Method

Participants

A letter was sent to all primary schools in the municipality of The Hague (the Netherlands) to recruit participants. This letter explained the design of the pull-out program and asked schools to nominate fifth- or sixth-grade gifted children to participate. Criteria for nomination were above-average intelligence, as indicated by the national pupil monitoring system, motivation to participate in extra activities, and absence of clinical diagnoses of behavioral or emotional problems. Fourteen schools responded by nominating a number of children for the program. For organizational reasons, only 40 of 57 nominated children could be involved in the program. Three children were excluded based on information gathered in conversation with parents and teachers. Of the other 54 nominated children, 40 children were selected by lot to participate in the enrichment program.

Table 1

Descriptive Statistics of the Participants

<table>
<thead>
<tr>
<th>Condition</th>
<th>Experimental</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td>$M$ (SD)</td>
<td>Range</td>
</tr>
<tr>
<td>Age</td>
<td>9 – 12</td>
<td>11.32 (0.77)</td>
</tr>
<tr>
<td>General intelligence scores</td>
<td>35 – 119</td>
<td>74.69 (18.67)</td>
</tr>
<tr>
<td>SES</td>
<td>3.5 – 4.0</td>
<td>3.94 (0.17)</td>
</tr>
</tbody>
</table>

Note. SES = socioeconomic status.

The program had two components, the implementation and the intervention. During the first half of the year, which was the implementation period, the teachers set up the program while children were already placed in the new learning environment in which the teacher acknowledged their abilities. After a couple of weeks, some children left the enrichment program and new children entered the program. During this implementation period, teachers were still adapting their methods of teaching to the needs of the children. In February, two stable groups of 20 children were formed, and teachers adapted their teaching to these groups prior to the start of the intervention.

Also prior to the start of the intervention period, primary schoolteachers of the participating children were asked to nominate nonselected classmates with similar general intelligence levels to take part in the study as a control group. In addition, the 14 pupils who were nominated but excluded from participation were included in the control group. Because
control group children were recruited from the same schools, teacher characteristics and type of education were constant across groups.

In total, 37 fifth- and sixth-grade children agreed to participate in the control group. During the school year, two children withdrew from participation in the pull-out program before all data were collected due to personal circumstances. In addition, six experimental group and three control group children were not present on one measurement occasion. These children were not included in the analyses; hence, the participants comprised 32 children (20 boys) in the experimental and 34 children (17 boys) in the control group.

Descriptive statistics for the control variables (i.e., intelligence level, socioeconomic status [SES], and age) are presented in Table 1. Pearson’s chi-square tests and independent-sample t tests revealed that participants in the two groups did not differ in intelligence level, as measured with the Dutch Intelligence Structure Test (IST; Liepmann, Beauducel, Brocke, & Amthauer, 2010), \( t(64) = 0.62, p = .539, d = .15 \), or SES, \( \chi^2(2, n = 66) = 2.07, p = .355 \), Cramér’s \( V = .18 \). The two groups were also equal in terms of gender, \( \chi^2(1, n = 66) = 1.05, p = .307 \), Cramér’s \( V = .13 \) and grade level, \( \chi^2(1, n = 66) = 1.99, p = .655 \), Cramér’s \( V = .06 \). Yet, the analyses revealed that the two groups did differ in age. Children in the control group were on average 5 months older compared with children in the experimental group, \( t(52.97) = 2.39, p = .020, d = .61 \). In the subsequent analyses designed to compare the experimental and control group data, age was therefore included as covariate.

**Procedure**

At the start of the implementation period, the researcher visited the classroom to administer the first tests (M0). Tests were administered over 2 weeks. Children were first given instructions on how to complete the test. They were assured that all answers would be treated confidentially and the tests would remain anonymous.

The second round of measurements (M1) was administered in February at the start of the intervention period, and the final measurement (M2) was administered in June at the end of the intervention period. These measurement occasions followed the procedure similar to the one used in the first measurement occasion.

To get an insight into the effects of the triarchic enrichment program, children in the control group completed all questionnaires at the start (M1) and at the end of the intervention
period (M2). These children received the same instructions on how to complete the test as well as the same information on confidentiality and anonymity. All control group measurements were completed in one morning. To get an insight into intelligence levels, children in both groups completed the IST in April.

**Intervention**

Children spent all of their time in the regular classroom, except for the morning of the week during which they participated in the pull-out program provided by a secondary school. The program comprised three successive 1-hr classes each week: robotics, mathematics, and research and design. Three qualified secondary schoolteachers with experience in teaching gifted children taught the successive classes.

In the research and design classes during the implementation period, children completed adapted versions of lessons developed for the Thinker-Tinker-Lab of the Center for the Study of Giftedness in Nijmegen (Schrover, 2010). The lessons were tested extensively with 8- to 12-year-olds. They aimed to provide children with a better insight into their thought processes. The assignments asked children to construct multiple solutions to a problem and adapt them repeatedly to obtain the best result. An example of such an assignment is “Build a boat out of aluminum foil that can hold as many marbles as possible.” All lessons were based on Sternberg’s (1985) theory of triarchic intelligence. This means that assignments were designed to enhance analytical, creative, and practical abilities. Moreover, teachers adapted their teaching to the triarchic principles. To support the analytical abilities, children were encouraged to think critically and judge and evaluate the solution to a problem (Sternberg & Grigorenko, 2004). For creative thinking, they role-modeled creativity in their own behavior and supported and encouraged creativity when displayed by pupils. In addition, all teaching addressed the practical needs of the children.

During the intervention period, the attention of the research and design classes shifted from stimulating triarchic abilities to stimulating higher order skills, such as reasoning and synthesizing knowledge. Longitudinal projects, like making a mind map or designing a city, were used as assignments that supported children’s analytical, creative, and practical abilities only implicitly.
In the robotic classes, children had to build and program robots using LEGO MINDSTORMS NXT Software. With this software, children could build a robot by themselves or by following a description provided with the software. During the implementation period, children were challenged to build and program their own model in a group, without the help from a manual. In these lessons, the teacher again supported triarchic abilities. During the intervention period, children built robots following standard lessons and detailed instructional models.

The mathematics lessons comprised three areas of mathematics: geometry, algebra, and arithmetic. The lessons were derived from first-grade secondary school mathematics, and they provided only a brief introduction into the important aspects of the three mathematical areas. Lessons and assignments were not predefined but were adapted to the questions and interests of the pupils. In addition, mathematic and logic puzzles were provided. Both the implementation and intervention period addressed metacognitive skills, such as logical reasoning and integration of knowledge.

Materials

**SES.** Mean educational level of both parents was used as an indicator of SES. Children indicated the highest level of their parents’ education (low = primary education or lower secondary education; intermediate = upper secondary education or vocational training; high = college or university).

**Intelligence.** The Dutch IST (Liepmann et al., 2010) was used to measure verbal, numerical, and figural intelligence levels. Each area of intelligence was assessed using three subtests comprising 20 multiple-choice items. With the first verbal intelligence subtest, Sentence Completion, children had to indicate which of five options could complete a sentence best. Second, in Verbal Analogies, three words were presented, and children needed to identify the correct fourth word out of five alternatives. With Similarities, participants were asked to select two words belonging to the same category out of a list of six words. The second set of subtests included Numerical Calculations, Number Series, and Numerical Signs. Numerical Calculations involved mathematical problems with realistic numbers. With Number Series, pupils were provided with a series of numbers and were asked to choose from five alternatives the correct number following this series. In the Numerical Signs items, an equation using rational numbers
is presented from which the mathematical operators have been omitted. Pupils had to choose from the four basic mathematical operators (add, subtract, divide, multiply) which was correct. For figural intelligence, the first subtest, Figure Selection, involved pupils indicating which 5 out of 10 whole shapes could be produced by fitting together presented geometrical pieces of shapes. Cubes presented children with three out of six possible patterns of cubes. Answering entailed designating the cube corresponding to the initial cube out of five answering possibilities. Finally, Matrices provided figures arranged according to a particular rule for which children had to point out the one figure that conformed to this rule. The reliability measures for all subtests ranged from .63 to .90 in our sample of participants. Together, scores on these nine subtests were combined to form a general intelligence score ($\alpha = .95$). Raw scores may be converted into standardized scores with a mean of 100 and a standard deviation of 15. Yet, because norm scores are only available for children aged 12 and older, the present study used raw scores.

**Triarchic abilities.** A Dutch translation of the original American Aurora Assessment Battery (Tan et al., 2009) was used to measure triarchic abilities. This battery is based on the theory of triarchic intelligence (Sternberg, 1985) and is being further developed in multiple countries around the world (e.g., Saudi Arabia, Greece, Portugal). The paper-and-pencil test of the Aurora Assessment Battery comprises 18 subtests measuring analytical, creative, and practical abilities. In the Netherlands, only a few subtests were properly translated with equivalencies with respect to meaning, psychometric properties, and item difficulty, so that only one subtest was used per type of ability. Based on a pilot study, the Boats subtest assessed analytical abilities, Book Covers assessed creative abilities, and Money assessed practical abilities. In the present study, internal consistency was found to be acceptable to good for all scales, with $\alpha = .82$ for analytical, $\alpha = .89$ for creative abilities, and $\alpha = .76$ for practical.

On the analytical subtest, children saw 10 items comprising a photograph of toy boats that were connected to each other with a cord. The children were told that the boats could float around on the water, but would stay connected in the same way. Subsequently, four photographs portrayed the toy boats after floating. Only three of these were correct. For all 10 items, the children had to indicate which of the four was false. Every right answer was worth one point, allowing participants to earn 10 points overall.

The Book Covers subtest measured creative abilities and consisted of five colored drawings of possible book covers. The children were asked about the possible content of the
book. In the instructions, they were told that they should use their imagination and that all answers were correct. Answers were coded according to two criteria: accuracy and ability-creativity. For accuracy, a score from 0 to 2 could be obtained, dependent on the degree to which the child carried out the given task appropriately. With the variable ability-creativity, it was coded whether the child created an original substantial story to accompany the picture. The scores for ability-creativity ranged from 0, representing nonsense, to 4 when two or more interpretive or elaborating elements were described in a sequence to create a novel story. An independent rater coded 19% of the answers so that reliability measures could be calculated. Creativity codings of the two raters correlated .848 ($n = 31, p \leq .001$).

Practical abilities were assessed using the $Money$ subtest. Children responded to five items asking about a situation involving money. A sample item included, “Oliver’s aunt is taking him to the zoo. A ticket for Oliver costs 5 euro; a ticket for his aunt costs 8.75 euro. Oliver will pay half of his ticket; his aunt will pay the other half. How much should every one of them pay?” All items include payment of debts between the persons involved and payment to a third person. In total, children had to write down the expenses of 13 persons, so that a total of 13 points could be earned.

**Socio-emotional factors.** The Dutch School Attitude Questionnaire (Vorst, Smits, Oort, Stouthard, & David, 2008) measured *Motivation* for schoolwork, *Self-concept*, and *Wellbeing*. For 80 items, children had to indicate on a 3-point scale the degree to which they agreed with the statement ($1 = do not agree; 2 = no opinion; 3 = agree$).

**Attitude toward science.** Attitude toward science was measured using a scale developed by Denessen and colleagues (2011). The test comprised 20 items equally divided over four subscales: *Importance of science to society*, *Difficulty of learning science*, *Enjoyment of learning science*, and *Aspirations to pursue a future career in science*. Children had to indicate on a 4-point scale the extent to which they agreed with statements provided ($1 = strongly disagree; 4 = strongly agree$).

**Data Analysis**

Children in the experimental group completed all questionnaires and tests on three measurement occasions (M0, M1, and M2), whereas children in the control group only completed the measures in the intervention period (M1 and M2). Because the number of measurements differed
for both groups, the data of children in the experimental group were first analyzed separately from the data of children in the control group. Second, the effects of the intervention were examined by including data for both groups in a repeated-measures ANOVA. Due to the large number of tests, alpha was set to .01.