. X. (2015). 'Learning About Genetic Engineering in an Outreach Laboratory: Influence of Motivation and Gender on Students' Cognitive Achievement'. *International Journal of Science Education, Part B* 6 (2), pp. 166–187. <https://doi.org/10.1080/21548455.2015.1031293>.
- Görlitz, D. (1987). 'Exploration and attribution in developmental context'. In: *Curiosity, imagination and play: On the development of spontaneous cognitive and motivational processes*. Ed. by D. Görlitz and J. Wohlwill. New Jersey, U.S.A.: Lawrence Erlbaum.

- Greenfield, P. M. (2009). 'Technology and Informal Education: What Is Taught, What Is Learned'. *Science* 323 (5910), pp. 69–71.
<https://doi.org/10.1126/science.1167190>.
- Guðbjörnsdóttir, G. (1995). 'Content variations and performance on formal operational tasks by gender, social class and ability'. *Scandinavian Journal of Psychology* 36 (4), pp. 327–342.
<https://doi.org/10.1111/j.1467-9450.1995.tb00991.x>.
- Harris, P. L. (2002). 'What do children learn from testimony?' In: *The Cognitive Basis of Science*. Ed. by P. Carruthers, S. Stich and M. Siegal. Cambridge University Press, pp. 316–334.
<https://doi.org/10.1017/cbo9780511613517.018>.
- Hautamäki, J. and Kupiainen, S. (2014). 'Learning to Learn in Finland. Theory and policy, research and practice'. In: *Learning to Learn. International perspectives from theory and practice*. Ed. by R. D. Crick, C. Stringher and K. Ren. London, U.K.: Routledge. <https://doi.org/10.4324/9780203078044>.
- Hernandez, P. R., Bodin, R., Elliott, J. W., Ibrahim, B., Rambo-Hernandez, K. E., Chen, T. W. and Miranda, M. A. de (2013). 'Connecting the STEM dots: measuring the effect of an integrated engineering design intervention'. *International Journal of Technology and Design Education* 24 (1), pp. 107–120.
<https://doi.org/10.1007/s10798-013-9241-0>.
- Hidi, S. and Renninger, K. A. (2006). 'The Four-Phase Model of Interest Development'. *Educational Psychologist* 41 (2), pp. 111–127.
https://doi.org/10.1207/s15326985ep4102_4.
- Holmes, J. A. (2011). 'Informal learning: Student achievement and motivation in science through museum-based learning'. *Learning Environments Research* 14 (3), pp. 263–277. <https://doi.org/10.1007/s10984-011-9094-y>.
- Hoskins, B. and Fredriksson, U. (2008). 'Learning to Learn: What is it and can it be measured?' In: *JRC Scientific and Technical Report*. European Commission, Centre for Research on Lifelong Learning.
- Hyde, J. S. (2007). 'New Directions in the Study of Gender Similarities and Differences'. *Current Directions in Psychological Science* 16 (5), pp. 259–263.
<https://doi.org/10.1111/j.1467-8721.2007.00516.x>.
- Jalil, P. A., Sbeih, M. Z. A., Boujettif, M. and Barakat, R. (2009). 'Autonomy in Science Education: A Practical Approach in Attitude Shifting Towards Science Learning'. *Journal of Science Education and Technology* 18 (6), pp. 476–486.
<https://doi.org/10.1007/s10956-009-9164-4>.
- Jang, H., Kim, E. J. and Reeve, J. (2012). 'Longitudinal test of self-determination theory's motivation mediation model in a naturally occurring classroom context'. *Journal of Educational Psychology* 104 (4), pp. 1175–1188.
<https://doi.org/10.1037/a0028089>.
- Kenney-Benson, G. A., Pomerantz, E. M., Ryan, A. M. and Patrick, H. (2006). 'Sex differences in math performance: The role of children's approach to schoolwork'. *Developmental Psychology* 42 (1), pp. 11–26.
<https://doi.org/10.1037/0012-1649.42.1.11>.
- Lavigne, G. L., Vallerand, R. J. and Miquelon, P. (2007). 'A motivational model of persistence in science education: A self-determination theory approach'. *European Journal of Psychology of Education* 22 (3), pp. 351–369.
<https://doi.org/10.1007/bf03173432>.

- Lehrer, R., Schauble, L., Strom, D. and Pligge, M. (2001). 'Similarity of form and substance: Modeling material kind'. In: *Cognition and instruction: 25years of progress*. Ed. by D. Klahr and S. Carver. Mahwah, NJ, U.S.A.: Lawrence Erlbaum Associates, pp. 39–74.
- Lindley, D. (2013). *Understanding uncertainty*. New Jersey, U.S.A.: John Wiley and Sons.
- Lipsey, M. and Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA, U.S.A.: Sage.
- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S. and Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more interpretable forms*. NCSER 2013-3000. Washington, DC, U.S.A.: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.
URL: <https://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf>.
- Lundeberg, M. (1992). *And others highly confident, but wrong: Gender differences and similarities in confidence judgements*. Conference paper presented at the Annual meeting of the American Research Association, San Francisco, CA, U.S.A.
- McClelland, D. (1951). *Motivation and personality*. New York, U.S.A.: Harper and Row.
- Meehan, A. M. (1984). 'A Meta-Analysis of Sex Differences in Formal Operational Thought'. *Child Development* 55 (3), p. 1110.
<https://doi.org/10.2307/1130164>.
- Metz, K. E. (2008). 'Narrowing the Gulf between the Practices of Science and the Elementary School Science Classroom'. *The Elementary School Journal* 109 (2), pp. 138–161. <https://doi.org/10.1086/590523>.
- Michalsky, T., Mevarech, Z. R. and Haibi, L. (2009). 'Elementary School Children Reading Scientific Texts: Effects of Metacognitive Instruction'. *The Journal of Educational Research* 102 (5), pp. 363–376.
<https://doi.org/10.3200/joer.102.5.363-376>.
- Mondak, J. J. and Davis, B. C. (2002). 'Asked and answered: Knowledge levels when we will not take "Don't know" for an answer'. *Political behavior* 23 (3), pp. 199–224.
- Muijtjens, A., Mameren, H., Hoogenboom, R., Evers, J. and Vleuten, C. van der (1999). 'The effect of a "don't know" option on test scores: number-right and formula scoring compared'. *Medical Education* 33 (4), pp. 267–275.
<https://doi.org/10.1046/j.1365-2923.1999.00292.x>.
- Nasir, N. S., Rosebery, A. S., Warren, B. and Lee, C. D. (2006). 'Learning as a Cultural Process. Achieving equity through diversity'. In: *The Cambridge Handbook of the Learning Sciences*. Ed. by R. K. Sawyer. Cambridge University Press, pp. 686–706. <https://doi.org/10.1017/cbo9781139519526.041>.
- National Research Council (2007). *Taking Science to School: Learning and Teaching Science in Grades K-8*. Washington, DC, U.S.A.: National Academies Press.
<https://doi.org/10.17226/11625>.
- Niemiec, C. P. and Ryan, R. M. (2009). 'Autonomy, competence, and relatedness in the classroom. Applying self-determination theory to educational practice'. *Theory and research in education* 7 (2), pp. 133–144.
<https://doi.org/10.1177/1477878509104318>.
- OECD (2014). *PISA 2012 Results: What Students Know and Can Do — Student Performance in Mathematics, Reading and Science (Volume I)*.

- Oppenheimer, F. (1968). 'A Rationale for a Science Museum'. *Curator: The Museum Journal* 11 (3), pp. 206–209.
<https://doi.org/10.1111/j.2151-6952.1968.tb00891.x>.
- Osborne, J. F. and Dillon, J. (2008). *Science education in Europe*. London, U.K.: Nuffield Foundation.
- Passmore, C., Stewart, J. and Cartier, J. (2009). 'Model-Based Inquiry and School Science: Creating Connections'. *School Science and Mathematics* 109 (7), pp. 394–402. <https://doi.org/10.1111/j.1949-8594.2009.tb17870.x>.
- Perry, D. (1994). 'Designing exhibits that motivate'. In: *What research says about learning in science museums*. Vol. 2. Ed. by R. Hannapel. Washington, DC, U.S.A.: ASTC, pp. 25–29.
- Perry, W. (1970/1999). *Forms of intellectual and ethical development in the college years: A scheme*. New York, U.S.A.: Holt Rinehart and Winston.
- Raven, J., Raven, J. C. and Court, J. (2000). Section 3. *Standard Progressive Matrices*, 2000 Edition. Oxford, U.K.: Elsevier.
- (2003). *Manual for Raven's progressive matrices and vocabulary scales*. Oxford, U.K.: OPP Limited.
- Reeve, J. (2002). 'Self-Determination Theory Applied to Educational Settings'. In: *Handbook of Self-Determination*. Ed. by E. L. Deci and R. M. Ryan. Rochester, NY, U.S.A.: The University of Rochester Press.
- Reeve, J., Ryan, R., Deci, E. L. and Jang, H. (2009). 'Understanding and promoting autonomous self-regulation'. In: *Motivation and self-regulated learning. Theory, research and applications*. Ed. by D. H. Schunk and B. J. Zimmerman. New York, U.S.A.: Routledge, pp. 223–244.
- Rennie, L. J. (2014). 'Learning Science Outside of School'. In: *Handbook of Research on Science Education, Volume II*. Ed. by N. Lederman and S. Abell. London, U.K. and New York, U.S.A.: Routledge, pp. 120–144.
<https://doi.org/10.4324/9780203097267.ch7>.
- Rennie, L. J., Feher, E., Dierking, L. D. and Falk, J. H. (2003). 'Toward an agenda for advancing research on science learning in out-of-school settings'. *Journal of Research in Science Teaching* 40 (2), pp. 112–120.
<https://doi.org/10.1002/tea.10067>.
- Renninger, K. A. (2000). 'Individual interest and its implications for understanding intrinsic motivation'. In: *Intrinsic and Extrinsic Motivation*. San Diego, CA, U.S.A.: Academic Press, pp. 373–404.
<https://doi.org/10.1016/b978-012619070-0/50035-0>.
- (2007). *Interest and motivation in informal science learning*. Unpublished report. URL: http://sites.nationalacademies.org/cs/groups/dbasse/site/documents/webpage/dbasse_080085.pdf (visited on 13th November 2015).
- Rogers, T. and McClelland, J. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA, U.S.A.: MIT Press.
 URL: <https://mitpress.mit.edu/books/semantic-cognition>.
- Rosnow, R. L. and Rosenthal, R. (2003). 'Effect sizes for experimenting psychologists.' *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* 57 (3), pp. 221–237.
<https://doi.org/10.1037/h0087427>.
- Ryan, R. M. and Connell, J. P. (1989). 'Perceived locus of causality and internalization: Examining reasons for acting in two domains.' *Journal of Personality and Social Psychology* 57 (5), pp. 749–761.
<https://doi.org/10.1037/0022-3514.57.5.749>.

- Salmi, H., Thuneberg, H. and Vainikainen, M.-P. (2016a). 'How do engineering attitudes vary by gender and motivation? Attractiveness of outreach science exhibitions in four countries'. *European Journal of Engineering Education* 41 (6), pp. 638–659. <https://doi.org/10.1080/03043797.2015.1121466>.
- (2016b). 'Learning with dinosaurs: a study on motivation, cognitive reasoning, and making observations'. *International Journal of Science Education, Part B* 7 (3), pp. 203–218. <https://doi.org/10.1080/21548455.2016.1200155>.
- (2016c). 'Making the invisible observable by Augmented Reality in informal science education context'. *International Journal of Science Education, Part B* 7 (3), pp. 253–268. <https://doi.org/10.1080/21548455.2016.1254358>.
- Schwarz, C. V. and White, B. Y. (2005). 'Metamodeling Knowledge: Developing Students' Understanding of Scientific Modeling'. *Cognition and Instruction* 23 (2), pp. 165–205. https://doi.org/10.1207/s1532690xci2302_1.
- Screven, C. (1992). 'Motivating visitors to read labels'. *ILVS Review: A Journal of Visitor Behavior* 1 (2), pp. 36–66.
- Sherman, S. W. (1976). 'Multiple choice test bias uncovered by use of an "I don't know" alternative'. Paper presented at the Annual meeting of the American educational research association, San Francisco, California, April 19–23. URL: <http://files.eric.ed.gov/fulltext/ED121824.pdf> (visited on 3rd February 2016).
- Shulman, L. S. (2005). 'Pedagogies of uncertainty'. *Liberal education* 91 (2), pp. 18–25. URL: <https://www.aacu.org/publications-research/periodicals/pedagogies-uncertainty>.
- Snow, C. E. (2010). 'Academic Language and the Challenge of Reading for Learning About Science'. *Science* 328 (5977), pp. 450–452. <https://doi.org/10.1126/science.1182597>.
- Stewart, J., Cartier, J. and Passmore, C. (2005). 'Developing understanding through model based inquiry'. In: How students learn. Ed. by M. Donovan and J. Bransford. Washington, DC, U.S.A.: National Research Council, pp. 515–565.
- Tan, L. W. H. and Subramaniam, R. (2003). 'Science and technology centres as agents for promoting science culture in developing nations'. *International Journal of Technology Management* 25 (5), pp. 413–426. <https://doi.org/10.1504/ijtm.2003.003110>.
- Thuneberg, H., Salmi, H. and Vainikainen, M.-P. (2014). 'Tiedenäyttely, motivaatio ja oppiminen [Science exhibition, motivation and learning]'. *Psykologia* 49 (6), pp. 420–435.
- Thuneberg, H., Hautamäki, J. and Hotulainen, R. (2014). 'Scientific Reasoning, School Achievement and Gender: a Multilevel Study of between and within School Effects in Finland'. *Scandinavian Journal of Educational Research* 59 (3), pp. 337–356. <https://doi.org/10.1080/00313831.2014.904426>.
- Vainikainen, M.-P. (2014). Finnish primary school pupils' performance in learning to learn assessments: A longitudinal perspective on educational equity. University of Helsinki, Department of Teacher Education Research Report 360. Helsinki, Finland: Unigrafia.
- Vainikainen, M.-P., Salmi, H. and Thuneberg, H. (2015). 'Situational Interest and Learning in a Science Center Mathematics Exhibition'. *Journal of Research in STEM Education* 1 (1), pp. 51–67.
- Visone, J. (2010). 'Science or reading. What is being measured by standardized tests?' *American secondary education* 1.

- Wiser, M. and Amin, T. (2001). "Is heat hot?" Inducing conceptual change by integrating everyday and scientific perspectives on thermal phenomena'. *Learning and Instruction* 11 (4-5), pp. 331–355. [https://doi.org/10.1016/s0959-4752\(00\)00036-0](https://doi.org/10.1016/s0959-4752(00)00036-0).
- Woolnough, B. E. (1994). 'Factors affecting students' choice of science and engineering'. *International Journal of Science Education* 16 (6), pp. 659–676. <https://doi.org/10.1080/0950069940160605>.
- Zoldosova, K. and Prokop, P. (2005). 'Analysis of Motivational Orientations in Science Education'. *International Journal of Science and Mathematics Education* 4 (4), pp. 669–688. <https://doi.org/10.1007/s10763-005-9019-2>.

Authors

Helena Thuneberg is Docent (Adjunct professor) in Special pedagogy, Researcher in Science centre education and in Centre of Educational Assessment, in Department of Education, Faculty of Educational Sciences, University of Helsinki. She has authored several scientific articles relating to informal science education, in addition to articles related to special education and school assessment. In her works she applies motivational concepts based on the Self-Determination theory. E-mail: helena.thuneberg@helsinki.fi.

Hannu Salmi is currently Professor in Science centre education in Department of Education, Faculty of Educational Sciences, University of Helsinki. As Research director he has been involved in many EU Science education projects and authored several scientific articles relating to informal science education. E-mail: hannu.salmi@helsinki.fi.

How to cite

Thuneberg, H. and Salmi, H. (2018). 'To know or not to know: uncertainty is the answer. Synthesis of six different science exhibition contexts'. *JCOM* 17 (02), A01. <https://doi.org/10.22323/2.17020201>.



© The Author(s). This article is licensed under the terms of the Creative Commons Attribution — NonCommercial — NoDerivatives 4.0 License. ISSN 1824-2049. Published by SISSA Medialab. jcom.sissa.it