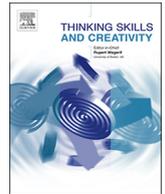




ELSEVIER

Contents lists available at ScienceDirect

## Thinking Skills and Creativity

journal homepage: [www.elsevier.com/locate/tsc](http://www.elsevier.com/locate/tsc)

# How creativity, autonomy and visual reasoning contribute to cognitive learning in a STEAM hands-on inquiry-based math module

H.M. Thuneberg<sup>a</sup>, H.S. Salmi<sup>a</sup>, F.X. Bogner<sup>b</sup>

<sup>a</sup> University of Helsinki, Faculty of Educational Sciences, FI-00014, Helsinki, Finland

<sup>b</sup> University of Bayreuth, Z-MNU (Centre of Science and Math Education), Chair of Biology Education, University Campus, NW-1, D-95447, Bayreuth, Germany



## ARTICLE INFO

## Keywords:

STEAM  
Math learning  
Inquiry-based  
Hands-on  
Art  
Informal learning  
Motivation  
Reasoning

## ABSTRACT

An informal mathematical module integrating Arts (modifying STEM to STEAM) and following an inquiry-based learning approach was applied to a sample of 392 students (aged 12–13 years). The three lesson module dealt with mathematical phenomena providing participants with the commercially available hands-on construction kit, aiming to advance STEAM education. Pupils built original, personal, and individual geometrical structures by using plastic pipes in allowing high levels of creativity as well as of autonomy. Tutors supervised the construction process and intervened only on demand. A pre-/post-test design monitored the cognitive knowledge and the variables of relative autonomy, visual reasoning, formal operations as well as creativity. Our informal intervention produced newly acquired cognitive knowledge which as a process was shown of being supported by a broad basis of (soft) factors as described above. A path analysis elaborated the role of creativity (measured with two subscale: act and flow) to cognitive learning (post-knowledge), when flow was shown to lead. Pre-knowledge scores were significantly influenced by both creativity subscales: act and flow. However, relative autonomy, visual reasoning and formal operations contributed, too. In consequence, cognitive learning within STEAM modules was shown dependent on external triggers. Conclusions for appropriate educational settings to foster STEAM environments are discussed.

## 1. Introduction

There is a lot of everyday knowledge supporting the new trend of enriching the STEM into STEAM education. One of the main approaches has been the adding of the elements of arts, skills, and creativity to the learning and teaching practices of mathematics especially in relation to other areas of knowledge and culture. However, there is only little evidence based studies to confirm these practical experiences. Thus, relevant and reliable research is urgently needed to expand the emerging STEAM movement and activities.

Learning by inquiry is assumed to follow the thinking paths of scientists. Rather than being told about science or just remembering facts, students are expected to learn how to think scientifically (Alberts, 2009; Faure et al., 1972; Gardner, 1991). Beyond the accumulation of disconnected facts, inquiry-based learning is expected to support some understanding of cause and effect, of relationships as well as of the power to predict, react and control (Illich, 1971; Others & Author, 2017a,b). This way of learning does not neglect individual knowledge, it also challenges individuals' everyday ideas about reality (Lakoff & Johnson, 1999; Tal, 2014). Learning follows the process by which scientists discover knowledge by collecting empirical evidence, by building upon critical analyses, by searching for independent confirmation and by integrating results from observations and/or experiments (e.g., Driver,

<https://doi.org/10.1016/j.tsc.2018.07.003>

Received 28 April 2018; Received in revised form 16 July 2018; Accepted 17 July 2018

Available online 18 July 2018

1871-1871/ © 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Leach, Millar, & Scott, 1996; Keselman, 2003; Kuhn, 2005; Burnard, 2015). Inquiry-learning may lead students to question and to share solutions, to extract valid conclusions from hands-on experiments, to formulate questions, to work with peers and to apply the usual research techniques (Lederman & Lederman, 2012; Scharfenberg & Bogner, 2013). Any inquiry process may integrate learning tasks, assessments, resources, environments, and teaching strategies, and typically may reflect commitment to student-centeredness and learner empowerment (White & Frederikson, 1998). Formal and informal settings may support higher-order learning experiences and lead to participation in scientific practices by using the discourse of science and working with scientific representations and tools (Rennie, 2014). Consequent hands-on experience might support observations (Oppenheimer, 1968) and lead to subsequent conclusions instead of offering occasional tours to such experiments (Salmi, 2003). Informal, open learning environments have also proved to be effective in learning mathematics (Vainikainen, Salmi, & Thuneberg, 2016). The integration of arts into STEM is supposed to reshape scientific education and humanities education, especially when further supported by problem-solving integration within trans-disciplinary frameworks (Salmi, Thuneberg, & Fenyvesi, 2017; Fenyvesi, Koskimaa, & Lavicza, 2015).

Scruggs and Mastropieri (1994) have shown that when students use hands-on methods they tend to enjoy learning more, to remember better and to consider hands-on as more effective for their learning than traditional classroom teaching methods, and especially as more efficient than learning only by seeing or listening. Teachers also rate the hands-on method as the most effective method (Ballantyne & Packer, 2009). Also, hands-on learning has been found effective for learners with learning difficulties (Brigham, Scruggs, & Mastropieri, 2011; Salmi & Thuneberg, 2017) or with more serious emotional disturbances (McCarthy, 2005). Informal, open learning environments have been proven as effective also for mathematics learning (Salmi, Vainikainen, & Thuneberg, 2015).

STEAM (Science, Technology, Engineering, Art, Mathematics education) can be defined as “education for increasing students’ interest and understanding in scientific technology and for growing STEAM literacy based on scientific technology and the ability to solve problems in the real world” (Kofac, 2017, p. 3)”. This definition has been one of the principles in successful education reforms in South-Korea. The definition is then operationalized into practical education by two key terms: 1. Education based on scientific technology, and 2. Ability to solve problems in the real world (Kofac, 2017). This same dilemma has been topic also in UK, where several attempts have been growing to broad the curricula towards a more responsive, dynamic, and also inclusive form of education. STEAM has been showing promising results in evidence-based practices (BERA-report, 2017).

The current STEAM initiatives may provide additional channels to learn abstract mathematical problem-solving tasks (BERA-report, 2017; Yakman & Lee, 2012). The interactive hands-on method in a workshop setting is assumed to encourage students to find creative solutions based on experimentation and observation, using the learning by doing principle (Dewey, 1938), which is important at the concrete operational stage (Piaget, 1977). The aesthetic elements of handicraft and art promote understanding of mathematical concepts by exposing students to concrete space and shape experiences (Dewey, 1980; Mack, 2006). Creativity as a supposedly complex construct comprises the difficulty to define and to quantify. As creativity (Burnard, 2015) is assumed to introduce new impulses into science education (STEM) and lead to better acceptance of science by adolescents, recent initiatives have proposed an integration of creativity (summarized in Arts), thus modifying STEM to STEAM. As creative processes are assumed to be complex, many studies contributed to prepare the field for quantification: Csikszentmihalyi (2000) defined two domains, the first was called “flow”, characterized by complete absorption in an activity: a person is regarded as fully immersed in a feeling of energized focus, accompanied with full involvement and enjoyment. Flow is perceived as linked to intrinsic motivation, particularly at young ages. Flow levels may tend to drop within knowledge-based classroom activities (Csikszentmihalyi, 2000). Act as the second pillar involved is covering conscious and trainable cognitive processes. Quantifying creativity of adolescents is following this two-pillar approach by following a Likert-scale questionnaire originated from Miller and Dumford (2016).

In consequence, our study had four objectives: First to assess the cognitive learning potential of participating students by monitoring pre- and post-knowledge levels; second, to identify the effect of visual reasoning, abstract thinking, experienced autonomy to predict the expected learning outcome; third, to characterize the influence of creativity in our STEAM-module; and finally to examine gender differences.

## 2. Methods and procedures

### 2.1. Participants

Participants came from the Helsinki capital area (N = 392), 52% girls (n = 204) and 48% boys (n = 188), age average 12 years and 4 months (Std.Dev. = 0.32). Altogether 11 schools contributed to our convenience sample. The study followed the empirical permission requirements and ethical principles.

### 2.2. Educational intervention

Our creative math & art workshop offered an opportunity to enrich traditional STEM into STEAM education. The pupils could build, test, explore and learn in small groups of 6–8 freely within a 3 × 45-minute time period. The workshop took place in an open learning environment of the university premises. The workshop followed the national curriculum obligations. At first, after a ten minutes introduction the concrete materials were presented providing the basic information about options of the creative hands-on construction materials. Two adults supervised as tutors mostly following, encouraging and providing information on demand. Pupils were encouraged to build their own structures by using the small plastic pipes and circles coming up with, for instance, machines, creatures, mobile equipment, or structures: They could produce, fabricate or create amusement. An overall plan was required

**Table 1**  
Item wording with the actual pre-test knowledge scores (first row).

Pre-test scores	
.10	The sum of the angles of a hexagon is 720 degrees. a) true, b) not true, c) I don't know
.11	5.1 cm distance on the map equals 1 km, when the scale is 1:20,000. a) true, b) not true, c) I don't know
.16	45 is 50% bigger than 30. a) true, b) not true, c) I don't know
.24	The surface area of a circle and square is exactly the same. Which one has the longer circumference? a) circle, b) square, c) exactly the same, d) I don't know
.26	$x = 3$ is the solution to equation $2x - 3 = 11 - 3x$ . a) true, b) not true, c) I don't know
.32	The sum of the angles of a triangle is 200 degrees. a) true, b) not true, c) I don't know
.42	A pyramid is a) cylinder, b) prism, c) cone
.57	Which fraction is bigger: $2/3$ or $4/7$ : a) $2/3$ is bigger; b) $4/7$ is bigger; c) equal; d) I don't know

consisting of simple drawings and short explanations. They were encouraged also to modify plans according to the empirical evidence they encountered during the process. In the open learning environment, there was a “wind tunnel”, a real scientific small-scale scientific test equipment. The pupils had an opportunity to test and try their hands-on construction materials and improve the structure by using the “wind tunnel”. In the end, all productions were presented to the peers and documented by photos or videos. The class teacher only offered help on demand.

The “math & art” workshop as STEAM-approach involved art, creative mathematics, engineering, and technology by, for example, building large artificial, moving animals a fusion of art and technology – with curiosity, imagination, and play (Authors et al., 2017). The commercially available hands-on construction kit is based upon the structural analysis and geometric formalization of building techniques in utilizing construction guidelines of traditional, light-structured buildings (Fenyvesi et al., 2015) as well as other aesthetic and artistic sources (Sochacka, Gyuotte, & Walther, 2016). Small groups of 3 pupils could use of one box containing hundreds of 2–30 cm long “tubes or straws” and various types of “connectors”, flexible enough to construct “unbreakable” modules or spatial formations. The material offered opportunities for conceptualizing, modeling, or analyzing structures relevant for instance to geometry and art (Fenyvesi et al., 2015). A step-by-step approach demands empirical testing, evidence, and analysis of each stage in the construction, thus offering a good platform for creative STEAM education.

### 2.3. Measures

Our ad-hoc knowledge test covered the workshop’s content areas. Pupils were presented with short verbal stimuli followed by one to four statements, which needed a judgement of correct or incorrect. Pilot analyses omitted two items from the further analyses. Thus, our final measure consisted of eight items containing understanding of mathematical, especially geometry and problem-solving phenomena. All items are listed in Table 1. Answering options ranged from 1 = true, 2 = untrue, 3 = I don't know, the subsequent coding was dichotomously. The final test scores for pre- and post-test consisting of the same items were calculated as sum scores.

The STEAM-activities in the hands-on workshop were clearly linked to the mathematical content of the curriculum (such as geometrical forms, simple equations and calculations). However, the open learning environment gave lots of freedom for the pupils how to solve the tasks.

All participants completed a pre-test (T0) two weeks before the teaching-unit, a post-test (T1) ten days after the intervention. All questionnaires were paper-and-pencil-tests completed under controlled conditions. Students were never aware of the retention testing schedules. Besides the knowledge test (to measure cognitive learning success) as described above, we applied (ii) a creativity measure consisting of 10 items originating from Miller and Dumford (2016) as well as (iii) a science motivation measure covering the subscales: IM “intrinsic motivation” and SD “self-determination”, both consisting of 4 items per subscale using a 5-point Likert scale pattern ranging from “never” (1) to “always” (5).

The Autonomous Motivation was based on the self-determination measure of Deci and Ryan (2002), which connects satisfaction for basic psychological need, autonomy, competence and relatedness, to motivation: the more the needs are fulfilled, the more autonomous is motivation and accordingly self-regulation. The scale consists of 32 items with four Likert-options (Ryan & Connell, 1989). The scale contains self-regulation styles, for example, reasons are monitored for why homework is done or why one tries to answer hard questions in class. The self-determination continuum included styles from non-autonomous towards autonomous: external, introjected, identified and intrinsic subscales. Ryan and Connell (1989) had presented the Relative Autonomy Index (RAI) by following the formula:  $2 \cdot \text{intrinsic} + \text{identified-introjected} - 2 \cdot \text{external}$ . The RAI therefore is supposed to describe the overall relative autonomy level of a participant. The plus-sign in RAI indicates that the experience is rather autonomous, minus-sign that one relies more on others than one trusts in one-self.

The cognitive measure was measured using the visual reasoning Raven Standard Progressive Matrices test (Raven, Raven, & Court, 2003) that quantifies visual non-verbal cognitive skills. Elements measured the capacity to learn and to embrace and remember the knowledge once learned (Raven et al., 2003). In each test item, the subject is asked to identify the missing element that completes a pattern. The test contains 60 items divided into five sets (A, B, C, D, E), each including 12 different tasks.

The Formal operational abstract thinking measure *Formula 1* is based on Inhelder and Piaget (1958), the core of which is the control of variables strategy (Shayer, 1979). This strategy has been shown (Kuhn, 2002; Neimark, 1975; Shayer, 2008; Sodian & Bullock, 2008) to be central to science and an attainable skill for children advancing to a formal operational level in their thinking.

The test measures the mastery of control variables is a second-generation modified group-version (Hautamäki, 1989, 2000; Author et al., 2014) of the original scientific reasoning task, The Pendulum. (Shayer & Wylam, 1978; Shayer, Kuchemann, & Wylam, 1976; Shayer, 1979; Wylam & Shayer, 1980), based on one of Inhelder and Piaget’s identified formal schemata (Inhelder & Piaget, 1958).

Formula 1 has been widely used in Finnish studies (Thuneberg, Hautamäki, & Hotulainen, 2014). In the test the subjects compare F1 drivers, cars, tires, and race tracks (four variables, each with two given values from which to select: Räikkönen, Schumacher; Ferrari, McLaren; Michelin, Bridgestone; Monaco, Silverstone). Subjects are given a set of values for the four variables (such as Räikkönen, Ferrari, Michelin, Monaco) and asked to construct another set that would clarify the role of a specified variable (say, cars). Subjects should produce a set of values for all four variables so that the focal variable is studied with an unconfounded pair (see Strand-Cary & Klahr, 2008). Subjects choose their selection from a given set of values for all variables, or answer by checking whether the pair is a good test of driver, car, tires, race.

In the present study the children are in the transfer phase from concrete operations to abstract thinking, formal operations. They are approaching the age (12–15) when formal operations are emerging and reasoning is possible independently of the presence of concrete objects (2006, Piaget, 1970).

2.4. Data analysis methods

The following statistical analysis methods were applied: Principal component analysis using oblimin rotation was applied to the CREAT variables. 2) The mean differences between boys and girls were tested by one-way analysis of variance (ANOVA) and the change between the pre- and post-knowledge test by General linear modelling (GLM) repeated measures, with the effect-size measured as the partial eta-squared coefficient ( $\eta^2 > .01$  small,  $> .06$  middle,  $> .14$  large), considered acceptable as the recommended generalized coefficient when only one grouping factor is used as in this case (see Bakeman, 2005). 3) The structural equation modeling (AMOS 22) SEM was based on theoretically valid variables and their bivariate correlations. The RAI, gender, RAVEN, formal operations test, CREAT\_act and CREAT\_flow were used as covariates to control their effects on measured pre and post knowledge test variables. The goodness of fit of the models was based on a chi-square-test ( $p > .05$ ) and indices of NFI, TLI, and CFI (good fit  $> .90$ , or better  $> .95$ ), RMSEA reasonable fit  $< .08$ , good fit  $< .05$  (see Byrne, 2010). Standardized beta coefficients and  $R^2$ -values reflect the total portion of variance explained by the predictor variables.

There were on average 6.8% missing values. The Maximum likelihood method and estimation of means and intercepts were used in the SEM path analysis due to the missing values.

3. Results

The reliabilities of the measures were good or satisfactory: The knowledge tests (8 items), Pre-test Cronbach’s scored  $\alpha = .613$ , post-test,  $\alpha = .65$ ; The Autonomous motivation,  $\alpha = .917$  (32 items); The Raven,  $\alpha = .882$  (60 items); Formula 1,  $\alpha = .817$  (26 items); The CREAT act,  $\alpha = .744$  (5 items). The reliability of the CREAT flow test was only just sufficient reliable, with Cronbach’s  $\alpha = .592$  (3 items).

The principal component analysis on the CREAT variables resulted to two components, which explained 50% of the variance. The first component was labelled CREAT\_act and it consisted of five variables (component loadings .428–.777). The first component CREAT\_act consisted of: 1. *Joined together dissimilar concepts to create a novel idea* (.778). 2. *Incorporated a previously used solution in a new way* (.738). 3. *Made a connection between a current problem or task and a related situation* (.670). 4. *Imagined a potential solution in a new way* (.626), and 5. *Tried to generate as many ideas as possible when approaching a task* (.428). The second component CREAT\_flow consisted of: 1. *Lost track of time when working intensely* (.746) 2. *Felt that work was automatic and effortless during an enjoyable task* (.731), and 3. *Been fully immersed in your work on a problem or a task* (.688). (Two items were omitted because of similar and high cross-loadings). The factor scores of these components were used in the further analyses.

The following Table 2 presents the descriptive statistics, both totally and separated for girls and boys.

The only significant difference ( $p = .001$ ) between boys and girls was that boys scored higher on the knowledge test measured before the Math and Art workshop.

The overall and within group change is explained in regard to the knowledge pre-test and post-test situation. The change between pre- and post knowledge test was significant (*Wilks’ lambda* = 6.076, *df* = 1,  $p = .014$ , partial eta-squared = .016). There was no time by gender interaction effect ( $p = .31$ , so the change was statistically identical for girls and boys. However, there was a significant

Table 2  
Statistical descriptives (M = Mean, SD = Standard Deviation).

	N girls	M	SD	N boys	M	SD	Total M	SD	Min	Max
Knowledge pre	203	.251	.174	187	.316	.202	.282	.191	0	1
Knowledge post	201	.289	.195	181	.329	.240	.308	.218	0	1
Raven	199	34.357	6.680	181	33.193	7.517	33.803	7.106	2	48
Formal op.	193	12.238	4.019	177	12.576	4.253	12.400	4.130	1	20
RAI	197	.211	2.066	170	.160	1.849	.187	1.966	-7.229	7.210
CREAT_act pre	181	.005	1.001	166	-.006	1.001	.000	1.000	-2.469	3.088
CREAT_flow pre	181	-.068	1.022	166	.074	.973	.000	1.000	-2.626	2.610

**Table 3**  
Bivariate correlations of the variables.

	Knowl. pre	Knowl. post	Raven	Formal operat.	RAI	Creative act
Knowledge post	.418***					
Raven	.226***	.174**				
Formal operations	.295***	.214***	.233***			
RAI	.218***	.166**	.042	.078		
Creative act	.248***	.155**	-.004	-.021	.137*	
Creative flow	.257***	.270***	-.029	.065	.306***	.415***

Note: \* =  $p < .05$ , \*\* =  $p < .01$ , \*\*\* =  $p < .001$ .

between –subjects effect, so the average level at the two time-points of the boys and girls differed,  $F(1,379) = 1122.837, p = .014$ , partial eta-squared = .748.

The Pearson correlation coefficients are presented in Table 3. The pre and post knowledge test variables correlated with all other variables, the pre-test more than the post-test, however. The cognitive variables Raven and formal operations correlated, in addition to the knowledge variables, only with each other. RAI correlated with the creativity variables, especially with CREAT\_flow. The correlations were small to medium. On the basis of this correlation analysis, the variables theoretically chosen for this study, were considered valid to use in the SEM path-model.

### 3.1. SEM path analysis

The Relative autonomy Index (RAI), gender, cognitive visual reasoning (Raven), formal operations, CREAT\_act and CREAT\_flow were used as covariates to control their effects on the measured variables Knowledge T1 and T2. The final model containing only significant effects fitted the data well:  $\chi^2 = 16.762, df = 15, p = .333; NFI = .949, TLI = .985, CFI = .994; RMSEA = .017$ .

The final path model is presented in Fig. 1. The magnitude of the paths (the standardized beta-coefficients) are shown with their levels of significance (\* =  $p < .05$ , \*\* =  $p < .01$ , \*\*\*  $p < .001$ ) and the total variance explained by  $R^2$ .

The variables totally explained 24% on pre Knowledge and 21% on post Knowledge. In the following, the main effects (after the role of other variables has been removed by analysis of covariance) are presented by standardized betas:

- (1) Visual reasoning directly predicted pre-Knowledge (beta = .18,  $p < .001$ ).
- (2) Autonomous motivation directly predicted pre-Knowledge (beta = .14,  $p < .01$ ).
- (3) Formal operations predicted both pre-Knowledge (beta = .24,  $p < .001$ ) and post-Knowledge (beta = .10,  $p < .05$ ).
- (4) CREAT act predicted pre-Knowledge (beta = .19,  $p < .001$ ).
- (5) CREAT flow predicted pre-Knowledge (beta = .12,  $p < .05$ ) and post-Knowledge (beta = .17,  $p < .001$ ).
- (6) Gender predicted only pre-Knowledge.

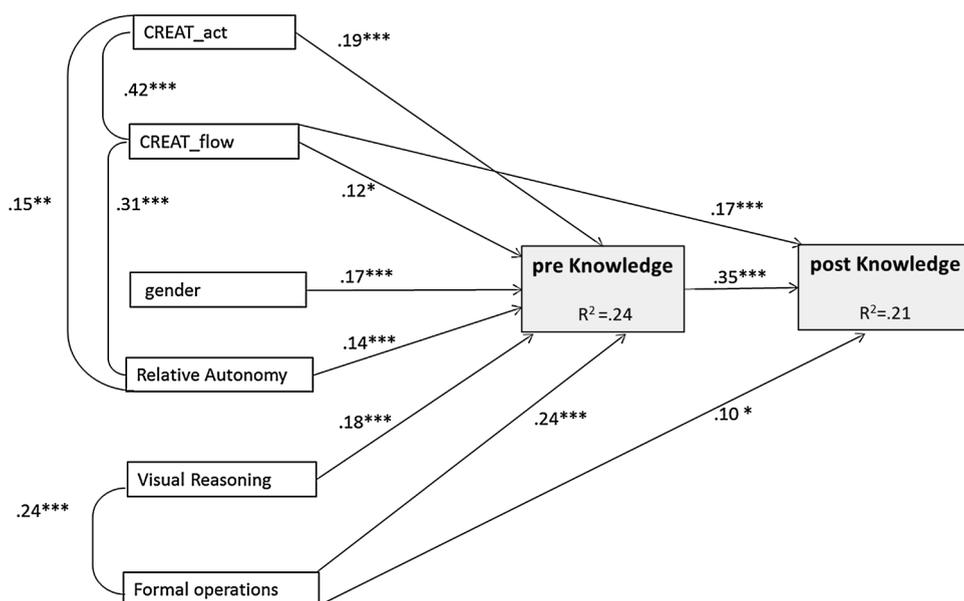


Fig. 1. The final path model.

(7) *Pre Knowledge* predicted *post-Knowledge* ( $\beta = .35, p < .001$ ).

When the moderation effect of gender was analyzed, the same model did not fit for both groups:  $\chi^2$  difference = 21.677,  $df = 8, p = .006$ . Hence the analysis continued by comparing the effects path by path. The result was that the only path, which was significantly different between boys and girls, was that from Formal operations to Knowledge pre-test ( $z = 4.122, p < .019$ ; girls'  $\beta = .051, p = .442$ , boys'  $\beta = .416, p = .000$ ).

#### 4. Discussion

Informal short-term hands-on modules produce cognitive learning with a substantial contribution by creativity. Effects are small when taking the measurements between the two testing points into account. A short-term, three hours intervention may not raise more potential. Such small but significant shifts are quite in line with earlier studies (Rennie, 2014; Thuneberg & Salmi, 2018). As students had increased their individual knowledge scores sustainably, our hands-on approach is to be successful, quite in line with a constructivist perspective, where individual hands-on activities appeared supportive to meaningful learning by requiring individual thinking (Mayer, 2004). Another cause for substantial knowledge acquisition may lie in the performance of small cooperative groups which have been found to invest more effort leading to higher achievements (Johnson & Johnson, 1999). Similarly, Lord (2001) had presented cooperative learner-centered modules as a suitable approach to support individual cognitive achievement.

As boys and girls apparently had learned equally based on the GLM-analysis, possible gender differences were of interests: The level of boys' knowledge profile was higher but the girls' mean knowledge caught up in the post-test. This gap has often been described in earlier studies: Meissner and Bogner, for instance, had reported a substantial knowledge increase, both in short-term and long-term perspective, but also a substantial gender-dependent difference in cognitive achievement: Active engagement in hands-on lab activities apparently were shown beneficial for girls, and thus to bridge gender differences. The SEM-model within the controlled situation by covariates revealed effects of the cognitive variables, relative autonomy and creative variables; highest was the effect of formal operations. In the post-test situation, CREAT\_flow and formal operations still had an influence on knowledge, and CREAT\_flow an even higher one than on the pre-knowledge. The analysis of the paths showed that there was only one path, which significantly affected to this result of the need of two models, namely the effect of formal operations on the pre knowledge test. In other words, the higher the level of abstract thinking among boys, the higher were the Math knowledge test scores before the Math and Art workshop. This path was non-significant in girls but highly significant and of medium size in boys.

It is interesting that Raven as a visual test had no connection with the creative variables. Instead, it seems natural that relative autonomy had an effect, because relative autonomy experience has been shown to support intrinsic motivation, enjoyment of task and creativity (Eccles & Wigfield, 2002; Reeve, 2002; Ryan & Deci, 2000; Wiest, Wong, & Kreil, 1998). The Raven test has been widely utilized both in practice and theoretical research. The test has turned out to be a reliable and valid tool in different age groups, schools and also in many socio-economic cultures. One of its benefits is that it does not give gender advantage for either male or female pupils. It quantifies the common cognitive ability and the capacity to learn. In addition, it forecasts the capacity to remember knowledge once learned. The most important reason for the use of Raven in our study was the fact that the test measures of non-verbal cognitive skills. STEM education has traditionally been tactile, but has become more and more visual during the digital era. However, many of the skills trained and taught are definitively not textual. Therefore, a mismatch between the structure of the knowledge and the structure of the media traditionally used may exist (Greenfield, 2009). The concrete hands-on workshop also offered for STEAM-learning an authentic environment, which has been found to be an effective motivating factor (Alberts, 2009; Braund & Reiss, 2007).

There was at least one source of bias in the analyses, the reliability of CREAT flow was just in the border of commonly accepted Cronbachs' alpha criterion. The reason of the rather low reliability might be in the small number (three) of items it consisted of.

#### 5. Conclusions

Pupils constructed original and individual geometrical structures by using plastic pipes and circles by allowing creativity at a high level of autonomy. As all constructions were created in team work providing collaborative learning and satisfaction. Different channels allowed sharing creative solutions with peers, for instance, by mobile phones with photos and videos. Earlier literature (Rennie, 2014) had reported this type of engagement as supportive for pupils to question and to share knowledge. Informal, open-learning environments seem to provide opportunities also for different types of learners (Burnard, Craft, & Grainger, 2006) and to offer opportunities to utilize a pupil's own learning strategies especially related to hands-on experiments (Thuneberg & Salmi, 2018). Additionally, curiosity, imagination, and play may encourage to formulate authentic questions, to work with peers and to apply the usual research techniques (Görlitz, 1987). The learning process in our module turned out to be interestingly complex in relation to the role of formal thinking, with an interaction with gender. In the pre-test, there was both a gender effect (boys > girls) and a differential effect in relation to formal thinking. Among boys, formal thinking made a difference. This was not true for girls: the formal level did not contribute to results in the pretest.

In the post-test, the general intervention effect for knowledge gains scored positive: the mean level was higher, but there was no gender gap. Here, the formal thinking appeared to have a moderating effect for girls, showing added value in this learning context. In the post-test, the outcome was positively related to formal level of thinking, not only for boys, but now also for girls. If the level of formal thinking was higher, the post-test outcome was also higher. This sleeper effect for girls was so effective that the gender gap diminished. For some reasons, something taking place in the intervention activated girls to use also their other resources, in this study

formal thinking. Tentatively, the learning situation provided gibsonian affordances (see Gibson, 1979), incentives to make more use of available modes of thinking.

Interdisciplinarity and aesthetics are essential elements of creative STEAM education especially related to mathematics (see also Fenyvesi & Lähdesmäki, 2017). Bridging the gap between mathematical problem-solving and hands-on learning was especially successful among the pupils with a higher level of abstract thinking abilities – an essential element of learning-to-learn capacities (see also Hautamäki & Kupiainen, 2014). The elements of autonomy, creativity, personal inquiries and original testing turned out to provide an opportunity to activate and utilize these capacities in learning mathematics in an informal, out-of-school setting. The expansion of STEM to STEAM seems worthwhile.

## Acknowledgements

This work was supported by the European HORIZON-2020 framework labelled CREATIONS: Developing an Engaging Science Classroom (Grant Agreement No.665917; <http://creations-project.eu>). Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the position of the founding institutions. Last but not least, we would like to thank all students and teachers who supported our study.

## References

- Alberts, B. (2009). Making a science of education. *Science*, 323(5910), 15. <https://doi.org/10.1126/science.1169941>.
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, 37(3), 379–384.
- Ballantyne, R., & Packer, J. (2009). Introducing a fifth pedagogy: Experience-based strategies for facilitating learning in natural environments. *Environmental Education Research*, 2(15), 243–262.
- BERA-report (2017). In L. Colucci-Gray, P. Burnard, C. Cooke, R. Davies, & J. Trowsdale (Eds.). *Reviewing the potential and challenges of developing STEAM education through creative pedagogies for 21st learning: How can school curricula be broadened towards a more responsive, dynamic, and inclusive form of education?* London: The British Educational Research Association. Retrieved latest 12.4.2018 <https://jotrowsdale.files.wordpress.com/2017/11/bera-research-commission-report-steam.pdf>.
- Braund, M., & Reiss, M. (2007). Towards a more Authentic Science Curriculum: The contribution of out-of-school learning. *International Journal of Science Education*, 28(12), 1373–1388.
- Brigham, F., Scruggs, T., & Mastropieri, M. (2011). Science education and students with learning disabilities. *Learning Disabilities Research*, 26(4), 211–230.
- Burnard, P. (2015). Re-positioning creativities in relation to effective arts pedagogy: UK perspectives on teaching for creativity and teaching creatively in the arts. In J. Fleming, R. Gibson, & M. Anderson (Eds.). *How arts education makes a difference* (pp. 249–264). London: Routledge.
- Burnard, P., Craft, A., & Grainger, T. (2006). Pedagogy and possibility thinking in the Early Years. *Thinking Skills and Creativity*, 1(2), 26–38.
- Byrne, B. (2010). *Structural equation modeling with AMOS. Basic concepts, applications, and programming* (second edition). New York: Routledge.
- Csikszentmihalyi, M. (2000). *Beyond boredom and anxiety* (25. Anniversary ed.). San Francisco: Jossey-Bass Publishers.
- Overview of self-determination theory: An organismic dialectical perspective. In E. Deci, & R. Ryan (Eds.). *Handbook of self-determination*. Rochester, NY: The University of Rochester Press.
- Dewey, J. (1938). *Experience & education. The Kappa Delta Pi Lecture Series*. New York: Collier Books.
- Dewey, J. (1980). *Art as experience*. Penguin, USA: The Berkley Publishing Group.
- Driver, D., Leach, R., Millar, P., & Scott, P. (1996). *Young peoples images of science*. Buckingham, England: Open University Press.
- Eccles, J., & Wigfield, A. (2002). Self-regulation beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132.
- Faure, E., Herrera, F., Kaddoura, A.-R., Lopes, H., Petrovsky, A. V., Rahnema, M., et al. (1972). *Learning to be: The world of education today and tomorrow*. Paris: UNESCO.
- Fenyvesi, K., & Lähdesmäki, T. (2017). *Aesthetics of interdisciplinary art and mathematics*. Cham, Switzerland: Birkhäuser 290 p.
- Fenyvesi, K., Koskimaa, R., & Lavicza (2015). Experiential education of mathematics: Art and games for digital natives. *Kasvatus ja aika*, 9(1), 107–134.
- Gardner, H. (1991). *The unschooled mind. How children think and how schools should teach*. New York: Basic Books.
- Gibson, J. (1979). *The ecological approach to visual perception*. Hillsdale, NJ: Lawrence Erlbaum Association.
- Görlitz, D. (1987). Exploration and attribution in developmental context. In J. Wohlwill (Ed.). *Curiosity, imagination and play: On the development of spontaneous cognitive and motivational processes, görlitz*. NJ, USA: Lawrence Erlbaum.
- Greenfield, P. (2009). Technology and Informal Education: What is Taught, What is Learned. *Science*, 323, 69–71. <https://doi.org/10.1126/science.1167190>.
- Hautamäki, J. (1989). The application of a rasch model on piagetian measures of stages of thinking. In P. Adey, J. Bliss, J. Head, & M. Shayer (Eds.). *Adolescent development and school science* (pp. 342–349). London: Falmer Press.
- Hautamäki, J., & Kupiainen, S. (2014). Learning to learn in Finland. Theory and policy, research and practice. In Ruth Deakin Crick, Cristina Stringher, & Kai Ren (Eds.). *learning to learn. International perspectives from theory and practice*. London: Routledge.
- Illich, I. (1971). *Deschooling society*. New York: Harper and Row.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Johnson, D., & Johnson, R. (1999). Making cooperative learning work. *Theory into practice*, 38(2).
- Keselman, A. (2003). Supporting inquiry learning by promoting normative understanding of multivariable causality. *Journal of Research in Science Teaching*, 40(9), 898–921. <https://doi.org/10.1002/tea.10115>.
- Kofac (2017). *Concept and definition of STEAM*. Seoul: The Korea Foundation for the Advancement of Science and Creativity - KOFAC. Retrieved 12 April 2018 [https://steam.kofac.re.kr/?page\\_id=11269](https://steam.kofac.re.kr/?page_id=11269).
- Kuhn, D. (2002). What is scientific thinking and how does it develop? In U. Goswami (Ed.). *Handbook of childhood cognitive development* (pp. 371–393). Oxford: Blackwell.
- Kuhn, D. (2005). *Education for thinking*. 218 p. Cambridge, MA: Harvard University Press.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind & its challenge to western thought* (new edition). Basic Books.
- Lederman, N. G., & Lederman, J. S. (2012). Nature of scientific knowledge and scientific inquiry: Building instructional capacity through professional development. In B. Fraser, K. Tobin, & C. McRobbie (Vol. Eds.), *Second International handbook of science education. Springer International handbooks of education: vol. 24* Dordrecht: Springer. [https://doi.org/10.1007/978-1-4020-9041-7\\_24](https://doi.org/10.1007/978-1-4020-9041-7_24).
- Lord, T. R. (2001). 101 reasons for using cooperative learning in biology teaching. *The American Biology Teacher*, 63(1), 30–38.
- Mack, A. (2006). A deweyan perspective on aesthetic in mathematics education. *Philosophy of Mathematics Education Journal*, 19(1), 27–42.
- Mayer, R. E. (2004). Should there be a three-shrikes rule against pure discovery learning? *The American Psychologist*, 59(1), 14–19.
- McCarthy, C. (2005). Effects of thematic-based, hands-on science teaching versus a textbook approach for students with disabilities. *Journal of Research in Science Teaching*, 42(3), 245–263.
- Meissner, B., & Bogner, F. X. (2012). Science Teaching based on Cognitive Load Theory: Engaged Students, but Cognitive Deficiencies. *Studies in Educational Evaluation*, 38, 127–134.

- Miller, A. L., & Dumford, A. D. (2016). Creative cognitive processes in higher education. *The Journal of Creative Behavior*, 50(4), 282–293. <https://doi.org/10.1002/job.77>.
- Neimark, E. (1975). Intellectual development during adolescence. In F. Horowitz (Ed.), *Review of child development research* (pp. 541–594). Chicago, IL: University of Chicago Press.
- Oppenheimer, D. (1968). A rationale for a science museum. *Curator the Museum Journal*, 11(3), 206–209.
- Piaget, J. (1970). *Structuralism*. New York: Harper & Row.
- Piaget, J. (1977). *Epistemology and psychology of functions*. 228 p. Dordrecht, Netherlands: D. Reidel Publishing Company.
- Piaget, J. (2006). Reason. *New Ideas in Psychology*, 24(1), 1–29.
- Raven, J., Raven, J. C., & Court, J. (2003). *Manual for Raven's progressive matrices and vocabulary scales*. Oxford, UK: OPP Limited.
- Reeve, J. (2002). Self-determination theory applied to educational settings. In E. Deci, & M. Ryan (Eds.), *Handbook of self-determination*. Rochester, NY: The University of Rochester Press.
- Rennie, L. (2014). Learning science outside of school. In N. Lederman, & S. Abell (Vol. Eds.), *Handbook of research on science education: Vol. II*, (pp. 120–144). London and New York: Routledge <http://hdl.handle.net/20.500.11937/35673>.
- Ryan, R., & Connell, J. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57(5), 749–761.
- Ryan, R., & Deci, E. (2000). Self-Determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, 55(1), 68–78.
- Salmi, H. (2003). Science Centres as learning laboratories. *International Journal of Technology Management*, 25(5), 460–476. <https://doi.org/10.1504/IJTM.2003.003113>.
- Salmi, H., & Thuneberg, H. (2017). Learning with the dinosaurs: a study on motivation, cognitive reasoning, and making observations. *International Journal of Science Education*, 7(3), 203–217. <https://doi.org/10.1080/21548455.2016.1200155> Part B.
- Salmi, H., Vainikainen, M., & Thuneberg, H. (2015). Mathematical thinking skills, self-concept and learning outcomes of 12-year-olds visiting a mathematics science centre exhibition in Latvia and Sweden. *Journal of Science Communication*, 14(4), 1–19.
- Salmi, H., Thuneberg, H., & Fenyvesi, K. (2017). Hands-On Math and Art Exhibition Promoting Science Attitudes and Educational Plans. *Education Research International*, 2017, 13 p.. <https://doi.org/10.1155/2017/9132791> ID 9132791.
- Scharfenberg, F.-J., & Bogner, F. X. (2013). Teaching gene technology in an outreach lab: Students' assigned cognitive load clusters and the clusters' relationships to learner characteristics, laboratory variables, and cognitive achievement. *Research in Science Education*, 43(1), 141–161.
- Scruggs, T., & Mastropieri, M. (1994). Current approaches to science education: Implications for mainstream instruction of students with disabilities. *Remedial and Special Education*, 14(1), 15–24.
- Shayer, M. (1979). Has Piaget's construct of formal operational thinking any utility? *The British Journal of Educational Psychology*, 49(3), 256–276.
- Shayer, M. (2008). Intelligence for education: As described by Piaget and measured by psychometrics. *The British Journal of Educational Psychology*, 78(1), 1–29.
- Shayer, M., & Wylam, H. (1978). The distribution of Piagetian stages of thinking in British middle and secondary school children 11–14/16-year olds and sex differentials. *The British Journal of Educational Psychology*, 48(1), 62–70. <https://doi.org/10.1111/j.2044-8279.1978.tb02370.x>.
- Shayer, M., Kuchemann, D. E., & Wylam, H. (1976). The distribution of Piagetian stages of thinking in British middle and secondary school children. *The British Journal of Educational Psychology*, 46, 164–173.
- Sochacka, N., Gyuotte, K., & Walther, J. (2016). Learning together: A collaborative auto-ethnographic exploration of STEAM (STEM + the Arts) education. *Journal of Engineering Education*, 105(1), 15–42.
- Sodian, B., & Bullock, M. (2008). Scientific reasoning: Where are we now? *Cognitive Development*, 23(1), 431–434.
- Strand-Cary, M., & Klahr, D. (2008). Developing elementary science skills: Instructional effectiveness and path independence. *Cognitive Development*, 23, 488–511.
- Tal, T. (2014). Out-of-school: Learning experiences, teaching and student's learning. In N. Lederman, & S. Abell (Vol. Eds.), *Handbook of research on science education: vol. II*, (pp. 1109–1122). London and New York: Routledge. [https://doi.org/10.1007/978-1-4020-9041-7\\_73](https://doi.org/10.1007/978-1-4020-9041-7_73).
- Thuneberg, H., Hautamäki, H., & 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), 337–356.
- Thuneberg, H., & Salmi, H. (2018). To know or not to know: uncertainty is the answer. *Journal of Science Communication*, 17(2), 1–28. <https://doi.org/10.22323/2.17020201>.
- Vainikainen, M., Salmi, H., & Thuneberg, H. (2016). Situational Interest an Learning in a science center mathematics exhibition. *Journal of Research in STEM Education*, 1(1), 15–29.
- White, B., & Frederikson, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, 16(1), 3–118.
- Wiest, D., Wong, E., & Kreil, D. (1998). Predictors of global self-worth and academic performance among regular education, learning disabled, and continuation high school students. *Adolescence*, 33(131), 601–618.
- Wylam, H., & Shayer, M. (1980). *CSMS science reasoning tasks: General guide*. Datchet, UK: NEFR.
- Yakman, G., & Lee, H. (2012). Exploring the exemplary STEAM education in the U.S. Practical educational framework for Korea. *Journal of the Korean Association for Science Education*, 32(6), 1072–1086.