Making the invisible observable by Augmented Reality in informal science education context

Hannu Salmi, Helena Thuneberg & Mari-Pauliina Vainikainen


To link to this article: http://dx.doi.org/10.1080/21548455.2016.1254358

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 25 Nov 2016.

Submit your article to this journal

Article views: 1487

View related articles

View Crossmark data

Citing articles: 1 View citing articles
Making the invisible observable by Augmented Reality in informal science education context

Hannu Salmi, Helena Thuneberg and Mari-Pauliina Vainikainen

ABSTRACT
The aim of the study was to analyse learning using Augmented Reality (AR) technology and the motivational and cognitive aspects related to it in an informal learning context. The 146 participants were 11- to 13-year-old Finnish pupils visiting a science centre exhibition. The data, which consisted of both cognitive tasks and self-report questionnaires, were collected using a pre- post-test design and were analysed by SEM path-analysis. The results showed that AR-technology experience was beneficial for all, but especially for the lowest-achieving group and for the girls. In general, pre-knowledge skills predicted post-knowledge test results. As expected, school achievement had an effect on pre-knowledge results. In addition, motivation turned out to be an alternative key route for learning. Being a boy predicted directly or indirectly all other motivational variables, enjoyment and interest, but girls had a higher relative autonomy experience (RAI). Situation motivation and attitude towards learning in the science exhibition were much more strongly inter-connected among boys than girls, and attitude predicted post-knowledge only for boys. AR seems to be a promising method by which to learn abstract phenomena using a concrete manner.

ARTICLE HISTORY
Received 4 May 2016
Accepted 25 October 2016

KEYWORDS
Augmented Reality; AR; informal learning; intrinsic motivation; SDT-theory; situational motivation

Introduction
Information and communication technologies have profoundly altered our everyday lives. For more than two decades, they have been considered as highly promising tools for improving education. Although the amount of publications related to qualitative and quantitative results has been growing rapidly, the demand and challenge published in the thematic issue of Science (1/2009) is still more than actual: Making a Science of Education demanded that a great deal of high-quality research should be performed by focussing on the utilisation and effects of the new technologies in both school and informal learning environments (Alberts, 2009; Rennie, 2014; Salmi, Kaasinen, & Kallunki, 2012). This study aims to contribute to the research literature of Augmented Reality (AR) both in practice and theory. The focus is in bridging the gap between formal education and informal learning.

The original definition of AR has been summarised for our purposes with four simultaneously required elements (Azuma, 1997): (a) a mixture of real-world and digital illusions, (b) interactivity with the real environment, (c) on-line effects and (d) three-dimensional impressions (Martin et al.,...
2011; Salmi et al., 2012). As in earlier research reports, there have been lots of opportunities to apply these technologies: ‘The more a virtual immersive experience is based on design strategies that combine actional, symbolic, and sensory factors, the greater the participant’s suspension of disbelief that she or he is “inside” a digitally enhanced setting’ (Dede & Barab, 2009, p. 66). However, only a few applications have been created and orientated towards educational purposes. In particular, there have still only been a few teaching applications of AR (Andújar, Mejías, & Márquez, 2011; Salmi et al., 2012), but they have started to grow fast (Bacca, Baldris, Fregnat, Graf, & Kinshuk, 2014; Bower, Howe, McCredie, Robinson, & Grover, 2014; Radu, 2014).

‘Mixed Reality’ has become the widely used term for combining AR with the environments of real objects. The separate environments of the real world and the virtual world form a reality – virtuality continuum of mixed reality (see Figure 1). While viewing from the very ends of the continuum, augmented reality and augmented virtuality appear. This continuum has been extended into a two-dimensional plane of ‘Virtuality’ and ‘Mediality’. This already classical continuum (Milgram & Kishino, 1994) can be derived as an educational model as has been shown in Figure 2. The model below describes formal education and informal learning from a new angle by adding the real ↔ virtual-dimension. The figure shows that different types of virtual learning solutions have brought a lot of new learning materials especially to informal learning settings. However, AR has become more common in current applications because not every little detail of reality has been needed to be modelled as they are already presented (Andújar et al., 2011; Salmi et al., 2012).

AR has been utilising the already existing environment and has added new visual information to it. According to its broadest definition, AR may be considered a term that describes technologies that allow the real-time mixture of computer-generated digital content and the real world (Haller,Billinghurst, & Thomas, 2007).

AR’s first utilisation – like many new innovations – only came about in military and high-tech solutions demanding big investments. Boeing inventor and technician Thomas Caudell has been attributed to having been the very first ‘civil’ solution in the beginning of the 1990s; at that time he started to use it as a new term to describe how the head-mounted displays that electricians use when assembling complicated wiring harnesses work.

Since then the AR has been a trendy technology which attracts more and more attention from HCI-Human Computer Interaction planners, designers and researcher (Salmi, Sotiriou, & Bogner, 2010). Today, trendy i-glasses, heads-up displays and car windshields with data have become perhaps the most well-known consumer AR products, but the technology has been used in many industries, including healthcare, public safety, gas and oil and tourism and marketing.

**Reported learning results by using AR-solutions**

One indicator of the AR boom in education is that there have been several recent meta-analyses about the topic. One such analysis is Bacca et al. (2014) meta-study ‘Augmented Reality Trends in Education: A Systematic Review of Research and Applications’, which analysed more than 30 research reports from 2003 to 2013. The review partly described the current state of AR research in education. Besides the main outcomes, this meta-study showed that there is a lack of review studies which focus on investigating factors such as the uses, advantages, limitations, effectiveness,
challenges and features of AR in educational settings. Personalisation for promoting inclusive learning using AR has also become a growing area of interest.

In Radu’s (2014) ‘Augmented reality in education: A meta-review and cross-media analysis’, 26 articles were selected for analysis. The articles mainly compared AR vs. non-AR solutions. Thus, the approach provided both a critical literature review and at the same time identifying the outcomes that AR can give learners. The positive outcomes were as follows: increased content understanding, learning spatial structure and function, learning language associations, retention in long-term memory, improved physical task performance and increased student motivation. On the other hand, the negative outcomes were as follows: attention tunneling, usability difficulties, ineffective classroom integration and learner differences.

The abovementioned meta-analyses by Bacca et al. (2014) and by Radu (2014) have only covered a very small part of the research area. Much of the reported research has been very pragmatic as it often contains research and development elements and has sometimes been also conducted and published outside the traditional academic field. It has, however, offered a firm starting point for further analysis. The meta-studies (Bacca et al., 2014; Radu, 2014) discussed also the trends and the vision of the future and opportunities for further research in AR for educational settings. As such, for this present research, we have brought three elements to the discussion, which have so far been underrepresented. We have focused on the following: (1) to bridge the gap between informal learning and formal education, (2) to take into account different learners, especially (3) the possible gender effects, (4) to analyse the learning effect of AR in pre- and post-design and (5) to find pathways for changing the strong interest and external situation motivation into intrinsic, content-based motivation.

Authenticity of science learning (Braund & Reiss, 2007) has been found to be an effective motivating factor. AR educational solutions have been trying to use it as the starting point while designing AR simulations. In a study (Rosenbaum, Klopfer, & Perry, 2007) related to role playing by AR, the topic to be learned was disease transmission that the students could affect through their own actions. According to the results, this kind of ‘participatory reality’ game has made new kinds of authentic

![Figure 2. Persistent dichotomies or blurring boundaries? (Salmi, Kaasinen & Kallunki, 2012).](image-url)
science inquiry experiences possible. This approach has become known globally during the Pokémon GO game trend in 2016. Furthermore, visual and immersive interfaces have been offering new approaches for utilising the visual reasoning ability to transfer learning processes from out of the classroom into real and authentic, open learning environments (Salmi et al., 2010).

There has been evidence that virtual learning environments promote higher learning results (Vilkonienė, Lamanauskas, & Vilkonis, 2008). However, in the case of learning natural sciences, either real learning itself or a combination of real and virtual environments (Bilek, Zemanova, & Turcani, 2007) has been preferred (Lamanauskas & Vilkonis, 2007).

AR has shown to improve learning achievement (Chiang, Yang, & Hwang, 2014). More specifically, several reports (Radu, 2014; Vilkonienė, 2009) have found that AR solutions could help gifted learners, but the results are somewhat mixed, as, for example, Freitas and Campos (2008) claim that high-achievers do not especially gain from AR-technology. Pupils with low motivation as well as pupils with lower school achievement have shown to benefit from the use of AR (Bower et al., 2014; Freitas & Campos, 2008; Nincarean, Bilal Alia, Halim, & Rahman, 2013; Salmi et al., 2010).

Science centre exhibitions and informal learning settings have been found to be more gender equal than formal classroom education settings (Osborne & Dillon, 2008; Salmi, Thuneberg, & Vainikainen, 2016). Video games and computer-based entertainment and serious pc-educational games (cf. Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012) have traditionally been more beneficial for the boys. In addition, results by Echeverría, Gil, and Nussbaum (2011) relating to educational use of AR-technology indicate that boys benefit from AR more than girls, as they performed better in the knowledge post-test.

In one of the few reports related to teacher education and teacher professional development (Salmi et al., 2012), questionnaire data of 292 teachers were statistically analysed. The main outcomes were as follows: moving from a teacher-controlled learning environment towards a pupil-orientated learning one, connecting ICT-AR with and between learning environments, seeing remarkable changes in the roles and responsibilities of students and teachers towards student centeredness, autonomy, and peer collaboration. Teachers were also seeing structural changes in the learning environments.

**Informal learning, open learning environments, science centre, and AR exhibits**

Learning in informal science education contexts has been far less studied than science education in traditional school settings. The present study has focused on both the cognitive and the motivational aspects of learning using AR exhibits as part of the science centre visit. Using this type of technology as mixed reality of concrete and virtual objects has become more common in the modern science exhibitions recently (Kamarainen et al., 2013; Martin et al., 2011).

The main results about informal learning and especially about science centre education have been related to the effects of motivation on learning (Falk & Dierking, 1992; Fenichel & Schweingruber, 2010; Osborne & Dillon, 2008; Salmi, 2003, 2011; Salmi et al., 2016, Salmi, Vainikainen, & Thuneberg, 2015; TanWee Hin & Subramaniam, 2003). Learning from informal sources and in an out-of-school environment has been found to be effective and motivating (Fenichel & Schweingruber, 2010; Osborne & Dillon, 2008).

This quantitative study has been based on the theories and literature findings bridging the gap between formal education and informal learning. A huge amount of information, especially about modern phenomena like the recent developments in mathematics, science and the new technologies, has been obtained in a personal way from family, friends and peer groups. Furthermore, the roles of television, libraries, magazines and newspapers have also been essential for public understanding of science. Museums and science centres have regularly had more visitors during the last decades. Despite this development, there has been less theoretical or empirical research in the informal sector (Osborne & Dillon, 2008). ICT- and web-based learning have challenged and changed the vision of traditional formal education (Salmi, 2011, 2012).
**Interest, situation and context of learning**

Researchers have shown that there is significant overlap between interest and motivation (Durik & Harackiewicz, 2007). It has been claimed that in the science centre learning context interest could be the link bridging situation and intrinsic motivation (Salmi, 2003). On the one hand, interest-based, self-determined forms of learning motivation provide the most favourable learning outcomes also in relation to use of AR-technology (Antonioli, Blake, & Sparks, 2014), and on the other hand the organisation of the learning environment can spark interest before personal motivation develops (Fenichel & Schweingruber, 2010; Renninger & Hidi, 2011). Thus, out-of-school learning – formal-like teaching in an informal context – can provide an excellent opportunity for enhancing pupils’ interest and motivation towards learning science and mathematics (Vainikainen, Salmi, & Thuneberg, 2015).

Novelty has been the key factor in creating the interest and situational motivation in science centres. The shock of a new setting can be addressed by the so-called ‘sniff around corners’ method (Balling & Falk, 1980). It works also as a ‘head-start’ for the learning process. It has also a clear link to intrinsic learning motivation. The content of an exhibition is essential: meaningful experiences are closely connected with a positive development of interest in the subject. Of course, the starting point of this process of growing interest lies within the person, but the content and interaction define the development of situational and individual interests (Hidi & Renninger, 2006). The conflict between earlier knowledge and capacities and the goals that can be reached by own learning activities start this process (Salmi, 2003). However, the informal learning settings might give strong stimuli and thus motivate also the students who otherwise would not be ‘academically’ orientated. Those pupils, who like what they have learnt, are then apt to develop an autonomous learning motivation (Hsiao, Chen, & Huang, 2012; Wagner, 2008), thus cultivating the more superficial interest into a deeper desire to continue learning.

**Description of the AR-exhibitions of the present study**

The pupils used the exhibits ‘Doppler Effect’ and ‘Molecule Movement in Gas’, the AR-application which enabled them to see something more than is possible by ordinary experiments (Salmi, 2012). While investigating thermal motion, the velocity of molecular nitrogen for example in a refrigerator or on a hotplate can be compared both by following the motion of augmented molecules in different places and also by comparing the different velocities of the molecules on a temperature-velocity graph (Gargalakos, Rogalas, Lazoudis, & Sotiriou, 2012). In a similar way, also the exhibit ‘Doppler Effect’ described for the pupils three-dimensionally the phenomenon of mixing the graphical elements with the original physical phenomenon: a mini-fire-truck passed the listener on the side of a road, then the listener heard first a higher sound when the fire truck approached, and after passing the sound heard is lower (Owen, Owen, Barajas, & Trifonova, 2012). In the case of this experiment the augmented information was the circular wave pattern that represents the 3-D spherical pressure wave (Giallouri, Lazoudis, & Sotiriou, 2012). The wave pattern changes according to the Doppler effect when the fire truck starts to drive (Larsen & Bogner, 2012). Also the theoretical and mathematical patterns and formulas have been included in visual screens, which can be seen by the users (Buchholz & Brosda, 2012). See the exact description of the project ‘Guide of Good Practice’ ed. by Lazoudis, Sotiriou, & Salmi, 2012.

**Research questions**

The referenced literature in the previous pages suggests that AR might be beneficial, for example, for pupils with special needs, gender, or pupils with low motivation, the goal is to test whether these assumptions would hold. The focus is to test motivational aspects and to analyse learning results relating to AR of the differently achieving (school success) pupils.

The aim of the study was to find out how pupils learn when using AR technology in a science centre context. The specific research questions were:
Q1: How did the contextual (formal vs. informal learning context) variables predict learning?

H1: The more interest pupils show towards the exhibition context, the better cognitive learning results they will receive.

Q2: How did the motivational variables predict learning?

H2: Situational motivation supports cognitive learning outcomes.

H3: Pupils intrinsic motivation and autonomy experience have an effect on cognitive learning in the exhibition.

Q3: What is the role of gender in the learning process?

H4: Boys gain more by using AR

H5: Girls have stronger autonomy experience.

Q4: What was the role of visual reasoning ability and school achievement in the learning process?

H6: Lower achieving pupils benefit from the AR technology the most.

H7: Visual reasoning scores predict pupils' knowledge learning results.

All these questions were analysed in relation to gender and school achievement.

**Method**

**Participants**

The participants consisted of 146 pupils, of which 51% (n = 75) were girls. The pupils were in average 12.3 years old. From the 32 schools in the Helsinki metropolitan area in Finland visiting the science centre exhibition, 7 schools were randomly selected.

**Venue**

The pupils visited a typical science centre exhibition, which was supported and fulfilled by five AR-technology supported exhibits. They were the (1) Doppler phenomenon, (2) Boltzmann’s molecule movement, (3) Young Experiment, (4) Aeroplane Mini Wing exhibit and (5) Rolling Double Cone.

**Measures**

Table 1 presents an overview of the design of the study, the measures and timing.

**Deci and Ryan autonomous motivation**

We used self-determination theory (SDT) to test the autonomous motivation. It was only administered as a pre-test because after a short visit, there should not be any major changes in the overall motivation, which is related to the whole personality.
The Deci-Ryan Motivation (SRQ-A: Self-Regulation Quality – Academic) test includes 32 standardised items on a Likert scale (1 = not at all true, 2 = not nearly true, 3 = somewhat true, 4 = totally true). The summative variables on the self-determination continuum are located in the order from external towards intrinsic: external, introjected, identified and intrinsic.

The SRQ-A test includes a formula (Ryan & Connell, 1989) by which the Relative Autonomy Index (RAI) can be calculated. In this study, the RAI was used as the indicator of the level of autonomous motivation. The positive sign in RAI indicates that the experience is rather autonomous, and the negative one that it is more externally controlled and relies more on adults or others than on one-self.

The reliability of the test was $\alpha = .92$ (32 items).

**Situation motivation**

Situation motivation was measured with a questionnaire consisting of 14 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 pupils experienced the exhibition. An example of an item: ‘Visiting the exhibition with AR was fun’. The reliability of the test was $\alpha = .91$ (14 items).

**Raven test**

The cognitive measure was a visual reasoning and learning capacity test: Raven Standard Progressive Matrices (Raven, Raven, & Court, 2003). This test has been widely utilised in both practice and theoretical research. The cumulative amount of data in different age groups, schools and socio-economic cultures has shown the Raven test to be a valid tool. In addition, it does not give gender advantage for either male or female pupils.

According to the test theory (Raven, Raven, & 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. This Raven test acts as the measurement of non-verbal based cognitive skills. It gives a reliable standardised tool for comparing the individuals’ learning abilities compared to the representative age group.

In each test item, the subject is asked to identify the missing element that completes a pattern. The test contains 60 items that have been divided into five sets (A, B, C, D, E). Each of these groups contains 12 different tasks. The reliability of the test was $\alpha = .79$ (60 items).

**Knowledge test**

The knowledge test consisted of 31 items related to the content areas of the school curriculum combined with the AR-solutions and the science exhibition. The questions were piloted two months before the actual pre-knowledge test. Three items were omitted in this phase as irrelevant, as they turned out to be too easy or all too difficult for the majority of the pupils. The answering options to the statements in the test were 1 = true, 2 = untrue and 3 = I don’t know. In the present study, only the number of correct answers was analysed. An example of an item: ‘While the air is heated, the gas molecules start to move more quickly’.

The post-knowledge tests (exactly the same as the pre-tests) were conducted one week after the science exhibition visit. The reliabilities of the pre-test and the post-test were $\alpha = .72$ and $\alpha = .77$ (31 items), respectively.

**School achievement**

The school achievement variable was the summary of the four school grades (Physics, Chemistry, Mathematics and Mother tongue) based on information obtained from the teachers. The pupils were classified into three categories according to their school achievement level:
A+ = Above Average School Achievement (A+; 25% of the pupils in each class),
A = Average School Achievement (A; 50% of the pupils in each class),
A− = Below Average School Achievement (A−; 25% of the pupils in each class).

**Learning context: school vs. AR in science centre**

The Semantic Differential method (Osgood, 1964) was used for measuring pupils’ interest towards learning science in school settings and in the science centre. They had to evaluate 14 pairs of adjectives on a five-point scale (e.g. studying science in school is (5) interesting to (1) boring). The reliability of the school science interest scale (pre-test) was $\alpha = .81$ (14 items), and $\alpha = .88$ (14 items) for the science centre science (post-test). The questions were the same, but the pupils had to relate their attitudes in the pre-test with the school science learning context and in the post-test with the science centre learning context.

**Analysis methods**

We used the path modelling (AMOS 22) to find out how the observed data would support the theoretical assumptions and set hypotheses. However, in order to enhance interpretation of the path-analysis results, it was necessary to start the analysis by other methods. First we compared the simple pre- and post-knowledge test means. The differences between the groups (boys vs. girls, and pupils with low vs. average vs. high school achievement) was tested by GLM Multivariate analysis of variance (MANOVA). In addition, the GLM repeated measures was applied for revealing the significance of the change. The effect size measure was partial $\eta^2$ (interpretation: >.01 small, >.06 middle, >.14 large (Cohen, 1988; Cohen, Miles, & Shevlin, 2001).

By the path modelling method, we could then identify the direct and indirect effects of the studied variables, the direction of the effects and, in addition, the size of the effects. Gender, RAVEN, RAI and school achievement were used as covariates to control their effects on the measured pre- and post-knowledge test variables and on situational motivation (only as a post-test) and attitudes towards learning in the school learning context and science centre learning context. The goodness-of-fit evaluation of the models was based on a $\chi^2$-test and NFI, CFI (good fit > .90, or better > .95) and RMSEA (reasonable fit < .08, good fit < .05) (see Byrne, 2010). Parametric bootstrapping was applied to obtain also the indirect effects and to test invariance of the models of gender and school achievement groups (indicator: $\chi^2$-tests the difference between the unconstrained and fully constrained models). In case the models were not invariant, pair-wise comparison of the regression weights (paths) between the models was analysed (indicator: $z$-test).

**Missing values**

There were 3.0–5.4% missing values depending on the tested variable. The list-wise method was used to remove the cases with missing values in the SEM, because the bootstrapping method in AMOS requires it.

**Results**

The analysis by the GLM MANOVA showed that there were no interaction effects of the gender and school achievement group on the knowledge-tests, which allows the simple analysis of the main effects of gender and school achievement group separately. The exact statistical descriptives are shown in Appendix 1 and Appendix 2. Gender had a significant effect on the pre-knowledge test, $[F(df = 1) = 11.83, p < .01, \text{partial } \eta^2 = .08]$, but the effect was non-significant in the post-knowledge test ($p = .055$; see Figure 3).
The school achievement group had an effect on the pre-knowledge test \( F(df = 2) = 7.37, p < .01, \) partial \( \eta^2 = .09 \), but the effect was non-significant in the post-knowledge test \( (p = .08) \). In the pre-knowledge test, the post hoc analyses revealed statistically significant differences between the lowest achieving group and the average \( (p < .01) \) and the highest \( (p < .001; \) see Figure 4).

**Figure 3.** Pre and post knowledge-test 95% confidence intervals of boys and girls.

The school achievement group had an effect on the pre-knowledge test \( F(df = 2) = 7.37, p < .01, \) partial \( \eta^2 = .09 \), but the effect was non-significant in the post-knowledge test \( (p = .08) \). In the pre-knowledge test, the post hoc analyses revealed statistically significant differences between the lowest achieving group and the average \( (p < .01) \) and the highest \( (p < .001; \) see Figure 4).

**Figure 4.** Pre and post knowledge-test 95% confidence intervals by school achievement group (below average, average, above average).
GLM Repeated measures test showed that overall (with no intervening variables) the change in knowledge was significant [Wilk’s lambda $F(df = 1) = 16.79, p < .001$, partial $\eta^2 = .10$]. The predictors of the change were then analysed by the path model presented in Figure 5 in order to take the cognitive and motivational direct and indirect effects into account. The final model with only significant paths fitted the data well ($\chi^2 = 23.02$, $df = 25$, $p = .58$; NFI = .94, CFI = 1.00, RMSEA = .00).

In the following, the statistical path-model results are explained first. The meaning and conclusions of the results are discussed in the Discussion and conclusions section.

The path-model results are presented by showing the significant direct effects and their directions by arrows in the Figure 5. In addition, the total effect percentage of all the involved direct effects is shown in relation to each variable in concern. The indirect effects are explained in the text by showing the standardised betas and the significance level. After the path-model, the possible moderation effects are analysed by comparing whether the same model would fit (a) to both the boys and girls, and (b) to the three school achievement groups.

The path model in Figure 5 showed that three variables predicted pre-knowledge directly: school achievement, school learning context and gender. Being a boy (boy coded as 0, girl as 1) predicted higher performance. Gender predicted knowledge also indirectly ($\beta = -.06$, $p < .01$). The exogenous variables together explained a fourth of the variance of the pre-knowledge test. The descriptives of the pre- and the post-tests are presented in Appendix 1 and Appendix 2.

RAI and gender predicted interest in learning science in school learning context by explaining 19% of its variance. Situation motivation at the science centre was directly predicted only by the interest in learning science in the school learning context. However, RAI had a small indirect effect on situation motivation, ($\beta = .14$, $p < .01$), and being a boy enhanced it indirectly, as well ($\beta = -.09$, $p < .01$).

Interest in learning science in the science centre learning context was largely and directly predicted by situation motivation. Overall, 71% of its variance was explained by the model and, in addition to

**Figure 5.** The final path model.
situation motivation, it was also indirectly explained by gender ($\beta = -0.08, p < .01$), RAI ($\beta = 0.12, p < .01$) and school learning context ($\beta = 0.29, p < .01$).

The variables of the model explained 28% of the post-knowledge test results. In addition to the direct effects of the pre-knowledge test and interest in science in the science centre learning context, there were several indirect effects: school achievement ($\beta = 0.15, p < .01$), gender ($\beta = -0.16, p < .01$), RAI ($\beta = 0.06, p < .01$), school learning context ($\beta = 0.16, p < .01$) and situation motivation ($\beta = 0.18, p < .01$). Cognitive competence as measured by Raven had no direct effects of its own, but in the model its role shows as a significant covariance between it and school achievement.

Next, the possible moderation effects of the gender and school achievement group were tested in order to find out whether the same model would fit across groups. Indeed, the models of boys and girls did differ. The difference between the unconstrained and fully constrained model was $\chi^2$ difference = 27.21, $df = 7$, $p = .000$. Two individual paths were different. The regression coefficient of situation motivation on interest in learning science in the science centre context was positive in both groups, but the relation was much stronger for boys ($z = -4.76, p < .01$). The second difference, the regression coefficient of interest in learning science in the science centre context on the post-knowledge test was not significant in the girls’ group, which was in contrast to the boys’ result ($z = -2.24, p < .05$).

The models of the school achievement groups were invariant, indicating no significant difference between the unconstrained and fully constrained model ($\chi^2$ difference = 18.16, $df = 16$, $p = .315$). Pre-knowledge was a significant predictor of post-knowledge in all school achievement groups (A−, $p < .001$, A $p < .001$, A+ $p < .05$), but the paths differed in that the regression coefficient was significantly stronger in the lowest achieving group than in the highest achieving group ($z = -2.59, p < .01$). Being a boy was related with higher interest in learning science in the school context more in the lowest achieving group than in the average achieving group ($z = 2.19, p < .05$).

**Discussion and conclusions**

The question related to the science centres is whether they are capable of managing to orient and enhance the momentary, strong situation motivation into a long-lasting intrinsic motivation. This is also one of the biggest challenges for open learning environments, such as science centres (Falk & Dierking, 1992; Rennie, 2014; Salmi, 2003; Salmi et al., 2010; Salmi et al., 2015, Salmi et al., 2016).

The post-tests were administrated one week after the exhibition visit. Of course, there are several factors which – theoretically – might have caused the learning effect. However, controlling these everyday experiences as variables is nearly impossible. In addition, if the pupils have gained more knowledge from other sources, it is a positive outcome too. This type of extra curriculum learning indicates an increase in intrinsic motivation, which is an important goal for bridging the gap between formal education and informal learning.

**Contextual variables predict learning**

Previous research shows that positive effects of informal environment on learning gains are not self-evident, but they vary according to the pupils’ prior interest in science and their readiness to take responsibility of setting their own goals (Renninger, 2000). The results of this study provide evidence that interest in school science is enhanced by situation motivation in the science exhibition context. This leads to better cognitive learning results in the post-knowledge test, which confirms the first hypothesis (H1). The results also show that in general pupils really learn in the science centre exhibition as their overall science knowledge scores increase after the AR experience. This result is in line with earlier findings (Bacca et al., 2014; Radu, 2014). We could also show that the learning effect was mid-size.
Two routes resulting to better performance outcomes

The present study indicates that there were two routes which enhance learning as measured by the post-test knowledge scores – a stronger one that depends on prior knowledge and another less effective one that depends on interest and motivation. Knowledge before the exhibition is a moderate direct explainer of the post-test results, but interest in learning in the science centre context predicts the outcomes directly, as well. School achievement, gender and autonomy experience, interest in learning science in the school context and situation motivation during the science centre experience predict the knowledge results indirectly to some extent. By these results, the second (H2) and third hypothesis (H3) were confirmed: both situational motivation and autonomy experience support learning. However, as shown, there are other factors involved, and both indirect and direct effects.

No gender gap after the AR-experience – somewhat different processes

Before the AR exhibition, boys had higher science knowledge scores than girls. However, in the post-tests, gender shows no direct effect anymore, which implies that then the boys and girls learned equally much. By this result, the fourth hypothesis (H4) cannot be confirmed. Thus, the results contradict previously found results.

The higher pre-test results of boys are partly due to the direct and indirect effect of boys’ higher interest in science learning at school. It is of interest to note that being a boy is generally related to having higher direct or indirect effects on all other motivational variables. However, in regard to RAI, being a girl is beneficial, which confirms the fifth hypothesis (H5), and implies that the non-significant difference in the post-knowledge test results between boys and girls can be partly attributable to the higher autonomy experience of girls.

Comparison of the girls’ and boys’ models shows that the coupling with situation motivation and interest in learning in the science centre is very strong among boys, and much more powerful than that of girls. It implies that if boys are situationally motivated and feel inspired there, they become more interested in science centre learning in general. That positive attitude enhances the knowledge-test results afterwards. On the other hand, the lower their situation motivation, the less interesting the informal learning option is perceived, and so these detrimental attitudes result in weaker knowledge outcomes. That attitude has no effect on the knowledge results of girls, but it does on the results of boys. This implies that intervention on the attitudes of boys would be a way to enhance their learning and prevent underachievement. On the other hand, it would be important to find out why there is no link between the girls’ interest in learning in the science exhibition and their learning outcomes.

Low-achievers gained from AR-technology: visual reasoning ability in indirect role

The concrete AR-learning session is most beneficial for the lowest achieving pupils, who do not know as much as the others before the exhibition. This is despite the higher achieving pupils also learning during the exhibition. This is an encouraging result for developing new methods for pupils in drop-out risk, and confirms the sixth hypothesis (H6). However, also the best-performing pupils were receiving new, better learning results. This is also in harmony with the earlier findings (Bacca et al., 2014) of personalisation for promoting an inclusive learning using AR is also a growing area of interest. Unlike it was assumed in the seventh hypothesis (H7) based on research literature, visual reasoning scores did not predict knowledge results.

Within the lowest achieving group, high pre-test results predict post-test results more than in the other achievement groups. That is, when a pupil has low school grades but good pre-test knowledge results; those results are of higher importance for the post-test results than they are in the groups that have higher school grades. Higher school achievement predicts neither interest in learning science at school nor the arousal of situation motivation nor interest in the science centre. It is interesting that
the path by path analysis reveals that being a boy is related with higher interest in learning science at school more in the lowest achieving group.

AR-technology itself has and will develop rapidly. However, during the last five years the development has concentrated on producing applications for tablets and smartphones. Meanwhile, only a few innovations exist of the optical head-mounted displays, and the few existing glasses are on the prototype level. The big advantage when comparing head-mounted displays to tablets and screens is that the glasses leave users’ hands free to do concrete things, which is exactly in keeping with Thomas Caudell’s original definition of AR.

Radu (2014) names two weaknesses of using AR in education: (1) ineffective classroom integration and (2) learner differences. But, actually, we use the re-framing approach (Mattila, 2001) to change the ‘weaknesses’ into strengths. The ‘ineffective classroom integration’ gives the opportunity for especially out-of-school methods to bridge the gap between formal education and informal learning via AR. Also, the ‘learner differences’ is also an important factor. AR clearly offers new types of learning paths for different types of pupils. AR seems to be turning out to be one of the very few and rare pedagogical solutions that benefits those pupils who are below average in traditional school success and achievement level.

Notes
1. AR used for teaching purposes of symbolic associations like the meaning of written words (Freitas & Campos, 2008).
2. Focusing attention to a particular channel of information, which might lead to neglect of relevant events on other channels (Wickens & Alexander, 2009).
3. Note! All found effects were positive. However, in case of gender, the minus-sign indicates significant effect of boys, because boy was coded as 0, girl as 1.

Disclosure statement
No potential conflict of interest was reported by the authors.

References


Appendix 1

Table A1. Descriptive statistics for gender groups.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Raven</th>
<th>RAI</th>
<th>SitMot</th>
<th>Sci_School</th>
<th>Sci_Centre</th>
<th>Pre_Know</th>
<th>Post_Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boy 71</td>
<td>M 37.17</td>
<td>−54</td>
<td>3.44</td>
<td>3.55</td>
<td>3.47</td>
<td>16.65</td>
<td>17.63</td>
</tr>
<tr>
<td></td>
<td>SD  5.71</td>
<td>1.76</td>
<td>.87</td>
<td>.56</td>
<td>.85</td>
<td>3.49</td>
<td>4.33</td>
</tr>
<tr>
<td></td>
<td>Min  22</td>
<td>−4.13</td>
<td>1.29</td>
<td>2.15</td>
<td>1.29</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Max   50</td>
<td>4.40</td>
<td>4.71</td>
<td>4.71</td>
<td>4.71</td>
<td>24</td>
<td>26</td>
</tr>
<tr>
<td>Girl 75</td>
<td>M 38.72</td>
<td>.09</td>
<td>3.43</td>
<td>3.33</td>
<td>3.40</td>
<td>14.63</td>
<td>16.32</td>
</tr>
<tr>
<td></td>
<td>SD  5.76</td>
<td>1.79</td>
<td>.69</td>
<td>.51</td>
<td>.48</td>
<td>3.61</td>
<td>3.88</td>
</tr>
<tr>
<td></td>
<td>Min  22</td>
<td>−3.44</td>
<td>1.64</td>
<td>1.79</td>
<td>2.36</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Max   52</td>
<td>4.67</td>
<td>4.86</td>
<td>4.57</td>
<td>4.43</td>
<td>23</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>N 146</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M 37.97</td>
<td>−21</td>
<td>3.43</td>
<td>3.44</td>
<td>3.44</td>
<td>15.61</td>
<td>16.96</td>
</tr>
<tr>
<td></td>
<td>SD  5.77</td>
<td>1.79</td>
<td>.78</td>
<td>.54</td>
<td>.68</td>
<td>3.68</td>
<td>4.14</td>
</tr>
</tbody>
</table>

*p < .05.
**p < .01.
***p < .001.

Appendix 2

Table A2. Descriptive statistics for school achievement groups.

<table>
<thead>
<tr>
<th>Sch_Achiev</th>
<th>Raven</th>
<th>RAI</th>
<th>SitMot</th>
<th>Sci_School</th>
<th>Sci_Centre</th>
<th>Pre_Know</th>
<th>Post_Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowest N = 32</td>
<td>M 34.38***</td>
<td>−.26</td>
<td>3.38</td>
<td>3.50</td>
<td>3.33</td>
<td>13.66***</td>
<td>16.16</td>
</tr>
<tr>
<td></td>
<td>SD  5.85</td>
<td>1.79</td>
<td>.73</td>
<td>.67</td>
<td>.71</td>
<td>3.51</td>
<td>4.23</td>
</tr>
<tr>
<td></td>
<td>Min  22</td>
<td>−4.13</td>
<td>1.64</td>
<td>1.79</td>
<td>1.43</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Max   48</td>
<td>4.21</td>
<td>4.79</td>
<td>4.71</td>
<td>4.43</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td>Average = 82</td>
<td>M 38.21</td>
<td>−3.8</td>
<td>3.43</td>
<td>3.39</td>
<td>3.47</td>
<td>15.85</td>
<td>16.73</td>
</tr>
<tr>
<td></td>
<td>SD  5.14</td>
<td>1.76</td>
<td>.83</td>
<td>.52</td>
<td>.71</td>
<td>3.27</td>
<td>4.20</td>
</tr>
<tr>
<td></td>
<td>Min  25</td>
<td>−3.67</td>
<td>1.29</td>
<td>2.15</td>
<td>1.29</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Max   50</td>
<td>4.67</td>
<td>4.86</td>
<td>4.57</td>
<td>4.71</td>
<td>23</td>
<td>26</td>
</tr>
<tr>
<td>Highest N = 32</td>
<td>M 40.94</td>
<td>.26</td>
<td>3.49</td>
<td>3.50</td>
<td>3.45</td>
<td>16.94</td>
<td>18.34</td>
</tr>
<tr>
<td></td>
<td>SD  5.48</td>
<td>1.85</td>
<td>.72</td>
<td>.46</td>
<td>.58</td>
<td>4.14</td>
<td>3.69</td>
</tr>
<tr>
<td></td>
<td>Min  29</td>
<td>−2.48</td>
<td>1.64</td>
<td>2.64</td>
<td>2.14</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Max   52</td>
<td>4.40</td>
<td>4.71</td>
<td>4.57</td>
<td>4.64</td>
<td>24</td>
<td>26</td>
</tr>
<tr>
<td>Total N = 146</td>
<td>M 37.97</td>
<td>−.21</td>
<td>3.43</td>
<td>3.44</td>
<td>3.44</td>
<td>15.61</td>
<td>16.96</td>
</tr>
<tr>
<td></td>
<td>SD  5.77</td>
<td>1.79</td>
<td>.78</td>
<td>.54</td>
<td>.68</td>
<td>3.68</td>
<td>4.14</td>
</tr>
</tbody>
</table>

Note: pair-wise post hoc comparisons: Raven lowest vs. average**, vs. highest***; average vs. highest*. Pre-knowledge lowest vs. average**, vs. highest**.

*p < .05.
**p < .01.
***p < .001.