Thinking Skills and Creativity 37 (2020) 100687

Contents lists available at ScienceDirect

Thinking Skills and Creativity

journal homepage: www.elsevier.com/locate/tsc

Computer-based training in math improves inductive reasoning of 9- to 11-year-old children

Mojahed Mousa[a,*], Gyöngyvér Molnár[b]

[a] Doctoral School of Education, University of Szeged, H–6722, Petőfi sgt. 30–34., Szeged, Hungary
[b] Department of Learning and Instruction, MTA–SZTE Research Group on the Development of Competencies, Center for Research on Learning and Instruction, University of Szeged, H–6722, Szeged, Petőfi sgt. 30–34., Hungary

ARTICLE INFO

Keywords:
Computer-based training
Inductive reasoning
Arabic context
Latent curve modelling
Assessment

ABSTRACT

This study focuses on a computer-based training program in inductive reasoning through tasks embedded into mathematical content for 9–11-years-old students (N = 118) and presents the results of the evaluation study. The online training consists of 120 playful problems based on Klauer's "Cognitive training for children" concept and on his theory of inductive reasoning (Klauer, 1989). Both the experimental and the control group in the study consisted of 118 participants. A computer-based inductive reasoning test comprising of 44 multiple-choice items was used in the pre- and posttest to measure the effectiveness of the training (Cronbach α = .91). Both the test and the training tasks were in Arabic context regarding language and used directions and were delivered via the eDia online assessment platform (Csapó & Molnár, 2019). On the posttest, after the six weeks of training the experimental group significantly outperformed the control group by more than one standard deviation. The effect size of the training was in international context high (Cohen d = 1.71). Non-significant variance of the latent slope indicated that there was no significant variability in responding to the intervention program. This study provided evidence that inductive reasoning could be developed even on class level and in a computerised environment very effectively at the age of 9–11 independent of students' original level of inductive reasoning, school-achievement, gender, and socio-economic status.

1. Introduction

Inductive reasoning (IR) considered as one of the most important component skill of almost all transversal skills (Molnár, Greiff, & Csapó, 2013; Söderqvist, Nutley, Ottersen, Grill, & Klingberg, 2012) such as problem solving (Wu & Molnár, 2018a, 2018b) and general intelligence (Klauer & Phye, 2008; Klauer, Willmes, & Phye, 2002; Sternberg and Gardner, 1983), it constitutes a driving force of cognitive development (Perret, 2015). It plays a central role in learning processes, such as knowledge acquisition and application (Bisanz, Bisanz, & Korpan, 1994; Hamers, De Koning, & Sijtsma, 2000; Klauer, 1990, 1996; Molnár et al., 2013; Pellegrino & Glaser, 1982) and is a good indicator of learning potential (Csapó & Nikolov, 2009; Resing, 1993). IR plays a part in a wide range of daily cognitive activities, including categorical generalization, analogical reasoning, causal reasoning or probabilistic judgement (Perret, 2015). "Inductive reasoning enables one to detect regularities, rules, or generalizations and, conversely, to detect irregularities. This is one way in which we structure our world" (Klauer et al., 2002, p. 1). This suggests that improving reasoning skills (Hattie & Anderman, 2013), especially, fostering IR should be part of formalized education, of the school activities explicitly (de

* Corresponding author.
E-mail addresses: mojahed.mousa@edu.u-szeged.hu (M. Mousa), gymolnar@edpsy.u-szeged.hu (G. Molnár).

https://doi.org/10.1016/j.tsc.2020.100687
Received 4 January 2020; Received in revised form 23 May 2020; Accepted 22 July 2020
Available online 05 August 2020

Konig, 2000; Molnár, 2011; Resnick, 1987).

At the beginning of this millennium, almost the same ideas emerged, but in a new wave of teaching 21st-century skills (Molnár & Csapó, 2019; Voogt, Erstad, Dede, & Mishra, 2013; Wegerif, 2006). Fostering thinking skills or more specifically, a set of cognitive skills essential for learning and creating new knowledge became interesting concepts for educational policy (Wegerif, Li, & Kaufman, 2017; Vainikainen, Hautamaki, Hotulainen, & Kupiainen, 2015).

It is commonly assumed that the development of reasoning skills are embedded in the ordinary school material (de Konig, 2000), focusing mainly on reading writing and mathematics (Molnár, 2011). However, according to Molnár and Csapó (2019) there was no appreciable development noticeable regarding students reasoning skills in reading and science between Grades 2 and 6 and "there was a steady increase detectable in mathematics, especially in the first four years of schooling [...] Overall, these results highlighted the importance, sensitivity and potential of the development of thinking skills in the early years of schooling". Thus enhancing thinking skills must become a real goal in education (Vainikainen et al., 2015).

One of the major challenges in classroom teaching comes from the large differences between students in terms of their abilities. Training programs are needed which can be applied even on classroom level and which can handle the individual differences, can fit the actual needs of students' cognitive level (Pásztor, 2016). Face-to-face programs are not very successful (Molnár & Lorincz, 2012) in this regard, but technology-based training programs can provide feasible solutions to address these challenges (Csapó, Lőrincz, & Molnár, 2012). Beyond adaptive fostering technology has the power of a higher level motivation by pupils of the 21st century (Csapó et al., 2012; McGaw et al., 2012)

In the present paper we address a technology-based training program of inductive reasoning for 9–11-year-old students and presents the direct result of the evaluation study. The online training consists of 120 playful problems based on Klauer's "Cognitive training for children" concept and on his theory of inductive reasoning (Klauer, 1989). All of the problems are embedded in Mathematics content, thus the training tasks are applicable during normal school hours as part of the Mathematics lesson. According to our knowledge, it does not exist any online available training program in Arabic, which is empirically proved and focuses on the development of students' inductive reasoning skills in educational context.

2. Inductive reasoning and its fostering

2.1. Definition and development of inductive reasoning

Despite of its importance there is no universally accepted definition of inductive reasoning (see e.g., Klauer, 1990; Osherson et al., 1990; Sloman, 1993; Gick & Holyoak, 1983; Sandberg & McCullough, 2010; Goswami & Brown, 1990; (Ropo, 1987; Hayes et al., 2010). According to the classical interpretation, IR covers the processes of moving from the specific to the general (direction), that it, generalizing, deriving broad rules based on single experiences and observations (Sandberg & McCullough, 2010); and IR is a form of reasoning under uncertainty (level of confidence), because it involves forming hypotheses about rules (Perret, 2015). The latter one is the reason, why it is frequently defined as an opposite to deductive reasoning, which mechanism are moving from the general to the particular (direction) and are logical (level of confidence). Probably, the most elaborated definition was published by Klauer (1993), who interpreted IR as the discovery of regularities through the detection of similarities, dissimilarities, or a combination of both, with respect to attributes or relations to or between objects. This results in total six operations of IR: generalization, discrimination, cross-classification, recognizing relations, discriminating between relations, and system formation.

The development of IR appears to emerge at a fairly young age (Perret, 2015), but IR becomes noticeably more efficient with age. According to Csapó (1997), Molnár and Csapó (2011) and Molnár et al. (2013) IR develops during a broad age range, covering the whole period of elementary and secondary education. The average pace of development is relatively slow at about one quarter of a standard deviation per year (Molnár et al., 2013) as a result of the lack of direct and explicit stimulation of IR in schools. The development occurs spontaneously as a 'by-product' of teaching rather than being guided by explicit instruction (de Konig, 2000; Molnár & Csapó, 2019).

2.2. Directions of fostering reasoning skills

Against the early views, which rejected the trainability of reasoning skills and enlarged the role of inheritance (Jensen, 1973), empirical studies indicated that these skills develop over time, mostly during compulsory schooling (Molnár et al., 2013). The pace of development is relatively slow, at about one quarter of a standard deviation per year (Molnár et al., 2013), but they are trainable. The modifiability offers opportunities and new perspectives for enhancement by educational interventions (Adey et al., 2007).

Two main groups, two main directions can be distinguished regarding the development of reasoning skills. Researchers identify with the first issue believe that reasoning skills must be and can only be taught explicitly (see e.g. Feuerstein et al., 1980; Klauer, 1989, 1991, 1993; Lipman, 1985), while researchers belong to the second group believe that operations should be embedded in traditional school subjects (e.g. the CASE project or see Dienes, 1963, 1973; Shayer & Adey, 1981). In the present study, we mix these issues by applying an explicit training of inductive reasoning strategies in Mathematics content.

2.3. The effectiveness of the IR training programs based on Klauer's model

The processes and mechanism of IR can be described with a well-structured system and model from educational perspective (Klauer, 1989, 1990, 1996; Klauer & Phye, 1994; Klauer et al., 2002). Student's level of IR can be modified very effectively, that is IR can be trained using tasks, developed in this approach (Klauer, 1996; Tomic, 1995; Tomic & Klauer, 1996). Barkl et al. (2012), de Koning & Hamers (1999), de Koning et al. (2002), Hamers et al. (1998), Klauer and Phye (2008), Klauer et al. (2002), Tomic & Kingma (1998) and Molnár (2011) empirically confirmed this statement and reported a significant improvement of inductive reasoning skills as a result of IR training programs developed according to Klauer's model using different contexts, various cultures, ages or different target groups (students with special need, average or gifted students), but all executed as a face-to-face intervention study.

3. Aims and research questions

The objective of this study is twofold. First we examine the applicability and effectiveness of an online inductive reasoning training program based on Klauer's "Cognitive training for children" concept at the age of 9-to 11 in Arabic educational context. We than examine the effectiveness of the intervention program on different groups of students, that is, on students having different starting level of IR, on students having different socio-economic factors and on students having different gender.

We thus intend to answer four research questions: (1) How effectively can Klauer's training concept be adopted to online environment using mathematical content in the learning tasks and how effective is the online training program of inductive reasoning in Arabic educational context at the age of 9-to 11? (2) Which starting level of IR is the most sensitive regarding the training program for? Thus, which level can we expect the largest effect on? (3) Which group of students can be enhanced the most via the training program based on students' socio-economic background? Does the training program favour girls or boys? (4) How generalizable are the results? Are the effects confirmed by latent level analyses using latent change model in the intervention group and a no-change model in the control group?

4. Methods

4.1. Participants

The sample of the study was drawn from fourth and fifth grade students (aged 9–11) in four Palestinian primary schools. A total amount of 236 students participated in the study: 118 students (55.1 % males) were assigned to the intervention group and 118 (52.5 % males) to the control group. The experimental group consisted of 57 fourth and 61 fifth graders, while the control group consisted of 60 fourth and 58 fifth graders. In the participating schools, like schools in Palestine generally, each grade is made up of several forms. Every school year, students are rearranged between the forms based on their school grades to achieve a balanced distribution of low, intermediate and high achievers within each form. Forms are taught the same way, using the same teaching methods. As entire forms were selected for the study, members of both the experimental and control groups were taught for the same number of school hours during the period of time when the training was administered.The retention rate was 100 % in both of the groups. Considering school achievement, 31.4/37.3/31.4 % of the students in the intervention group and 33.4/44.9/22.0 % in the control group belong to the low/intermediate/high achievers. Mother's school level was approximately the same between the two groups (see Table 1).

4.2. Instruments

Similarly to Klauer's original program, the training consisted of 120 learning tasks in total, 20 problems in each class of inductive reasoning (generalization, discrimination, cross-classification, recognizing relations, discriminating between relations, and system formation). All of the learning tasks could have been solved through the application of appropriate inductive reasoning processes and all of the tasks were embedded in various mathematical contents, corresponding to the targeted population: even and uneven numbers, Roman numerals, relationship between numbers and quantities, fundamental operations of arithmetic, usage of relational math symbols, measurements, conversion of units of measurements, series (completing, ordering), pairs of data, correlations among triplets, concepts in geometry, geometric transformations, measurement of time, knowledge of the clock.

The training tasks contained pictures that correspond to the targeted age. Students had to indicate their answer by using the mouse. Operations based exclusively on clicking and drag-and-drop from left to right. Based on earlier research results (Mousa and Molnár, 2019a) the size and number of objects they had to click on or drag and drop could have not influenced significantly the success and difficulty of the training tasks.

According to Klauer's framework (Klauer & Phye, 2008, p. 88) we have used the following inferential formats by the six

Table 1
The distribution of the sample based on mothers' education.

	Control group (%)	Experimental group (%)
Uneducated	12.7	12.7
Elementary school	24.6	25.4
High school certificate	27.1	22.0
Diploma	31.4	28.0
BA and MA	4.2	11.9

operations of IR: class formation, class expansion and finding common attributes (by learning tasks using the process of generalisation), identifying disturbing items (by discrimination), 4, 6 or 9-fold schemes (by cross-classification), series completion, order series and simple analogies (by recognising relationships), disrupted series (by differentiating relationships) and matrices with complex analogies (by system formation). The instructions for the tasks were provided in written form on screen. The language of the instruction was simplified Arabic and we have used the Hindu numeral system. The 120 tasks (20 per operation) were divided into five sessions, that is, one training session contained 24 training tasks. Fig. 1 shows examples of the training tasks for each of the operations. The training is based on a Hungarian training program of IR (Pásztor, 2016).

The effectiveness of the intervention program was measured with a computer-based, internationally widely used test of inductive reasoning, a separate context with different task types (figural series, figural analogies, number series and number analogies) for avoiding the near transfer effect of the training program (see Molnár et al., 2013; Pásztor, Hotulainen, Kupiainen, Molnár, & Csapó, 2018; Wu & Molnár, 2018a, 2018b). The original IR test was adapted to Arabic in both directions and language (see Fig. 2). Since the Palestinian school textbooks use the Hindu numeral system, the items that contained numbers, have been changed into the Hindu numbers. Fig. 2 shows the original and the adapted version of the same item from the test. The test consisted of 44 items in total, comprising no interactive elements. The reliability index of the whole test was in the pretest Cronbach $\alpha = .812$ and in the posttest Cronbach $\alpha = .912$.

4.3. Design and procedures

The study followed a quasi-experimental design, with both the intervention and control groups assessed at two different time points: Before intervention and after the training. Students were randomly assigned to intervention and control group.

The computer-based IR training were administered via the online training platform eLea (Molnár, Pásztor, & Csapó, 2019) and the pre-and-post-test were delivered through the eDia online assessment platform (Csapó & Molnár, 2019; Molnár & Csapó, 2019). The test and the training were administered in the computer labs (using desktop computers) of the participating schools. The training lasted six weeks and happened during regular school hours after taking official acceptance from the directorate of education in the region (Bethlehem district) and consulting with the schools' principals. Each session took approximately one school lesson, that is, forty minutes. Classroom teachers supervised the training.

There was a team of teachers who supervised the training sessions. One week before the training, the implementers received a short (three-hour-long) on-site group training held by the researchers. The aim of this meeting was to provide full information about the theory, platform and process of the training: (1) how thinking skills and abilities, especially inductive reasoning, develop during the age range of schooling; (2) what the characteristics of inductive reasoning targeted in the training are; (3) what the structure of the training looks like; (4) how much time is required for the project; (5) what the training tasks look like – with examples; and (6) how the online training (eLea) and test (eDia) platforms can be used; that is, the teachers were taught how to run the system. During the training, it was agreed how to behave during the training sessions – when to interact with the children and what type of support is allowed during the training if the students ask for assistance – so as to make sure that all the students have the same opportunity to receive assistance. Teachers were allowed to resolve any technical issues and answer questions raised by the students but were not allowed to provide solutions for or make any reference to the right answer on any of the tasks. Teachers also learned how to track students' progress during the training.

Each student got a password and logged in to the system with the help of the class teacher and the ICT teacher who gave a hand for any technical issues and facilitated the process regarding technical issues during the training. Students were provided with instructions at the beginning of the training to learn to use the program and to learn how the training runs. The instruction included the followings: in each session a yellow bar at the top of the screen indicates how far along they are on the test; if the students finished a training task, they have to click on the "next" button to check the goodness of the answer or after a positive feedback moving on to the next task; that is they received feedback immediately after each task whether the answer was right or wrong. If the answer was right, the student received number of stones represents how many stones (right answers) he or she collected so far. If the answer was wrong, the student received extra supporting information beyond the feedback informing him or her that the answer was wrong and they have another chance to solve the training task again. If the student failed for the second time again, he or she had a final, a third chance to solve the same task with much more help. When the student reached the end of a session, they received information about their actual achievement in the training.

In case of the pre- and posttest the assessment system eDia and during the intervention program the training platform eLea logged, recorded and scored automatically the answers of the students, which made immediate feedback on test and on task level possible. Both the pre- and posttest took about one hour to be completed and similarly to the training sessions classroom teachers supervised the datacollection, which happened in the ICT rooms of the participating schools.

Beyond the analyses using observed variables (e.g. ANOVA-family analyses), which have several limitations, based on the paper of Alessandri, Zuffianò, and Perinelli (2017), we applied second order multiple group latent curve modelling in the structural equation modelling framework using latent variables too. Analyses underlying classic parametric tests have the assumption that the measurement structure of the construct under investigation is invariant across groups and/or time and that the data being analyzed have equal population variances and are normally distributed. As the latter one was not completely confirmed by the results, we decided to run SEM analyses to confirm results got on manifest level.

There were only two points in time available in the analyses, that is, it was possible to estimate two latent curve models: A no-change model (or strict stability model, see Alessandri et al., 2017) and a latent change model. In the no-change model both the mean and variance of the second-order intercept factor were freely estimated across groups. It included only a second-order intercept factor

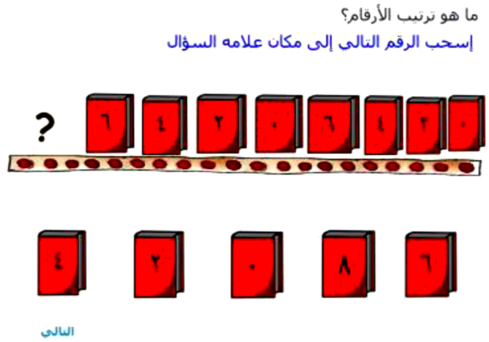
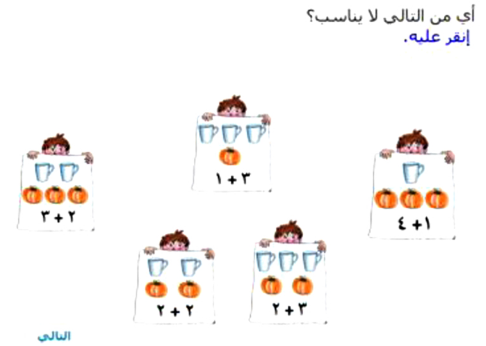
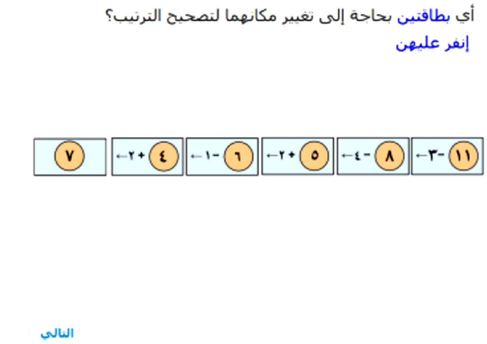
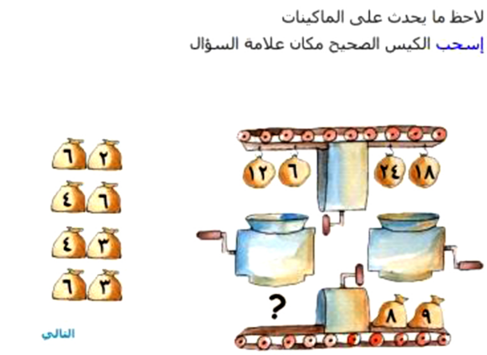
	Attributes of objects	Relations among objects
Similarity	Generalization  Form three groups. Pull the items belonging to the same group onto the selves	Recognizing relations  What is the order of the numbers? Pull the next number to the question mark's place
Differences	Discrimination  Which of the following does not fit? Click on it.	Discriminating relations  Which two cards are needed to be inverted for the correct sequence? Click on them.
Similarity and differences	Cross-classification  The books on the shelves have been arranged according to a certain rule. What can be the rule? On which shelf would you put the separate book according to the rule? Click on it.	System formation  Observe what happens in the machines. Drag the correct bags on the question mark.

Fig. 1. Examples of training tasks.

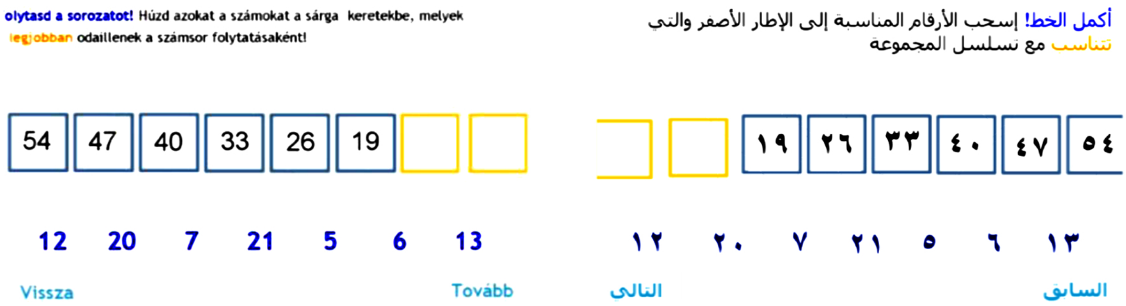


Fig. 2. The original (Hungarian) and the adapted (Arabic) version of the same test item.

representing the students' initial skill level. In the latent change model a slope growth factor was estimated.

Generally, two waves of data are insufficient to estimate latent change models. A possible way to do this, if the latent change model is over-identified, that is, if at least two observed indicators of the construct of inductive reasoning are available at each time point, e.g. there are two scales assessing the same construct in the study. However, the construct of inductive reasoning was assessed by one scale (but the test was composed by several items). But according to Steyer, Eid, and Schwenkmezger (1997), it was possible to partitioning the items composing the scale into two parcels that can be treated as parallel forms. We followed the procedure described in Little, Cunningham, Shahar, and Widaman (2002) than we run three alternative model: (1) no-change model in both groups (experimental and control), (2) latent change model for the experimental and no-change model for the control group, (3) latent change model for both of the groups. We compared model fit indexes of CFI (Comparative Fit Index), TLI (Tucker-Lewis Index) with associated 90 % confidence intervals and RMSEA (Root Mean Square Error of Approximation) and the changes in fit indexes between the different models. We accepted CFI and TLI values > 0.90, RMSEA values < 0.08 (see Kline, 2016).

5. Results

5.1. The effectiveness of the online IR training using mathematical content at the age of 9-to 11

No significant differences were found between the performance of the experimental and the control groups prior to the experiment (Mcont = 33.4, SDcont = 15,0; Mexp = 35.0, SDexp = 13.5, t = 1.3, p = .18). There was no significant development detectable in case of the control group (Mcont = 34.0, SDcont = 14.1), while the experimental group significantly outperformed the control group by more than one standard deviation (Mexp = 58.6, SDexp = 14,5, t = 13.1, p < .001).

The effect size of the training program – using Cohen's (1988) convention for describing the magnitude of effect size shows a clear large effect d = 1.71. Placing the training program of inductive reasoning into an international context according to its effect size even independent of its delivery media, it allows us for favourable conclusions. The effect size was as high as the effect size published in the literature (Klauer & Phye, 2008; Molnár, 2006) regarding face-to-face training programmes in non-academic context using pre- and-posttest developed in accordance of the Klauer's model. Our result indicates that the newly developed online training program of inductive reasoning can be effectively used in mathematics lessons for improving students IR skills and the development was even detectable via IR tests, which were not developed in accordance of the Klauers' model for avoiding the near transfer effect of the training.

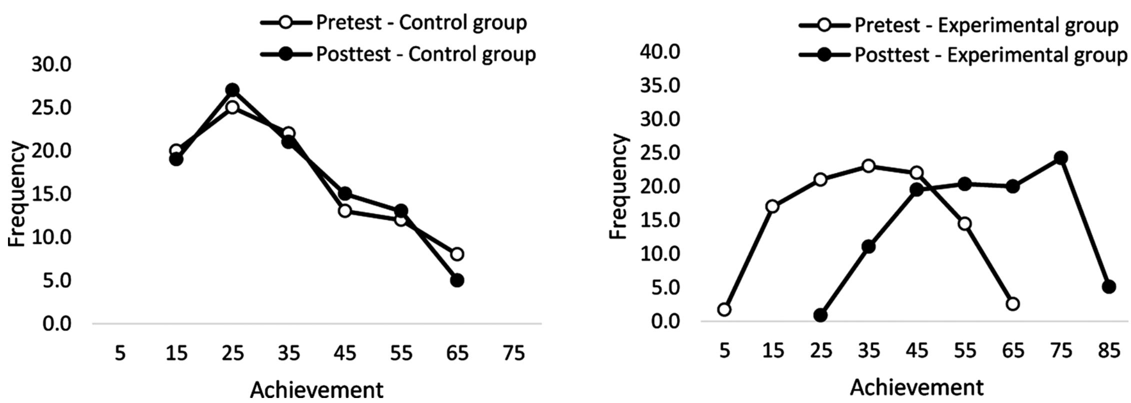


Fig. 3. Distribution curves of the control and the experimental groups in the pre- and posttest.

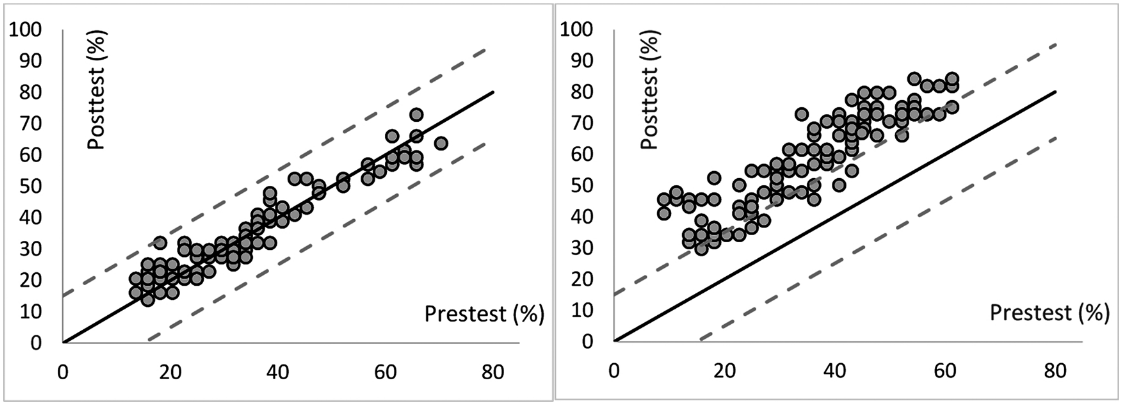


Fig. 4. Student level changes of the achievement from pretest to posttest in both of the control and the experimental groups.

5.2. The changes in performance with regard to students' original level of inductive reasoning skills

In the case of the control group, the distribution curve for both the pre- and posttest (see Fig. 3) is inclined to the left. The two curves are laying on each other, indicating no "spontaneous "development in student's IR skills in the given period of time based on the normal school learning. The posttest distribution curve of the experimental group turned into one skewed to the right reflecting an immense improvement at each level of IR by the members of the experimental group. Based on the group level distribution curves we might hypothesise that each member of the experimental group attained significant improvement in performance as a result of the training.

The group level results above are supported by the student level analyses, visualized in Fig. 4, where the performances in the pre- and posttest are projected onto each other. The abscissa shows the achievement obtained in the first data collection and the ordinate displays this from the second. The dots are representing the students. Those students, whose symbols are falling on the mean line or between the two broken lines (representing one standard deviation) performed identically in the two cases. If the symbol is positioned above the broken line, it means that the given student showed significant development from the pretest to the posttest, while if it is below the broken line, it represents significantly worse performance from the first to the second data collection.

In the case of the control group (graph on the left), the dots are distributed around the mean line, between the broken lines; that is, all of the students in the control group performed significantly at the same level, quite similarly in the two data collection. A completely different tendency is displayed on the right-hand graph, showing the performance of the experimental group before and after the training. Almost all of the symbols are located above the broken line, that is, on one hand there are no students whose performance dropped significantly from pretest to posttest in the given time frame, on the other hand almost all of the students' IR level in the experimental group improved by more than one standard deviation as a result of the training. To sum up, the training resulted in a significant ($p < .001$) improvement for the experimental group in the field of inductive reasoning.

The influential factor of students gender, school achievement and socio-economic background on the magnitude of their development in inductive reasoning

The paired *t*-test results indicated no gender level differences in the training effect, that is, there were no significant changes detectable in the control group, while there were statistically significant differences detectable in the experimental group in both subgroups. According to the independent *t*-test there were no significant gender level differences in the pretest or in the posttest results (see Table 2), therefore the training proved to be gender fair.

Mother's educational background correlated strongly with the students' level of IR in both time points ($r_{pre} = .667$, $r_{post} = .555$, $p < .001$) and in both groups (control group: $r_{pre} = .718$, $r_{post} = .741$; experimental group: $r_{pre} = .626$, $r_{post} = .642$; $p < .001$). It was confirmed by the ANOVA analyses too. According to the Tukey B analyses there were four significantly different achieving groups distinguished in both time points: group of uneducated or on elementary level educated mothers' children ($M_{pre} = 20.9/24.2$; $M_{post} = 32.4/36.1$), the group of children, whose mother has high school certificate ($M_{pre} = 35.6$, $M_{post} = 46.60$, diploma ($M_{pre} = 44.5$, $M_{post} = 55.6$), BA or MA degree ($M_{pre} = 49.5/52.3$, $M_{post} = 65.6/77.3$). That is, students' level of IR skills was generally higher if their mother education was better and vice versa.

Table 2
Within and between gender level differences in the pre- and posttest.

Group		N	Correlation	Diff. in Mean	SD	t_pre_post (p)	t_pre (p)	t_post (p)
Control group	Male	62	.964	.84	4.4	1.52 (.14)	−.97 (.34)	−.95 (.34)
	Female	56	.958	.52	4.5	.87 (.39)		
Experi-mental group	Male	65	.887	24.05	7.4	26.11 (.001)	−.65 (.52)	.54 (.59)
	Female	53	.924	21.01	5.0	30.71 (.001)		

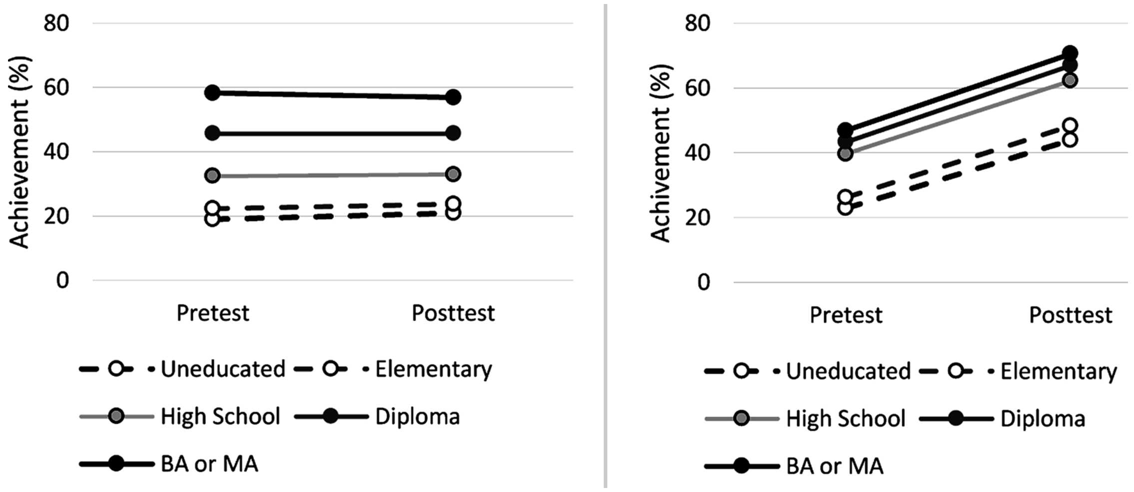


Fig. 5. Trajectories of inductive reasoning skills for control group (graph on the left) and for the intervention group (graph on the right) according to the mothers education.

Fig. 5 compares performance differences of students having mothers with different educational background. In case of the control group (graph on the left), no change can be detected in either of the groups. While, in case of the experimental group (graph on the right) the trajectories visualising the effects of the training are running parallel, indicating similar training effect in all of the groups. That is, non of the subgroups was favoured regarding the efficacy of the training in the experimental group.

Finally, students were divided into three subgroups regarding their achievement in their schools. There was a strong relation between school achievement and the level of IR detectable (Pretest: M_advanced = 53.78, M_average = 33.99; M_low = 19.67, F = 460.43, $p < .001$; r = .889, $p < .001$) (posttest: M_advanced = 67.06, M_average = 44.89; M_low = 31.07, F = 136.16, $p < .001$; r = .726, $p < .001$). Fig. 6 illustrates the trajectories of IR in the control and experimental group based on students' school achievement, respectively. The developmental curves are running parallel in case of the experimental group, that is, the training had the same effect on all of the students independent of their school achievement and level of starting IR skills.

5.3. Evaluating the effect of the intervention program in the latent curve modelling framework

First we tested a measurement model for inductive reasoning with all indicators combined under one general factor. We used the preferred estimator for categorical variables, Weighted Least Squares Mean and Variance adjusted (WLSMV; Muthén & Muthén, 2010). The measurement model based on the pretest results showed a good fit ($\chi2$ = 974.9; df = 902; $p < .05$; CFI = .940; TLI = .937; RMSEA = .019 (CI: .003,.027)).

We created two parallel forms of the inductive reasoning scale based on the factor loading values. Cronbach's alphas were good ($\geq$0.79), and correlations above .89. Table 3 shows the fit indexes for the three alternative models. As hypothesized, the mixed, that is, the second model was the best fitting model (no-change model for the control group and latent change model for the experimental group), however, the RMSEA value were still higher than accepted for a good model fit. All of the fit indices of the other two models fall below the accepted values.

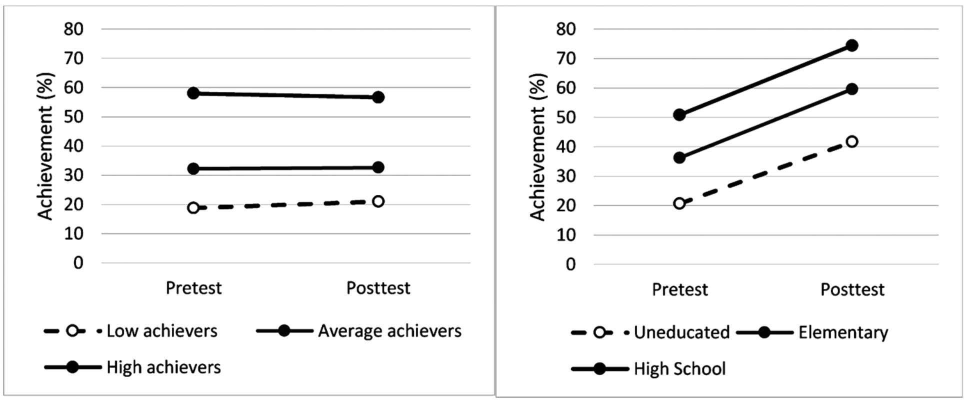


Fig. 6. Trajectories of inductive reasoning skills for control group (graph on the left) and for the intervention group (graph on the right) according to the mothers education.

Table 3
Goodness-of-fit indices for the tested models.

Model	χ^2 (df)	CFI	TLI	RMSEA [90 % CI]
No-change model for both of the groups	615.8 (12)	.602	.602	.653 [.610, .697]
No-change model for the control and latent change model for the experimental group	47.7 (11)	.974	.961	.205 [.151, .263]
Latent change model for both of the groups	312.6 (10)	.801	.761	.506 [.459, .555]

The results also confirmed that there was no significant variability among students in responding to the intervention program, as indicated by the non-significant variance of the latent slope. Non-of the students were more sensitive or responsive to the intervention delivered.

6. Discussion

This study presented a computer-based training program of inductive reasoning for 9–11-year-old students in educational context and addressed the direct result of the evaluation study. The online intervention program was build according to Klauer's model and "Cognitive training for children" concept of inductive reasoning, thus it consists of 120 playful problems, but in online environment. All of the problems were embedded in Mathematics content to make the integration and the applicability of the training during normal school hours as part of the Mathematics lesson easier.

The study followed a quasi-experimental design. The developmental level of the control and experimental group did not differ prior to the experiment, thus, the control group meets the requirement for a control group. As a result of the online training, the inductive reasoning skills of the experimental group showed significant improvement, while students' IR skills remained at the same level in the control group. This result confirms previous research results regarding (1) the usability and effectiveness of technology-based trainings at school level (see Mousa & Molnár, 2018, Mousa and Molnár, 2019a, 2019b; Molnár & Pásztor, 2015); (2) the possibility of explicit fostering of inductive reasoning during normal school hours (de Koning et al., 2002; Molnár, 2011; Adams, 1989; Nisbet, 1993; Resnick & Klopfer, 1989), (3) and the lack of explicit fostering in the school (de Konig, 2000; Molnár, 2011). Thus we confirmed that Klauer's concept of cognitive training for children can be applied not only in face-to-face, but in an online environment and not only with everyday, but in school subject context – in the present case mathematics – too.

The effect size of the training was exceptionally good not only in Arabic context but internationally as well. It proved to be unrelated to gender, to mother's education (as a socio-economic indicator (Wegerif, Li, & Kaufman, 2017)– the mother plays the major role in the education of the children at home in Palestina), to school achievement and to the original IR level of the students; i.e. it had a similar effect on boys and girls, on students having different educated mothers and got different stimulus and instruction at home, on students with low, average or high school achievement and on students having low or high starting level of IR at the age of 9–11. This result was confirmed by the structural equation modelling analyses, where we applied three model, three different combinations of no-change and latent change models in both the intervention and control group. The best fitting trajectory (no-change model in the control group and latent change model in the experimental group) confirmed results on manifest level. The non-significant variance of the latent slope indicated and proved previous results that there was no significant variability among students in responding to the intervention program. Generally, non-of the students were more sensitive to the intervention delivered.

To sum up, the results suggest that the elaboration of this online domain-specific intervention program can be considered successful. It develops 9-to-11-years-old students' inductive reasoning in a playful way. The findings suggest that inductive reasoning skills can significantly and effectively develop not only in traditionally face-to-face, but in computer-based environment too between the ages of 9 and 11. An online test of inductive reasoning was also constructed in Arabic as part of the program package. According to its structural validity and reliability it can be effectively applied to assess the developmental level of elementary students' inductive reasoning skills even independently of the training program.

7. Limitations of the study

Limitations of the study include the procedure of dividing the students into control and experimental groups, resulting in two groups with the same level of inductive reasoning skills, but different socio-economic background factors, e.g. mother's level of education. Further repetition is required to validate the results with a larger sample and groups with not only the same average level of inductive reasoning skills, but also other background factors.

To avoid the near-transfer effect of the training, different models and different types of tasks were used in developing the pre- and post-test tasks. While the training tasks were developed according to Klauer's model of inductive reasoning to activate generalisation, discrimination, recognising relations, discriminating between relations and cross-classification, the test consisted of tasks involving figural and numerical analogies and series. This resulted in a more valid assessment of the effect size of the training. It also ruled out the option of analysing the effect size of the training based on the pre- and post-test data on the dimensional level. The dimensional-level analyses require logfile analyses of the training sessions, which might provide more detailed information on the training tasks themselves and help to improve the effectiveness of the programme. This forms part of our future plans, but is also among the limitations of the present analysis.

Funding

This study was funded by the Hungarian Academy of Sciences through the MTA–SZTE Research Group on the Development of Competencies, by OTKA K115497, by the Center for Research on Learning and Instruction, and by the EFOP 3.2.15 project.

CRediT authorship contribution statement

Mojahed Mousa: Software, Data curation, Writing - original draft, Validation. **Gyöngyvér Molnár:** Conceptualization, Methodology, Software, Visualization, Investigation, Supervision, Validation, Writing - review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Preparation of this article was funded by the Hungarian Academy of Sciences through the MTA–SZTE Research Group on the Development of Competencies, by OTKA K115497, by the Research Group on Learning and Instruction and by the EFOP 3.2.15 project.

References

Adams, M. J. (1989). Thinking skills curricula: Their promise and progress. *Educational Psychologist, 24*(1), 25–77.

Adey, P., Csapó, B., Demetriou, A., Hautamäki, J., & Shayer, M. (2007). Can we be intelligent about intelligence?: Why education needs the concept of plastic general ability. *Educational Research Review, 2*(2), 75–97. https://doi.org/10.1016/j.edurev.2007.05.001.

Alessandri, G., Zuffianò, A., & Perinelli, E. (2017). Evaluating intervention programs with a pretest-posttest design: A structural equation modeling approach. *Frontiers in Psychology, 8*, 223. https://doi.org/10.3389/fpsyg.2017.00223.

Barkl, S., Porter, A., & Ginns, P. (2012). Cognitive training for children: Effects on inductive reasoning, deductive reasoning, and mathematics achievement in an Australian school setting. *Psychology in the Schools, 49*(9), 828–842.

Bisanz, J., Bisanz, G., & Korpan, C. A. (1994). Inductive reasoning. In R. Sternberg (Ed.). *Thinking and problem solving*San Diego: Academic Press. https://doi.org/10.1016/B978-0-08-057299-4.50012-8.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Csapó, B. (1997). The development of inductive reasoning: Cross-sectional measurements in an educational context. *International Journal of Behavioral Development, 20*(4), 609–626.

Csapó, B., Lőrincz, A., & Molnár, G. (2012). Innovative assessment technologies in educational games designed for young students. *Assessment in Game-Based Learning,* 235–254. https://doi.org/10.1007/978-1-4614-3546-4_13.

Csapó, B., & Molnár, G. (2019). Online diagnostic assessment in support of personalized teaching and learning: The eDia system. *Frontiers in Psychology*. https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01522/full.

Csapó, B., & Nikolov, M. (2009). The cognitive contribution to the development of proficiency in a foreign language. *Learning and Individual Differences, 19*, 209–218. https://doi.org/10.1016/j.lindif.2009.01.002.

de Konig, E. (2000). *Inductive reasoning in primary education: Measurement, teaching, transfer.* Zeist, Kerckebosch.

De Koning, E., & Hamers, J. H. M. (1999). Teaching inductive reasoning: Theoretical background and educational implications. *Teaching and learning thinking skills,* 157–188.

de Koning, E., Hamers, J. H., Sijtsma, K., & Vermeer, A. (2002). Teaching inductive reasoning in primary education. *Developmental Review, 22*(2), 211–241.

Dienes, Z. P. (1963). *An experimental study of mathematics learning.* London: Hutchinson.

Dienes, Z. (1973). *(Building up mathematics) Építsük fel a matematikát.* Budapest: Gondolat Kiadó.

Feuerstein, R., Rand, Y., Hoffman, M. B., & Miller, R. (1980). *Instrumental enrichment: An intervention programme for cognitive modifiability.* Baltimore, MD: University Park Press.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), 1–38. https://doi.org/10.1016/0010-0285(83)90002-6.

Goswami, U., & Brown, A. L. (1990). Higher-order structure and relational reasoning: Contrasting analogical and thematic relations. *Cognition, 36*(3), 207–226. https://doi.org/10.1016/0010-0277(90)90057-Q.

Hamers, J. H. M., De Koning, E., & Sijtsma, K. (1998). Inductive reasoning in third grade: Intervention promises and constraints. *Contemporary Educational Psychology, 23*(2), 132–148.

Hamers, J. H. M., De Koning, E., & Sijtsma, K. (2000). Inductive reasoning in the third grade: Intervention promises and constraints. *Contemporary Educational Psychology, 23,* 132–148. https://doi.org/10.1006/ceps.1998.0966.

Hattie, J., & Anderman, E. M. (Eds.). (2013). *International guide to student achievement.* Routledge.

Hayes, B., Heit, E., & Swendsen, H. (2010). Inductive reasoning. *Wiley Interdisciplinary Reviews Cognitive Science, 1*(2), 278–292. https://doi.org/10.1002/wcs.44.

Klauer, K. (1996). Begünstigt induktives Denken das Lösen komplexer Probleme? Experimentellen Studien zu Leutners sahel-Problem [Has inductive Thinking a positive effect on complex problem solving?]. *Zeitschrift für Experimentelle Psychologie, 43*(1), 361–366 (Dutch version).

Jensen, I. L. (1973). *Acts: An Inductive Study.* Moody Publishers.

Klauer, K. (1989). *Denktraining für kinder [Cognitive Training for Children 1].* Göttningen: Hogrefe (Dutch version).

Klauer, K. J. (1990). Paradigmatic teaching of inductive thinking. In H. Mandl, E. De Corte, S. N. Bennett, & H. F. Friedrich (Eds.). *Learning and instruction. European research in an International context. Analysis of complex skills and complex knowledge domains* (pp. 23–45). Oxford: Pergamon Press.

Klauer, K. J. (1991). *Denktraining für Kinder II.* Göttingen: Hogrefe.

Klauer, K. J. (1993). *Denktraining für Jugendliche.* Göttingen: Hogrefe.

Klauer, K. J., & Phye, G. D. (1994). *Cognitive training for children. A developmental program of inductive reasoning and problem solving.* Seattle, WA: Hogrefe & Huber.

Klauer, K. J., & Phye, G. D. (2008). Inductive reasoning: A training approach. *Review of Educational Research, 78*(1), 85–123.

Klauer, K., Willmes, K., & Phye, G. (2002). Inducing inductive reasoning: does it transfer to fluid intelligence? *Contemporary Educational Psychology, 27*(1), 1–25.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th edn.). New York, NY: The Guilford Press.

Lipman, M. (1985). Thinking skills fostered by Philosophy for Children. In J. W. Segal, S. F. Chipman, & R. Glaser (Eds.). *Thinking and learning skills* (pp. 83–108). Hillsdale, NJ: Lawrence Erlbaum Associates.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*, 151–173. https://doi.org/10.1207/S15328007SEM0902_1.

McGaw, B., Griffin, P., & Care, E. (2012). *Assessment and teaching of 21st century skills.* NY: Springer (eds.).

Molnár, G. (2006). Az induktív gondolkodás fejlesztése kisiskolás korban [Fostering inductive thinking in elementary school pupils]. *Magyar Pedagógia, 1*, 63–80.

Molnár, G., & Csapó, B. (2011). Az 1–11 évfolyamot átfogó induktív gondolkodás kompetenciaskála készítése a valószínuségi tesztelmélet alkalmazásával. *Magyar Pedagógia, 111*(2), 127–140.
Molnár, G., & Csapó, B. (2019). *How to make learning visible through technology: The eDia-online diagnostic assessment system. Proceedings of the 11th International Conference on Computer Supported Education (CSEDU 2019), Vol. 2*, 122–131.
Molnár, G., & Lorincz, A. (2012). *Innovative assessment technologies: Comparing "face-to face" and game-based development of thinking skills in classroom settings. 2012 International Conference on Management and Education Innovation, 37*, 150–154.
Molnár, G., & Pásztor, A. (2015). Feasibility of computer-based assessment at the initial stage of formal schooling: the developmental level of keyboarding and mouse skills in year one. *Conference on educational assessment,* 118.
Molnár, G., Greiff, S., & Csapó, B. (2013). Inductive reasoning, domain specific and complex problem solving: Relations and development. *Thinking Skills and Creativity, 9*(1), 35–45.
Molnár, G., Pásztor, A., & Csapó, B. (2019). The eLea online training platform. In E. K. Molnár, & K. Dancs (Eds.). *XVII. Pedagogical assessment conference - 17th conference on educational assessment. Program and abstracts* (pp. 64). .
Mousa, M., & Molnár, G. (2018). *Introducing computer-based assessment among 2nd to 4th grade students in Palestine. 16th Conference on educational assessment, 16*, 71.
Mousa, M., & Molnár, G. (2019a). The feasibility of computer-based testing in palestine among lower primary school students: Assessing mouse skills and inductive reasoning. *Journal of Studies in Education, 9*(2), 1. https://doi.org/10.5296/jse.v9i2.14517.
Mousa, M., & Molnár, G. (2019b). Applying computer-based testing in palestine: Assessing fourth and fifth graders inductive reasoning. *Journal of Studies in Education, 9*(3), 1. https://doi.org/10.5296/jse.v9i3.14799.
Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide.* Los Angeles, CA: Muthén & Muthén.
Nisbet, J. (1993). The thinking curriculum. *Educational Psychology, 13*(3–4), 281–290.
Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological review, 97*(2), 185.
Pásztor, A. (2016). *Technology-based assessment and development of inductive reasoning. Ph.D. Thesis.* Doctoral School of Education, University of Szeged.
Pásztor, A., Hotulainen, R., Kupiainen, S., Molnár, G., & Csapó, B. (2018). Comparing Finnish and Hungarian fourth grade students' inductive reasoning skills. *EARLI SIG 1 Conference*.
Pellegrino, J. W., & Glaser, R. (1982). Analyzing aptitudes for learning: Inductive reasoning. In R. Glaser (Ed.). *Advances in instructional psychology* (pp. 269–345). Hillsdale, NJ: Lawrence Erlbaum Associates.
Perret, P. (2015). Children's inductive reasoning: Developmental and educational perspectives. *Journal of Cognitive Education and Psychology, Journal of Cognitive Education and Psychology, 14*(3), 389–408. https://doi.org/10.1891/1945-8959.14.3.389 2015.
Resing, W. C. M. (1993). Measuring inductive reasoning skills: The construction of a learning potential test. In J. H. M. Hammers, K. Sijtsma, & A. J. J. Ruijssenaars (Eds.). *Learning potential assessment: Theoretical, methodological and practical issues* (pp. 219–242). Amsterdam: Swets and Zeitlinger.
Resnick, L. (1987). *Education and learning to think.* Washington, DC: National Academy Press.
Resnick, L. B., & Klopfer, L. E. (1989). *Toward the thinking curriculum: Current cognitive research. 1989 ASCD yearbook*. 1250 N. Pitt St., Alexandria, VA 22314-1403: Association for Supervision and Curriculum Development.
Ropo, E. (1987). Skills for learning: A review of studies on inductive reasoning. *Scandinavian Journal of Educational Research, 31*, 31–39.
Sandberg, E.H., & McCullough, M.B. (2010). The development of reasoning skills. A clinician's guide to normal cognitive development in childhood, 179-189.
Shayer, M., & Adey, P. (1981). *Towards a science of science teaching: Cognitive development and curriculum demand.* London: Heinemann Educational Books.
Sloman, S. A. (1993). Feature-based induction. *Cognitive psychology, 25*(2), 231–280.
Söderqvist, S., Nutley, S. B., Ottersen, J., Grill, K. M., & Klingberg, T. (2012). Computerized training of non-verbal reasoning and working memory in children with intellectual disability. *Frontiers in Human Neuroscience, 6*, 271. https://doi.org/10.3389/fnhum.2012.00271.
Sternberg, R. J., & Gardner, M. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology General, 112*, 80–116.
Steyer, R., Eid, M., & Schwenkmezger, P. (1997). Modeling true intraindividual change: True change as a latent variable. *Methods Psychol. Res. Online, 2*, 21–33.
Tomic, W. (1995). Training in inductive reasoning and problem solving. *Contemporary educational psychology, 20*, 483–490.
Tomic, W., & Kingma, J. (Eds.). (1998). *Conceptual Issues in Research in Intelligence*. JAI Press.
Tomic, W., & Klauer, K. J. (1996). On the effects of training inductive reasoning: How far does it transfer and how long do the effects persist? *European Journal of Psychology of Education, 3*, 283–299.
Vainikainen, M.-P., Hautamaki, J., Hotulainen, R., & Kupiainen, S. (2015). General and specific thinking skills and schooling: Preparing the mind to new learning. *Thinking Skills and Creativity, 18*, 53–64. https://doi.org/10.1016/j.tsc.2015.04.006.
Voogt, J., Erstad, O., Dede, C., & Mishra, P. (2013). Challenges to learning and schooling in the digital networked world of the 21st century. *Journal of Computer Assisted Learning, 29*(5), 403–413.
Wegerif, R. (2006). A dialogic understanding of the relationship between CSCL and teaching thinking skills. *International Journal of Computer-Supported Collaborative Learning, 1*(1), 143–157.
Wegerif, R., Li, L., & Kaufman, J. C. (2017). *The Routledge international handbook of research on teaching thinking.* Oxford.
Wu, H., & Molnár, G. (2018a). Computer-based assessment of chinese students' component skills of problem solving: A pilot study. *International Journal of Information and Education Technology, 8*(5), https://doi.org/10.18178/ijiet.2018.8.5.1067 381-356.
Wu, H., & Molnár, G. (2018b). Interactive problem solving: Assessment and relations to combinatorial and inductive reasoning. *Journal of Psychological and Educational Research, 26*(1), 90–105.