Thinking skills intervention for low-achieving first graders

Risto Hotulainen\textsuperscript{a}, Riikka Mononen\textsuperscript{b} and Pirjo Aunio\textsuperscript{b}

\textsuperscript{a}Department of Teacher Education, Centre of Educational Assessment, University of Helsinki, Helsinki, Finland; \textsuperscript{b}Department of Teacher Education, Special Education, University of Helsinki, Helsinki, Finland

\textbf{ABSTRACT}
This paper reports the results of the improving thinking skills (ITS-1) intervention study on the thinking skills of low-achieving first graders. The intervention programme consists of 12 lessons, each lasting for 45 min. Lessons offer enriched-discovery learning activities and tasks to be solved through inductive reasoning. We used a quasi-experimental approach, with pre-, immediate post- and delayed post-tests conducted among intervention and control groups. The following groups were formed from a total of 149 first graders on the basis of the thinking skills measure: low-achieving (<−1 SD) intervention group (LowI) (n = 9) and two control groups, comprising both low-achieving (LowC) (n = 18) and well-performing (≥−1 SD) groups (WellC) (n = 122). Thinking skills, mathematical skills, listening comprehension skills and reading fluency were measured. The results showed that in the beginning of the study, there were differences in thinking skills, mathematical skills, listening comprehension skills and reading fluency between the LowI and WellC groups, but the LowI group was able to reach the level of their well-performing peers at the end of the intervention in all measures. The discussion focuses on the implications of intervention research, educational practice and responsiveness to intervention.

\textbf{ARTICLE HISTORY}
Received 28 September 2015
Accepted 7 January 2016

\textbf{KEYWORDS}
Inductive reasoning; intervention; low achieving; thinking skills; first graders

\textbf{Introduction}

Everyday life provides numerous situations in which people are both voluntarily and involuntarily engaged in problem-solving. For example, people compare the ripeness of watermelons according to their colour, hardness, smell, heaviness and size and make decisions on whether similar criteria can be applied to determine the ripeness of other fruits. Of course, some people are more adroit than others in applying such criteria, probably depending on their earlier experiences and the guidance and teaching they received (Kuhn 2010). In psychology, these comparison and decision-making processes are often viewed as higher order thinking skills (Klauer and Phye 2008), which are often operationalised as inductive and deductive reasoning (as done in this study), but many other definitions and their combinations are available as well (see e.g. Young 1997; Neber 2010). Briefly, inductive reasoning...
refers to deriving a general principle from specific examples, whereas deductive reasoning means applying and modifying these derived principles to understand specific cases (e.g. Klauer 1989).

Higher order thinking skills also occupy a central place in the literature on learning (Whitehead 1958). Many researchers regard knowledge building in formal schooling as based on perceiving similarities and differences by using skills in categorising and organising information, which in turn has been shown to have a transfer effect on learning school subjects (Bloom et al. 1956; Holland et al. 1989; Adey and Shayer 1994; Tomic and Klauer 1996; Adhami and Yates 2008). The recent evidence that these skills can be successfully taught and learners can be trained has increased the interest in studying how the promotion of thinking skills among low-achieving children works to enhance their performance level and prevent plausible learning difficulties (Feuerstein et al. 1980; Klauer 1989; Klauer and Phye 2008; Organisation for Economic Co-operation and Development [OECD] 2010; Molnár 2011).

The purpose of this study was to demonstrate how thinking skills can be taught successfully to low-achieving, Finnish first graders through an invented improving thinking skills in first grade (ITS-1) intervention programme and to examine if our programme has lasting transfer effects. Here, an intervention programme is defined as a studied practice of providing high-quality instruction that matches the students’ need to attain the set, age-appropriate goals and supporting low achievers accordingly before they fail and end up being labelled as students with special education needs (SEN) (see e.g. Slavin 1996; Batsche, Kavale, and Kovaleski 2006). The thinking-oriented education has recently received considerable interest in the literature and the organisational level, but such practices and their effects, especially on low achievers, have been under-researched (European Parliament and Council 2006; Rocard et al. 2007; OECD 2010).

**Finnish three-tiered support model**

The most recent changes in educational legislation (642/2010) regarding remedial support and special education in Finland took place in 2011 (Graham and Jahnukainen 2011). During the reform, the traditional division into general and special education was replaced by a three-tiered support model, comprising a new tier two (intensified support) between tiers one (general support) and three (special support). Thus, there is a strong emphasis on first, developing and second, using field tested, supportive skill-based activities and programmes for levels one and two to prevent the occurrence of possible learning challenges before students are officially transferred to tier three, which corresponds to the earlier SEN status (cf. Kearns and Fuchs 2013). So far, most free-of-charge and field-tested supportive activities and intervention programmes for five- to eight-year-old children have targeted the improvement of students’ math or reading skills, but the programmes for thinking skills have been missing in Finland. The ITS-1 was planned to respond to this need by providing small group activities to improve the thinking skills of children, and its development was funded by the Finnish Ministry of Education and Culture. The ITS-1 was piloted in the spring of 2012, with 17 low-achieving and 25 average children (Hotulainen and Linnansaari 2012).
Low achievers and need for thinking skills intervention

The ultimate goal of education is for students to develop their skills in school, achieve their full potential and attain success in life. The school-based curricula do not seem to sufficiently emphasise the systematic cultivation of thinking skills (Tomic and Klauer 1996; Kuhn 2005; Molnár 2011). Ongoing curriculum reform of basic education underlines thinking-oriented learning, but details regarding teaching or assessment approaches, are still missing (Finnish National Board of Education [FNBE] 2004, 2014). The promotion of thinking skills is infused into the curricula, and the curricula are expected to enhance the children’s thinking skills with increasing content complexity (spiral curriculum) (e.g. Bruner 1960). However, studies have shown that the promotion of thinking skills is most efficient when these are integrated purposefully in the curricula or explicitly taught (Nisbet 1993; Adey and Shayer 1994; Kuhn 2005; Dewey and Bento 2009; Bradley et al. 2012).

Earlier studies related to reading and math have shown that student performance levels vary considerably due to their informal learning experiences and their parents’ socioeco-nomic status (SES), and learning difficulties and differences in reading and math performance levels seem to remain throughout formal schooling (Adey, Robertson, and Venville 2002; Turkheimer et al. 2003; Dion et al. 2010; Welsh et al. 2010; Aubrey, Ghent, and Kanira 2012; Oliver, Venville, and Adey 2012; Aunio, Hautamäki, and Van Luit 2005). Thus, both existing and experienced difficulties in learning can lead to negative cumulative effects, such as cognitive and affective barriers to learning, which can jeopardise the fulfilment of the above-mentioned goals on the individual level (Slavin et al. 1991; Slavin 1996; Hotulainen et al. 2010; Möller and Pohlmann 2010). For this reason, educational interventions should be targeted as early as possible, for example, closely tied to the school entry (Molnár 2011).

Accordingly, if formal schooling aims to improve the children’s general thinking ability, it would be reasonable to identify and support the students who have weaknesses in this area (Adey et al. 2007). We argue that by promoting the development of thinking skills, we simultaneously construct the platform for formal and informal learning (which transfers its effects across the school curricula), as well as for higher levels of learning (Adey, Robertson, and Venville 2002; Higgins et al. 2005; Shayer 2008).

Intervention programmes designed to promote thinking skills

The skill-based programmes designed to promote thinking skills can be divided into two categories, according to the explicitness of the training, as follows: the content-based method of intervention programmes, where thinking skills are taught implicitly within the school subjects, and intervention programmes, where thinking skills are taught explicitly outside any school subject (McGuinness and Nisbett 1991). The first category, the content-based method, consists of programmes that ensure that thinking skills are taught within conventional curriculum areas (e.g. mathematics, science, etc.). For instance, the Cognitive Acceleration (CA) Programme (Adey and Shayer 1994) represents this method. The original programme was integrated into both the science and the math curricula (Adey et al. 2007). The CA’s principles for boosting cognitive development are based on Piagetian cognitive stages, from concrete to late formal operations, along with encounters with planned cognitive conflicts and Vygotskian scaffolded teaching (Piaget 1970).
The programmes belonging to the second category require children to reflect and to determine what thinking skills and related problem-solving strategies are needed when proceeding stepwise towards a solution. Feuerstein et al.'s (1980) Enrichment Programme is one of the first that is based on explicit teaching. Klauer’s (1989, 1991) Cognitive Training for Children I (Denktraining für Kinder I) programme is based on his own theory of inductive reasoning. It consists of lessons aiming to identify the similarities between attributes and relations and to show how to solve the identified problems by inductive reasoning. Klauer and Phye (2008) introduced the results of 74 training experiments conducted with nearly 3600 children in various age groups, comprising average children and children with SEN. Klauer and Phye (2008) used their explicit teaching method and argued that there was ample evidence that Klauer’s programme would permit a reasonable transfer effect on both fluid intelligence and various academic subjects.

Klauer’s programme shares some features of strategic learning that have been demonstrated as advantageous for children with SEN (Forness, Kavale, and Blum 1997; Forness 2001). On the other hand, the CA has been shown to work with all children; nonetheless, some children benefit more than others, and the gains made have not been found to depend on the starting levels of cognitive development (Adey and Shayer 1993). While our project aimed to develop a thinking skills intervention programme for low-achieving first graders, we concluded that it would be beneficial to use the principles of both above-mentioned programmes. Here, we have deliberately left out cognitive-focused approaches, considering them as representing a more stringent way to tackle problems of those children who do not benefit from skill-based interventions (Kearns and Fuchs 2013).

Research questions

This study aimed to examine if the ITS-1 intervention programme would work for the low-achieving children. Previous studies involving young, low-achieving children suggested the effectiveness of training programmes that emphasised the explicit teaching of thinking skills (Feuerstein et al. 1980; Klauer and Phye 2008).

The following research questions guided our investigations:

(1) To what extent does the ITS-1 intervention programme promote thinking skills in the intervention group comprising low-achieving children?
(2) Do the children belonging to the intervention group obtain higher gain scores in the post- and delayed post-tests used to measure thinking, mathematical and verbal skills compared to the children belonging to the other groups in the study?

Methodology

Participants

A total of 149 (48% female) first graders ($M_{age} = 74.02$ months, $SD = 3.62$ months) from three southern cities of Finland participated in this study. This study was part of a larger research project, which also included mathematics intervention programmes for kindergarten, first- and second-grade pupils (name of the project). Written authorised permissions to participate in the study were gathered from the both municipalities and the children’s parents. Two
qualified teachers out of the participating teachers \( n = 12 \) freely helped in conducting the ITS-1 intervention programme. This meant that we had only two intervention groups. To identify the low-achieving children in the group, we set a cut-off point for scoring with the minus one standard deviation \((-1 SD)\) below the total mean score (Murphy et al. 2007). For the group identification, we used the assessment of thinking skills in the first grade (ATS-1) measure, which was later employed in the immediate post- and delayed post-test phases as well. Thus, three groups were formed, as follows: an intervention group of low-achieving children (LowI) who were receiving thinking skills intervention \( n = 9 \), and two control groups who were not receiving the intervention, comprising both low-achieving children (LowC) \( n = 18 \) and children performing better \((x \geq -1 SD)\) in the thinking skills measure (WellC, \( n = 122 \) ) (Table 2 presents the pre-test results by grouping). Both the LowI and LowC groups had equal gender ratios. The family educational level was categorised into three types: (1) basic level of education (i.e. nine years of compulsory education), (2) upper secondary education (i.e. three-year vocational or academic track) and higher education (i.e. university-level degree). For the children’s mothers, the percentages from the basic to the higher levels were 7.2, 40.1 and 52.7%, respectively; for the children’s fathers, the percentages were 15.0, 47.2 and 37.8%, respectively, which corresponded to the average percentages among Finnish adults (Myrskylä 2009).

**Measures**

Thinking skills were assessed by using the ATS-1 (Authors 2012). It is a group-based, paper-and-pencil test that includes inductive reasoning tasks, such as classifying items according to attributes and performing comparisons. All the questions used \( n = 32 \) are formulated at the age- and skill-appropriate levels, proving normally distributed outcomes. The related items are figures consisting of either concrete or simple geometric pictures to reduce the verbal load of the measurement (Authors 2012). The Cronbach’s alpha values for the pre-, post- and delayed tests were 0.70, 0.70 and 0.79, respectively.

Mathematical skills were evaluated by using the assessment of mathematical skills in the first grade (AMS-1) (Authors 2013). It is a group-based, paper-and-pencil test that measures relational skills, counting skills, word problem skills and single-digit addition and subtraction fluency. The Cronbach’s alpha values for the pre-, post- and delayed tests were 0.91, 0.92 and 0.92, respectively, excluding the addition and subtraction tasks.

The following verbal measures were used for the pre- and post-tests but not for the delayed test – to ease the measurement load of the teachers. The listening comprehension skills were assessed by using the assessment of reading and listening comprehension and reading fluency measure (Kajamies et al. 2003). The test consisted of two short stories (non-fiction and fiction) with six and seven verbal claims, respectively. After listening to a story, the student was asked to choose whether the claims were true or false by crossing a happy or an unhappy face on the answer sheet. The Cronbach’s alpha values were 0.71 and 0.53, respectively. The reading fluency was measured by a standardised reading test for primary grades (Lindeman 2005). This five-minute test comprised 80 items, each including a picture and four different words. The child read the choices and chose the word corresponding to each picture. The Kuder–Richardson reliability of the test was 0.85, based on the study used for norming the test (Lindeman 2005).
Intervention programme

As outlined in the section on the theoretical bases, the ITS-1 was planned according to the principles of two well-known thinking skills programmes, namely, the CA (Adey and Shayer 1994) and Klauer’s (1989) Cognitive Training for Children I. The intervention consisted of a series of activities designed to provide age-appropriate problems, accompanied by a sturdy lesson structure. The lessons were delivered in a way that provided opportunities for social construction and metacognition. The total programme for this experiment included 12 activities, as well as one introductory lesson to orient the children about the working methods used in the activities.

Activities

Each activity was related to one of the concrete operation schemata described by Adhami and Yates (2008). The list of the activities based on the related schemata are as follows: Activity (1) Collecting animals (rules of game & classification), (2) Let’s reason 1 (classification & seriation), (3) Finding the right order (seriation), (4) Let’s classify (classification), (5) Treasure Island (time sequence), (6) City 1 (points of view), (7) Forest animals (classification), (8) Train trip (time sequence), (9) Envelopes (causality), (10) City 2 (points of view), (11) Lights and shades (causality) and 12) Let’s reason 2 (classification & seriation). In practice, the teachers taught from one to two lessons per week over a eight-week period, and each lesson activity followed the same phase order – orientation, problem and reflection. For example, the first activity (Collecting animals) was executed in the following way. The teacher and the children sat around a table. First, the teacher oriented the children about the lesson's theme by asking what collecting meant. Next, the teacher put a map (game board) on the table, and they discussed about the map. Then the teacher spread the cards (pictures of various animals with different colours) on the table. After all the children became familiar with the animals and related concepts, the teacher introduced the game rules, which followed the rules of the famous UNO® card game. After the game, the children reflected on what was difficult and how they overcame such challenges. Additionally, each session was enriched with problem-related, inductive reasoning tasks. For example, the classification lessons (activities 1, 2, 4, 7 and 12) included generalisation tasks, such as class formation, class expansion and finding common attributes (Klauer 1989), while the points-of-view lessons (activities 6 and 10) were enhanced with mental rotation tasks (Newcombe and Frick 2010). This lesson structure was believed to respond to Fuchs and Fuchs (2008) arguments that problem-based learning would not be enough for the children with weaknesses in learning, but the teachers need to reveal the lesson's goal, as well as guide and deliver scaffolded instruction to the children through the experiments. This process is in turn expected to strengthen their deductive reasoning.

Procedures

This study applied a research design that encompassed the control of the pre-test, intervention, post-test and delayed post-test phases. We measured the thinking and mathematical skills three times: just before the intervention (in September 2013), immediately after the intervention (in December) and three months after the intervention (in March). The language skills were measured twice: just before and immediately after
the intervention, due to the limited differentiating capacities of these measurements. The assessment sessions were held during the school day and took approximately three 45-min sessions on separate days. The intervention sessions were conducted in small groups of four and five children (one group for each teacher). With the classroom teachers’ assistance, the children's demographic descriptives (age, gender, language and family educational status) were collected from their parents during the pre-test period, using a questionnaire that the children brought home. The teachers participated in a six-hour training day before the intervention phase, concentrating on the study procedures and the implementation of the assessment and intervention tools. The teachers were given detailed lesson and measurement manuals with materials, which were introduced and explained on the training day. The authors and trained teachers with many years of teaching experience conducted all assessment sessions. The intervention sessions were conducted by two female special education teachers. All measures were graded by the authors.

**Statistical analysis**

To study the relationships between the children's background variables and their measured skills, the Spearman correlation test was performed. The parametric assumptions for analyses were tested with the comparisons between the standard error of the skewness range and the skewness value, the graphical observation of the distribution and the Shapiro–Wilk test. When the parametric assumptions were not met, the corresponding non-parametric analyses were performed and reported. The gain scores were used in measuring the intervention effects. The gain scores for each measure were calculated by subtracting the pre-test scores from the post-test scores, which provides a reliable measure when no randomised controlled trials are available (Oakes and Feldman 2001). To examine the differences among the groups of study, the ANOVA test with the post hoc Tukey was applied. The effect sizes (ES) of these comparisons were calculated by using Hedges’ $g$, due to the small study groups (Turner and Bernard 2006). The gain ES were calculated by subtracting the pre-test ES from the post-test ES. An ES value of 0.30 can be interpreted as a small effect, 0.50 as a medium effect and 0.80 as a large effect.

**Results**

**Pre-test time**

**Whole sample**

To study the possible interaction effects of the gender and the grouping, the t-test for the gender was performed. The boys and girls performed equally well on all the pre-test measures (all $p$-values $> 0.05$). When examining how the age, family SES and pre-test measures correlated to each other, the following relationships were found (Table 1).

Table 1 shows that a child's age had a statistically significant, positive correlation ($p < .05$) to both the thinking skills and reading fluency measures. The mother's and the father's educational levels had a statistically significant correlation to each other ($p < 0.001$) and to the mathematical skills measure ($p < 0.05$). The thinking skills measure had a statistically significant correlation ($p < 0.001$) to all the other pre-test measures. The highest correlation
(r = 0.61, p < 0.001) emerged between the mathematical skills measure and the reading fluency measure.

**Subgroups**
There were no statistically significant differences among the Lowl, LowC and WellC groups, neither in terms of the children's ages nor on their mother's and father's educational levels. Table 2 shows the descriptives of the study for each group, based on the measurement points (pre-, post- and delayed tests).

**Group comparisons in pre-test**
A one-way ANOVA was used to examine whether the children's scores on the pre-test were a function of their grouping. The one-way ANOVA of the test scores at the pre-test phase (see Table 2) revealed statistically significant main effects on the thinking skills measure, $F(2, 148)=82.44$, $p < 0.001$; mathematical tasks, $F(2, 141)=4.01$, $p < 0.05$, listening
comprehension tasks, $F(2, 142) = 11.95$, $p < 0.001$; and reading fluency tasks, $F(2, 141) = 4.54$, $p < 0.05$.

Post hoc comparisons using the Tukey procedures were used to determine which pairs of the three groups' means differed. These results (Table 3) indicated that the children in the WellC group had statistically significant higher scores on all four pre-test measures compared to the children who belonged to either of the low-achieving groups except mathematics measure in where WellC had not statistically significantly higher scores than the LowC group. There were no statistically significant differences between the low-achieving groups. It is noteworthy that the largest variations occurred in the LowI group in every pre-test measure except mathematics.

### Group comparisons in post-test

The one-way ANOVA among the three groups on their test scores at the post-test phase (see the descriptives in Table 2) revealed significant main effects on the thinking measure, $F(2, 146) = 19.37$, $p < 0.001$; listening comprehension tasks, $F(2, 142) = 6.02$, $p < 0.01$; and reading fluency tasks, $F(2, 142) = 4.54$, $p < 0.05$. Post hoc comparisons using the Tukey procedures were used to determine which pairs of the three groups' means differed. These results indicated that all the above-mentioned statistically significant differences were between the LowC and the WellC groups. The LowI group no longer differed from the WellC or the LowC group on any measure.

### Intervention effect in post-test

A one-way ANOVA was used to examine whether the children's gain scores on the post-test and the delayed post-test were a function of the intervention. The independent variable represented the three subgroups. The dependent variables were the children' gain scores on the thinking, mathematics, listening comprehension and reading fluency tasks. For the listening comprehension tasks, an additional Kruskall–Wallis test was performed to confirm the parametric results. The one-way ANOVA of the gain scores on the post-test measurement revealed a statistically significant main effect on the thinking measure, $F(2, 148) = 17.05$, $p < 0.001$ and the reading fluency tasks, $F(2, 141) = 3.89$, $p < 0.05$. Post hoc comparisons using the Tukey procedures were used to determine which pairs of the three groups' means differed.

The results indicated that the children in the Lowl group had a statistically significant, higher improvement on both the thinking ($p < 0.01$, gain ES 0.86) and reading fluency measures ($p < 0.05$, gain ES 1.28) than the children in the LowC group in post-test phase.
Additionally, on the thinking measure, both the LowI group ($p < 0.001$, gain ES 1.39) and the LowC group ($p < 0.01$, gain ES 0.69) demonstrated statistically significant improvements that were higher than those of the WellC group. The Kruskal–Wallis test on the listening comprehension items showed between-group differences ($p < 0.05$), and the post hoc pairwise comparisons showed that the LowI group significantly improved ($Z = -2.14, p = 0.016$, gain ES 0.75) its performance on the listening comprehension measure compared to the WellC group. There were no statistically significant gain score differences in mathematics.

**Between-group comparisons on delayed post-test**

The one-way ANOVA of the test scores on the delayed post-test phase (see the descriptives in Table 2) revealed the statistically significant main effect on thinking measures, $F(2, 146) = 21.27, p < 0.001$. Post hoc comparisons using the Tukey procedures were used to determine which pairs of the three groups’ means differed. These results indicated that the WellC group outperformed the LowC group ($p < 0.001$) on the thinking measure.

**Intervention effect on delayed post-test**

The one-way ANOVA for the gain scores on the pre-test and the delayed post-test showed the statistically significant main effect on the thinking measure, $F(2, 146) = 21.27, p < 0.001$. Post hoc comparisons using the Tukey procedures indicated that the LowI group ($p < 0.001$, gain ES 1.39) and the LowC group ($p < 0.05$, gain ES 0.69) had statistically significant, higher gain scores on the thinking measure than that of the WellC group. Focusing on the gain ES changes between the LowI and the LowC groups on the thinking measure would show that between the pre-test and the delayed post-test, an average LowI child improved about three quarters of the standard deviation (gain ES 0.71) more than an average LowC child; accordingly, an average LowI child improved almost one and a half standard deviation (gain ES 1.39) more than an average WellC child. The magnitude of the LowI group’s gain on the thinking measure remained at the corresponding levels in both the post-test and the delayed post-test.

**Discussion**

This study aimed to examine the effects of the ITS-1 intervention programme on the low-achieving first graders. The intervention programme showed positive effects on the low achievers’ performance at both post-test and delayed post-test measurement points. The statistically significant differences in the thinking skills, mathematical skills, listening comprehension and reading fluency that were found between the LowI and WellC groups at the pre-test point disappeared at the post-test point, and the differences in the thinking skills and mathematical skills that were found at the pre-test phase did not reappear on the delayed post-test. In contrast, such differences remained between the LowC and the WellC groups during the entire study period except mathematical skills in where there were no differences between these two groups in pre-test phase. Focusing on the gain scores and the gain ES of the low-achieving groups would show that the LowI group's performance improved more in thinking skills and reading fluency from the pre-test to the post-test phases compared to the LowC group. Beyond the improved scores of the LowI group alone, the programme had positive effects on the low-achieving children's thinking skills. Earlier studies had shown that thinking skills, including reasoning, spatial ability and problem-solving, constituted one of...
the most central factors that explained and predicted academic achievement through the elementary years (Kuhn and Holling 2009; Hu et al. 2012). The contents of various school subjects are constructed on such categories of concepts; therefore, when children attain higher levels of identifying categories, they are simultaneously primed for knowledge building in formal schooling. To express it in Piagetian language, children who are lifted to the required cognitive level with the help of direct and scaffolded instruction are able to assimilate the content and the topic on this level (Adey 2008; Morris et al. 2012). Furthermore, the first graders, who just entered the school, may face the challenge of adapting to the complexity of formal schooling; they may have not had time to develop their learning habits, nor even had learned how to behave and study at school not to speaking their test-taking skills.

Consequently, intervention (targeted small group time with structured programme) offered by a trained (special education) teacher may have a positive impact not only on children’s thinking skills but also their learning habits and motivation. This speculation is supported by the remarkable improvements in reading fluency in the LowI group compared with much limited gains by the LowC that at first glance does not seem to be related to their advances in thinking skills. However, another explanation for the favourable mediation effect of improved thinking skills and the early reading skills comes from the longitudinal studies, which have shown that levels of fluid intelligence accompanied with phonological awareness are related to the development of early reading and listening comprehension skills (Evans et al. 2001; de Jong and van der Leij 1999; see also Spencer et al. 2013). Although this conclusion is really tentative, owing to non-existing phonological awareness measures, it raises interesting question related to promotion of the early reading skills in conjunction with thinking skills. Interestingly enough, there were no such gains for the LowI over the LowC in performance of mathematical tasks in post-test. We assume that due to fact that for the first grade aims of curriculum for mathematics (FNBE 2004) contains learning of small numbers and simple calculations (e.g. number symbols, properties of numbers, addition and subtraction) in all simple tasks, we were not able to trace improvements with this regard.

To deepen the discussion, thinking skill improvements computed by gain scores were observed among the children participating in the intervention – and not in the same degree as in the rest of the children. During the intervention programme, the children were taught how to compare and contrast specific items as a general problem-solving strategy. However, true learning competence increase can be detected by improvements not only in thinking skills but also in different ways, such as learning subjects with retention (Hasselhorn 1995). This was not fully supported due to relatively small gains of the LowI group in mathematics. Nevertheless, present study’s results strongly support the idea that interventions in thinking skills should target the low performers in the early school years. The study supports the contention that preventing learning difficulties can be achieved, not only by reading, writing and mathematics interventions but via thinking skills intervention as well. The data show that the ITS-1 benefits low-achieving children before they fail. Slavin’s (1996) term for this approach is ‘neverstreaming’, in the sense that it helps those children who risk failing to achieve success and maintain it through grade one and onwards.

Most children in the LowI group improved significantly although some did not respond to the intervention. This ‘resistant to intervention’ (i.e. non-respondents) phenomenon is well known in intervention studies (cf. Adey and Shayer 1994; Toll and Van Luit 2012; Fuchs, Fuchs, and Compton 2013). Many possible reasons explain this effect, starting from the intervention programme itself. Perhaps some elements of the intervention programme, such as
the concepts used, the teaching methods, duration and intensity, do not match the responsiveness of a child. We do not know if the low-achieving children present a specific reading (dyslexia) or arithmetic (dyscalculia) problem. However, if they did, it would be necessary to take into account their specific difficulties (phonological loop, phoneme–grapheme correspondence, working memory, etc.) and to consider other, for example, cognitive-focused approaches to tackle these issues (see e.g. Kavale and Flanagan 2007; Fuchs and Fuchs 2008).

In future research, proper amounts of information, accompanied with dynamic assessment techniques, can improve the understanding of the antecedents of low-level performance and provide detailed knowledge about the learner’s characteristics and their effects on the responsiveness of the intervention (Phye and Johnson 2009).

This study has limitations that are worth noting for replicability purposes. First, the size and the formation of the study’s sample raise a potential question. The studies with smaller sample sizes have been shown to produce remarkably higher ES than those of the more representative studies ($n > 250$) (Marzano 1998; Slavin 2008). It is demanding and rare to be able to reach the proposed number of randomly chosen participants among the low-achieving children (Higgins et al. 2005). Another possible problem arises from the fact that only two special education teachers volunteered to implement this intervention. They might have been more motivated, skilful and research oriented than the other teachers. Cronbach et al. (1980) called this phenomenon a ‘superrealization bias’, meaning that the study on hand is not replicable in reality. Finally, the lack of delayed measures in reading fluency and listening comprehension could weaken the evaluation of the transfer effect of the intervention.

**Conclusion**

Our intervention programme for low-achieving first graders leads to the improvement of their thinking skills and academic achievement. As mentioned, the differences found between the intervention group and the well-performing control group in the beginning of the study disappeared at the post-test phase and did not reappear in the delayed measure to the same extent. However, such differences between the LowC and the WellC groups, respectively, remained detectable. In the future, we will develop and test similar programmes for the kindergarten and second grade pupils, which in total will provide a 36-h training programme for low-achieving school starters. We hope that when we develop this programme, test its suitability for six- to eight-year-old children with varying ability levels and make the materials and methods more generally accessible, we can make additional contributions in practice and in theory in the field of thinking-oriented education.

**Acknowledgment**

This research was part of the ThinkMath project, which was funded by the Finnish Ministry of Education and Culture (2011–2015). The opinions expressed are those of the authors and do not represent views of the Finnish Ministry of Education and Culture. We wish to thank all the participating teachers, children and their parents.

**Disclosure statement**

No potential conflict of interest was reported by the authors.
Notes on contributors

Risto Hotulainen works as an associate professor (Educational Assessment) at the University of Helsinki, Finland. He has worked earlier in the field of special education as an university lecturer and currently he works as a director of the Centre for Educational Assessment at the University of Helsinki, Finland. He is a specialist in the development of attention, self-concept, thinking skills and learning to learn skills and related interventions.

Riikka Mononen works as an associate professor (Mathematical learning difficulties) at the University of Oslo, Norway. Her research focuses on the mathematical skills development, mathematical learning difficulties, and interventions.

Pirjo Aunio works as a professor (Special Education) in University of Helsinki. She is also a visiting professor (Mathematical learning difficulties) in University of Oslo (Norway) and a visiting professor (Early Childhood Education) in University of Johannesburg (South Africa). Her research interests are the development and learning of mathematical skills, learning difficulties in mathematics, assessment of mathematical performance and development, and mathematical and related cognitive skills interventions.

References


