

Educational Psychology

An International Journal of Experimental Educational Psychology

ISSN: 0144-3410 (Print) 1469-5820 (Online) Journal homepage: <http://www.tandfonline.com/loi/cedp20>

Progress in the inductive strategy-use of children from different ethnic backgrounds: a study employing dynamic testing

Wilma C. M. Resing, Kirsten W. J. Touw, Jochanan Veerbeek & Julian G. Elliott

To cite this article: Wilma C. M. Resing, Kirsten W. J. Touw, Jochanan Veerbeek & Julian G. Elliott (2017) Progress in the inductive strategy-use of children from different ethnic backgrounds: a study employing dynamic testing, *Educational Psychology*, 37:2, 173-191, DOI: [10.1080/01443410.2016.1164300](https://doi.org/10.1080/01443410.2016.1164300)

To link to this article: <https://doi.org/10.1080/01443410.2016.1164300>



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



Published online: 06 Apr 2016.



Submit your article to this journal [↗](#)



Article views: 572



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 3 View citing articles [↗](#)



Progress in the inductive strategy-use of children from different ethnic backgrounds: a study employing dynamic testing

Wilma C. M. Resing^a, Kirsten W. J. Touw^a, Jochanan Veerbeek^a and Julian G. Elliott^b

^aFaculty of Social Sciences, Department of Psychology, Section Developmental and Educational Psychology, Leiden University, Leiden, The Netherlands; ^bSchool of Education, Durham University, Durham, UK

ABSTRACT

This study investigated potential differences in inductive behavioural and verbal strategy-use between children (aged 6–8 years) from indigenous and non-indigenous backgrounds. This was effected by the use of an electronic device that could present a series of tasks, offer scaffolded assistance and record children's responses. Children from non-indigenous ethnic backgrounds, starting at a lower level, profited as much from dynamic testing as did indigenous children but were unable to progress to the standard of this latter group. Irrespective of ethnic group, dynamic testing resulted in greater accuracy, fewer corrections, and reduced trial-and-error behaviour than repeated practice. Improvements in strategy-use were noted at both the group and individual level. After dynamic training, children from both ethnic groups showed a superior capacity for inductive reasoning although indigenous children subsequently used more inductive strategies. The study revealed individual differences between and within different ethnic groups and variability in the sorts of help required and subsequent strategy progression paths.

ARTICLE HISTORY

Received 13 August 2015
Accepted 7 March 2016

KEYWORDS

Dynamic testing; ethnic groups; series completion; inductive reasoning; (meta) cognitive training

Introduction

For over half a century, a variety of dynamic assessment and test procedures has been developed and evaluated (e.g. Haywood & Lidz, 2007; Lidz & Elliott, 2000; Sternberg & Grigorenko, 2002; Tzuriel, 2013). Whereas conventional, static test procedures are characterised by testing without the provision of any form of feedback, dynamic testing (or assessment) is based on the assumption that test outcomes following some form of (scaffolded) feedback or intervention are more likely to provide a better indication of a child's level of cognitive functioning than conventional, static test scores alone. The primary aims of research in dynamic testing have been to examine improvements in cognitive abilities following training between test session(s), to consider behaviour related to the individual's potential for learning, and to gain insight into learning processes at the moment they occur. Dynamic tests differ from static tests because in a dynamic test situation, testees are given feedback or guided instruction enabling them to show individual differences in progress when solving equivalent tasks. Some proponents of dynamic testing go further and argue that such information could potentially guide or inform recommendations about appropriate intervention in the classroom (Grigorenko, 2009; Jeltova et al., 2011). However, attempts to measure learning processes while children are being trained and

tested in detail can be overwhelming. For this reason, the present study sought to gather information during the assessment sessions about individual learning processes (how children respond to guided instruction and scaffolding, and, in the light of this, how their strategy-use progresses towards a more advanced way of task solving).

Findings about children's cognitive abilities cannot be fully understood without considering their cultural or ethnic context (Carlson & Wiedl, 2013; Sternberg, 2004). Children from non-indigenous ethnic backgrounds tend to score lower on cognitive tests than their peers from the indigenous, dominant culture (Carlson & Wiedl, 2013; Stemler et al., 2009; Tharp, 1989). In the Netherlands, the educational performance of non-Western immigrant children has been reported to be generally poorer than that of those of Dutch heritage with a smaller proportion entering higher education (Central Bureau for Statistics [CBS], 2007). The reasons for this are likely to be many, with a lower level of proficiency in the majority languages and sociocultural factors being key factors (Backus, 2004). In learning their first language, children develop an understanding of their world that is underpinned by a system of words and meanings, concepts and symbols. These can be very different for children from other cultures who then also often have to struggle with the challenges of a second language (Bialystok, 2001). Such difficulties are also often found on performance on the so-called culture reduced tests of cognitive ability (e.g. Cattell, 1979), where one can often witness a performance gap between dominant and minority culture students (e.g. Brown & Day, 2006). Brouwers, Van de Vijver, and Van Hemert (2009), for example, found in their 45 country meta-analysis of cultural differences in Raven's Progressive Matrices scores (supposedly a relatively culture-free measure), that cognitive variability between cultural groups can best be thought in terms of differences and variability in the ways that individuals approach and solve problems (e.g. Siegler, 1994, 2007) rather than in terms of stable differences.

A second difficulty for non-Western children concerns a possible deficit in their relative experience of testing, a phenomenon sometimes known as 'test-wiseness' (Williams, 1983). This refers to the ability of the participant to use an understanding of the test or item format to receive a higher score (Millman, Bishop, & Ebel, 1965). In addition to the unfamiliar nature of the measures employed, the formal interviewer–interviewee relationship, in which highly standardised modes of interaction are prescribed, can prove unsettling to children from other cultures and lead to poorer performance.

The seminal work of Feuerstein, Rand, and Hoffman (1979) highlighted the tendency of many immigrant children to underperform on traditional cognitive tests. To remedy this situation, they argued that the testing format should be refashioned in such a way that the true potential of children should be revealed. Their approach, dynamic assessment, drawing upon Vygotskian theory, emphasised assisting and guiding the child in the test situation, within their zone of proximal development. Such training can help to reduce the influence of language and culture on the child's performance, for example, by compensating for differences in factors such as test-wiseness, learning opportunities or non-native instruction language (e.g. Bridgeman & Buttram, 1975; Serpell, 2000; Van De Vijver, 2008). Guidance contingent on the child's performance on the test can help the student to gain information about the type of performance that is valued on the test (Sternberg et al., 2002).

It is important to emphasise that while such approaches may be particularly valuable for children from non-dominant cultures, or who experience social disadvantage, the proponents of dynamic testing argue that the approach is helpful for the examination of all children. Particularly helpful for all test users is the opportunity that is provided by the approach to observe the nature, rate and extent of the child's improved performance when assistance is provided (Hessels, 2000; Sternberg & Grigorenko, 2002, 2004). However, whether dynamic approaches involving just a few tester–testee sessions offer additional value for assessing the cognitive potential of children from ethnic minorities is a moot question that was key to the present study. Thus, even if there were initial gains in performance, these might not be sustained in situations where only minimal intervention has taken place. Our expectation was that, for significant differences to emerge between the groups, more intensive training would be necessary. However, educational (school) psychologists rarely have the opportunity to undertake very lengthy, time-consuming assessments of children with learning difficulties and a speedy yet productive dynamic approach would seem to be highly desirable.

Our dynamic-testing approach, utilising a pretest-training-post-test format with guided instruction and observation of learning during testing, draws upon the use of graduated prompts which are provided to the testee whenever they encounter difficulties in solving the tasks (Campione & Brown, 1987; Fabio, 2005; Resing, 2000; Resing & Elliott, 2011). Such help should be restricted to the minimum number of prompts and scaffolds necessary to effect progression on the presented training task. Changes in the number and quality of prompts needed during training, and in strategy-use when solving the tasks, can be considered to be indices of a child's potential for learning.

The graduated prompts approach has often employed inductive reasoning tasks and training procedures. These tasks, for example, analogical reasoning, categorisation, seriation, and all require rule finding processes that can be achieved by searching for similarities and differences between the objects, or in the relations between the objects under examination (Goswami, 1996; Klauer & Phye, 2008; Sternberg, 1985). Changes in the use of cognitive strategies after training or repeated testing have been found in inductive reasoning studies using class-inclusion tasks (Siegler & Svetina, 2006), and matrices/analogies (Alexander, Willson, White, & Fuqua, 1987; Siegler & Svetina, 2002; Tunteler, Pronk, & Resing, 2008). In contrast, dynamic testing research using series completion tasks is sparse (e.g. Ferrara, Brown, & Campione, 1986; Holzman, Pellegrino, & Glaser, 1983; Sternberg & Gardner, 1983) and has mostly focused on the detection of task components underpinning adult series completion (e.g. Simon & Kotovsky, 1963).

The present study examined whether children from non-indigenous backgrounds would show different forms of progression in solving patterns and in strategy-pathways than indigenous children when presented with an adapted version of the schematic-series completion task (Resing & Elliott, 2011) within a dynamic testing context. The task presented was based on a process model of series completion in which children were helped, as necessary, to complete several series of schematic-puppets. In an earlier publication, Resing, Xenidou-Dervou, Steijn, and Elliott (2012) demonstrated children's progression in verbal and behavioural strategies after dynamic testing. The present paper reports a different aspect of the same study, here focusing upon the extent to which children from non-indigenous backgrounds differed from indigenous children in progression and strategy-use.

For the original schematic-series completion task (Resing & Elliott, 2011), we used Simon and Kotovsky's (1963) letter series completion model to construct the item pool. Our first studies, however, had shown that the theoretically predicted and empirically found item difficulties did not correlate highly. It was concluded that the pictorial schematic-series completion task we were employing requires a more complex solution procedure than for letter or number series. Indeed, even solving letter and number series with the same underlying 'rules' seems to require different processes that cannot be captured within one and the same task analytical model (Quereshi & Seitz, 1993). Solving schematic-series completion tasks requires a more complex procedure than letter or number tasks because the letters of the alphabet have a fixed relation to each other, as do numbers that have other relational possibilities, whereas the various elements of schematic pictorial series do not have a known *a priori* relationship to each other. For every new task-item, the child must search for yet unknown strings of regularly repeating elements, in combination with unknown changes in the relationship between these elements, a process that does not have to be linear. In contrast to number and letters, our 'elements (puppets)' of the series are not integral single objects but instead have to be constructed from a number of blocks. An increasing number of changing aspects within each puppet [series position] renders the series more complex. Within one row, a variety of parallel periodicities and several transformations of elements have to be identified. Distractors can also have an impact on the difficulty level of items, particularly for young children (e.g. Richland, Chan, Morrison, & Au, 2010). For the present study, it was decided to construct a new series, based on 'puppets' with a greater number of characteristics, and a refined solution model based on a hierarchy of both the number of changes and the periodicity (see 'Method'). This was the starting point for defining and detecting variation in behavioural solution strategies.

A well-known problem of dynamic testing, especially in one-to-one assessment situations, is that the procedures involved are time-consuming for educational (school) psychologists and teachers and have yet to demonstrate utility for informing intervention. New and attractive educational electronics that

are rapidly evolving (Resing & Elliott, 2011) offer the potential to shed light on the learning processes of individual children in real time.

In the present study, children were tested with new, revised, and brightly coloured, transparent 3D-electronic tangibles. Children 'played' with the various pieces and were encouraged to place them freely on an electronic console (see Verhaegh, Fontijn, Aarts, & Resing, 2013). The tangible interfaces were combined with speech-technology and some visual support (white lights). The console recorded the nature and timing of the children's responses in what appeared to be a natural setting (Verhaegh, Fontijn, Aarts, & Resing, 2013). In comparison to the more typical use of a computer screen, tangible objects offer more possibilities to utilise adaptive prompt structures and scaffolding procedures, thus creating a more authentic assessment environment for the child (e.g. Revelle, Zuckerman, Druin, & Bolas, 2005).

According to Siegler (2007), learning is reflected by changes in strategy-use after a particular learning or intervention episode. Sternberg and Gardner (1983) described a shift in strategy-use between young and older children solving series completion based on schematic-picture tasks. Accuracy rates tended to increase with age and older children showed a more integrated, unitary encoding strategy. In contrast, younger children displayed a stop-and-go encoding process, and regularly employed strategies leading to an incorrect outcome. However, little research has been conducted on strategy-use and changes in strategy-patterns during dynamic testing of groups of children from different ethnic backgrounds.

In the present study, we sought to employ dynamic testing to assess the extent to which children would subsequently adopt more advanced behavioural strategies and offer superior verbal explanations of the reasoning behind their actions (verbal strategies). Although our primary research aim was to gain insight into the extent of the children's progression in strategy-use, we were also interested to examine a number of other ways by which the measurement of children's improved performance could be made visible. Here, the electronic console offered us the opportunity to capture various process measures of children's performance, such as action sequences.

We tested a number of hypotheses. Firstly, we sought to examine children's potential for learning. We anticipated that training by dynamic testing would lead to greater advances in children's series completion problem-solving behaviour than for controls, measured in terms of their (a) accuracy, (b) completion time, (c) correctly placed number of pieces and (d) number of corrections (Resing & Elliott, 2011). We also expected that the progression rates on these measures (a–d) would be similar for both groups of children who received training irrespective of their ethnic background and there would be no significant interaction effect between condition and ethnic group. Our expectations were derived from findings from recent studies using different tasks that have suggested that children from non-indigenous ethnic backgrounds, while tending to start at a lower level, often profit from training in a comparable way to indigenous children (Calero et al., 2013; Stevenson, 2012; Wiedl, Mata, Waldorf, & Calero, 2014).

Secondly, we explored the number and types of prompts that were required. We assumed that children would not show equal progression pathways as a consequence of dynamic testing. This issue was explored by splitting the dynamically tested children into different groups. We compared those who needed many, versus those requiring few, prompts during training, in combination with those who had lower, and those who had higher, post-test scores. Based on earlier research (Resing & Elliott, 2011; Resing, Tunteler, De Jong, & Bosma, 2009), we also expected that children would need fewer prompts in training 2 than in training 1, and that some of the children would only require metacognitive prompts, while others would also need cognitive prompts to be able to complete the tasks.

Third, we sought to examine the influence of training on strategy-use by looking at (a) how children explained their solutions *verbally* during testing and (b) how children's action sequences changed at the *behavioural* level. On the basis of earlier research (e.g. Resing & Elliott, 2011), it was anticipated that children from both indigenous and non-indigenous backgrounds would employ more sophisticated strategies after training compared with control-group children (Resing et al., 2009). We hypothesised that dynamic testing would enable the children to develop the quality of their explanations from non-inductive or what we termed 'partial inductive' reasoning towards more advanced inductive reasoning

strategies (Tunteler et al., 2008). However, we predicted that dynamically tested children from non-indigenous backgrounds would show less progress with respect to the quality of their verbalisations than dynamically tested indigenous children, when asked to provide an oral account of how they tackled the tasks. This hypothesis differs from our first hypothesis where we expected to find equal progression in solving behaviour. For the latter hypothesis, we anticipated that differing levels of linguistic competence might potentially prove to be a mediating variable between the two groups of children (Wiedl et al., 2014). We also anticipated that children's strategy-use measured at the behavioural level would become more advanced for both groups of dynamically tested children (indigenous and non-indigenous) than for the control-group (Resing, Xenidou-Dervou, et al., 2012). It was further predicted that both dynamically tested groups would display comparable progression in behavioural strategy-paths (Stevenson, 2012; Wiedl et al., 2014) because these were considered to be less likely to be influenced by verbal processes than their oral explanations.

Fourth, it was explored whether it would be possible to distinguish various subgroups differing on the basis of their (non-)progression in inductive, non-inductive and variable strategic behaviour and explanations.

Method

Participants

The study employed 116 children (60 boys and 56 girls; mean age 7.0 years; SD = 5.2 months) from grades 1 and 2 of five regular primary schools located in midsize and large towns in the western part of the Netherlands. Participants and schools were selected on the basis of their ethnic mix and their middle and lower socio-economic backgrounds. Sixty children were from a non-indigenous background and the remaining 56 were from the indigenous (Dutch) culture. Dutch was the primary language in the schools for all children. Parental permission to participate was obtained for each child. One child dropped out of the study during the extended period of testing, due to school absence. The verbal explanations of four children and the behavioural strategies of 14 children could not be scored because of partially missing data.

Design

The study utilised a 2 × 2 pre-test-post-test control group design with randomised blocks (see Table 1). Blocking was, per ethnic group, based on Exclusion, a visual inductive reasoning subtest from a Dutch child intelligence test, administered before the pre-test. After blocking, children were randomly allocated to the control static testing condition or the dynamic testing condition. All children were given static pre- and post-tests that were administrated without any form of feedback although they were shown how to work with the electronic console. During the sessions between pre-test and post-test, children from the dynamic testing condition were provided with graduated prompts training and scaffolds by the console whenever they failed to solve an item correctly. Children in the control condition instead were asked to solve paper-and-pencil 'mazes' during these sessions.

Table 1. Scheme of the design of the study (ct = control task).

		Exclusion task	Pre-test	Training 1	Training 2	Post-test
Ethnic minority	Dynamic testing group	X	X	X	X	X
	Control group	X	X	ct	ct	X
Indigenous	Dynamic testing group	X	X	X	X	X
	Control group	X	X	ct	ct	X

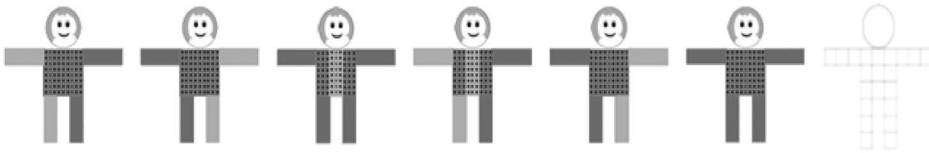


Figure 1. Item example: completed items consist of eight pieces (arms (2), legs (2), body (3) plus head).

Materials

Exclusion

Exclusion is a visual inductive reasoning subtest of a Dutch child intelligence test (RAKIT: Bleichrodt, Drenth, Zaal, & Resing, 1984). The task consists of 50 items each represented by four abstract figures. The child has to discover a rule in which three of the four figures belong together and must then subsequently exclude the other figure. The subtest requires the child to infer rules, an ability that is assumed to be important for successful inductive reasoning.

Series completion task

In order to measure inductive reasoning in children, we developed a highly structured dynamic visual-spatial 'puzzle' series completion task using tangible objects. The theoretical basis, and task construction principles for this instrument, has been described in Resing and Elliott (2011). The original task was evaluated and reprogrammed for the tangible console. An example item from this task is given in Figure 1.

Inductive reasoning tasks can be solved by detecting rules regarding commonalities and differences in the task elements, and in the relations between these elements (Klauer & Phye, 2008). Therefore, the series completion 'puzzle' task was constructed with respect to number and types of task elements, and the relationships between these elements in the series. One challenge is that there should be relatively few tangible task elements (puzzle pieces), while, at the same time, it should be possible to construct many different designs. Sternberg and Gardner's (1983) schematic-picture puppet materials served as a model for the task construction because the analogies and series tasks they developed are attractive to children and include only a limited number of elements that, nevertheless, yield a considerable number of different possible design transformations.

The series completion 'puzzles' consisted of six puppet pictures in line and an empty 'box'. Each puppet was comprised of 8 blocks (1 head, 2 × 2 legs and arms, and 3 body parts). To solve each problem, the child had to construct the next puppet in the line, by determining the nature of the systematic changes in the row of puppets and then detecting and formulating the underlying solution rule(s). This inductive rule-finding process included the following specific task features: (changes in) gender (male, female); colour (pink, yellow, green, blue); and design (stripes, dots or plain). The children could choose from 14 different types of puppet pieces (85 pieces in total) to create the correct solution. The difficulty of each series was based on the number of changes in relationships and changes of periodicity over two, three or four figures included in the series. The pre- and post-test stages of the series completion task consisted of one example item and 12 test items each, ranging from easy to difficult. The two training sessions included six items each, ranging from difficult to easy. In order to make parallel booklets for the pre- and post-test and the two training sessions, we changed various elements, for example, female puppets were changed into males, or pink items were changed into green pieces. In this way, the solving principles and the nature of the series remained identical.

Dynamic testing procedure

The activities for the dynamic testing phase were comparable to the pre- and post-test sessions but help was given by the console. Here, the children received structured prompts (see Table 2) on how to solve the series completion task, starting with general metacognitive prompts, followed by more

Table 2. Schematic structure of the graduated prompts offered by the electronic console during the dynamic training sessions.

Type of prompt	Verbal prompt	Visual prompt
1. Metacognitive	How did you solve the previous series?	
2. Metacognitive	Look at the pictures in the series What's alike and what changes? Look at girl/ boy, colour and pattern	
3. Task-specific/cognitive	How does this body part change?	Body part lights up in white
4. Task-specific/cognitive	Which body part should be placed here?	Body part lights up in white Body part lights up in right colour/pattern
5. Cognitive/modelling	Step by step explaining the right solution	Body part lights up in right colour/pattern

task-specific, cognitive scaffolds. The final prompt, which was provided by visual and verbal feedback, involved the modelling of the solution process. These structured prompts were developed along the lines of the graduated prompts approach developed by Campione and Brown (1987) and subsequently extended by Resing (2000) and Resing et al. (2009). The prompts and scaffolds were provided by the electronic console when the child appeared to be unable to solve the problem independently. After the completion of each item, the child was asked to explain why he or she thought their solution was the correct one. These verbal explanations were deemed to offer additional insight into the solution process.

Electronic tangible interface

During the test sessions, the children were asked to solve each series completion task by constructing their solution puppet (eight pieces) on the console. This electronic device consisted of a 12 × 12 electronic grid, connected to a computer. The console provided all the instructions to the child, such as what to do, and how to place the puzzle pieces on the console (during pre-test, training and post-test), visual prompts by using LED-lights (during training), and prompts using voice recordings and different sounds (during training). Each puzzle piece had an electronic identification code making it possible to register which piece was placed where, and which formed the basis of the prompts/scaffolds given by the console. On the basis of this identification code, the position of each piece on the console, the sequences of movements, the timing of responses and the change of placement of the puzzle pieces could be monitored. This information was logged by the computer for subsequent analysis.

General procedure

The pre-test, both training sessions, and post-test were administered individually in a quiet room in the child's school with intervals of approximately one week. Sessions each took approximately 25–40 min. At the start of each session, children received standardised instructions concerning the test problems and the working method of the tangible console. Subsequently, the series completion items were presented one by one and displayed in an exercise booklet. The children were then asked to construct their solution, explain why this was the correct puppet and organise the puzzle pieces before the next series was presented. The console guided the children through the items of the session. No time limit was applied to the tasks.

Variables and scoring

The data gathered by the electronic board were compiled into log files, recoded into numeric data (Excel) and then transferred into SPSS for further analysis. Logged variables of interest were, for each item: accuracy of solving behaviour (number of accurately solved items), number of moves the child required for solving an item, completion time, behavioural strategies and explanations (verbal strategies). It was considered that these outcome scores would cover possible changes in accuracy, speed, efficiency and task solving behaviour as a consequence of dynamic testing.

Determination of the child's verbal strategy-use was based on their explanations after completing each item. Children were asked to reflect on their solution strategy by answering the question 'Why do you think this is the correct puppet?' These explanations were divided into three *verbal* strategy categories (1) non-inductive, non-serial; (2) partially inductive; and (3) inductive explanation (see detailed information in the upper part of Appendix 1). *Behavioural* strategy-use was scored on the basis of the detected sequences in the child's puppet part placements. These could be placed either idiosyncratically (no system could be detected, i.e. pieces appeared to be placed randomly) or by more systematic ways of arranging or analytically grouping the puppet pieces. An item could be solved by arranging the pieces in groups based on the item-features or by more systematic placement patterns following more analytical but also flexible paths. Three levels were distinguished (1) non-analytical; not following an inductively generated rule other than 'head first'; (2) partial analytical, and (3) full analytical; systematically following a grouping rule (see Appendix 1 for detailed information). For each item, the most advanced behavioural or verbal strategy was registered. Based on their verbal and behavioural strategy-use throughout the session, children were assigned to particular strategy-groups. Two groups of five classes of strategy-use were distinguished, depending on the inductive and analytical strategy-level most used at pre-test and post-test. The lower part of Appendix 1 presents an overview of this classification scheme for both verbal and behavioural strategy-use.

Results

Descriptive statistics

Before analysing the data, two one-way ANOVAs were conducted to examine possible differences between groups for initial level of inductive reasoning (Exclusion), and age. The results of the analysis with exclusion as dependent variable revealed no significant differences in level of inductive reasoning (Exclusion) and distribution of ethnic background between the different groups of children participating in the study. The analysis regarding age differences did not reveal significant differences between children's mean ages in the two treatment groups.

Series completion solving behaviour and the effect of training

The description of our analyses followed the order of hypotheses formulated in the introduction.

Firstly, we were eager to ascertain whether dynamic testing (DT) would lead to a different change in the children's series problem-solving behaviour, measured in terms of their (a) accuracy in solving items, (b) completion time, (c) number of pieces correct and (d) corrections, compared to the control-group children (C). We were particularly interested to see if these factors varied by ethnicity. Descriptive statistics are presented in Table 3 (Mean, SD at pre-test and post-test for the four dependent variables). A multivariate repeated measures ANOVA was run with Session as a within factor (Session: pre-test/post-test) and Condition (dynamic testing/control) and Ethnicity (indigenous/non-indigenous background) as between factors, and with the various dependent variables. Outcomes of this analysis can be found in Table 4. The multivariate analysis revealed significant within factor effects for Session (Wilk's $\lambda = .526$, $F(4, 105) = 23.67$, $p < .001$, $\eta^2 = .47$), Session $\times$ Condition (Wilk's $\lambda = .702$, $F(4, 105) = 11.14$, $p < .001$, $\eta^2 = .30$) and Session $\times$ Ethnicity (Wilk's $\lambda = .877$, $F(4, 105) = 3.67$, $p = .008$, $\eta^2 = .12$). Furthermore, a significant between-subjects effect for Condition (Wilk's $\lambda = .787$, $F(4, 105) = 7.10$, $p < .001$, $\eta^2 = .21$) was revealed but this was not the case for Ethnicity, Condition $\times$ Ethnicity or Condition $\times$ Session $\times$ Ethnicity. Univariate (within subjects) results related to our hypotheses are further described in the paragraphs below except in those cases where the multivariate analysis revealed non-significance.

Accuracy

The repeated measures MANOVA revealed significant univariate main effects for Session ($p < .001$, $\eta^2 = .39$) and Condition ($p < .001$, $\eta^2 = .19$), and a Session $\times$ Ethnicity interaction ($p = .018$, $\eta^2 = .05$; see Table 4). Most important for answering our hypothesis was the significant interaction effect between

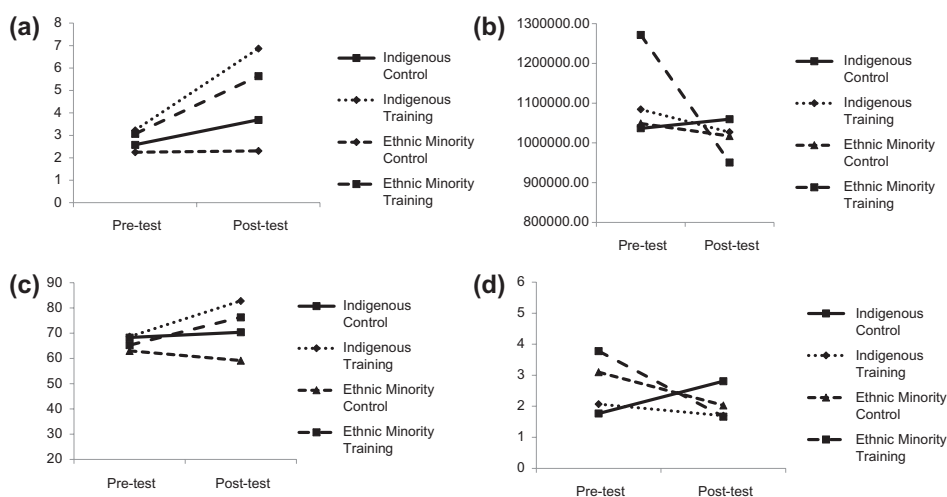


Figure 2. Mean number of correct solutions (a), completion time (b), number of pieces correct (c), and corrections (d) made per condition and ethnicity group.

Condition and Session ($p < .001$, $\eta^2 = .23$). Inspection of Tables 3, 4 and Figure 2(a) led to the conclusion that, as expected, dynamically tested children, irrespective of their ethnic background, demonstrated significantly greater progression in series completion than control-group children. The slopes of the progression lines of the two dynamically tested groups of children did not significantly differ, indicating that children in both ethnic groups, as anticipated, made comparable progress in accuracy.

Completion time

A second aspect of children's solving behaviour concerned the time they needed to complete all items of the series completion task. We expected that completion time would increase for the DT-children (on the grounds that they would give greater consideration to strategy), but remain unchanged or even decrease for the control-group children. The univariate part of the repeated measures MANOVA showed neither significant main nor significant interaction effects for completion time. The findings are displayed in Table 3, 4 and Figure 2(b). Contrary to our expectations but, nevertheless, in line with our earlier findings, neither DT-children nor controls, irrespective of their ethnic background, showed significant changes in their completion time over sessions.

Number of pieces correct

The third sub-question concerned the number of accurately laid pieces. Univariate results of the repeated measures MANOVA showed a significant main effect for Session ($p < .001$, $\eta^2 = .13$), Condition ($p = .007$, $\eta^2 = .07$) and more importantly, a significant interaction effect for Session $\times$ Condition ($p < .001$, $\eta^2 = .15$). As can be seen in Table 3, 4 and Figure 2(c), both groups of dynamically tested children showed greater progress in the total number of accurately laid pieces from pre-test to post-test compared to the children in the control group. The progression lines of the two DT-groups of children did not have significantly different slopes, indicating that children in both groups profited from training in a comparable way.

Number of corrections

We expected the dynamically tested children to show a general reduction in trial-and-error behaviour manifested by a reduction in the number of movements of pieces for each item. Univariate results showed a significant main effect for Session ($p = .049$, $\eta^2 = .03$), indicating that the number of corrections diminished over time, and a significant Session $\times$ Ethnicity interaction ($p = .003$, $\eta^2 = .08$). A trend

was found for Session $\times$ Condition ($p = .055$, $\eta^2 = .03$). Inspection of Tables 3, 4 and Figure 2(d) again indicated that dynamically tested children, irrespective of their ethnic background, tended to reduce their number of corrections from pre-test to post-test more than control-group children.

The need for prompts

Need for prompts over training sessions and prompt categories

Secondly, the number of prompts children required to solve the tasks was considered to be an indicator of their potential for learning. A decrease in the number of prompts needed from training session 1 to training session 2 was expected: if the first training session had proven effective. Findings from earlier studies led us to anticipate that children from non-indigenous backgrounds would need more prompts than their indigenous peers (see Table 5 for an overview of the prompts provided).

A repeated measures ANOVA with the required number of prompts per session as the dependent variable and Training Session as a within-factor, showed a significant effect for Training Session $F(1, 55) = 22.29$, $p < .001$, $\eta^2 = .29$, indicating that, in accordance with our prediction, children needed significantly fewer prompts in the second training session than the first. A specific analysis of the need for cognitive versus metacognitive prompts revealed that children also required fewer prompts from training session 1 to 2, $F(1, 55) = 11.8$, $p = .001$, $\eta^2 = .18$ and $F(1, 55) = 25.29$, $p < .001$, $\eta^2 = .31$, respectively. Additional repeated measures ANOVA's with one within (Training Session) and one between factors (Ethnicity), with the number of cognitive/metacognitive prompts provided as dependent variables, did not reveal a significant main effect for Ethnicity or interaction effects. These findings indicate that indigenous and children from non-indigenous backgrounds did not need significantly different numbers of (specific) prompts. Nevertheless, it should be noted that the large standard deviations shown in Table 5 are indicative of large individual differences in the number of prompts needed by individual children.

Group differences in needs for prompts and effect of training

Based on the number of prompts required and their pre-test score, children were allocated to one of four categories: low pre-test score and high number of prompts needed; low pre-test scores, low number of prompts needed; high pre-test score, high number of prompts needed; high pre-test scores, low number of prompts needed.

We examined whether the children in these four categories showed different performance patterns from pre- to post-test. A repeated measures analysis was performed with the number of accurately solved items as the dependent variable, Session as a within factor, and Prompts Category (four categories) as a between factor. This revealed significant effects for Session $F(1, 52) = 126.37$, $p < .001$, $\eta^2 = .71$, for Prompts Category $F(3, 52) = 36.38$, $p < .001$, $\eta^2 = .68$ and for the interaction Session $\times$ Prompts Category $F(3, 52) = 14.82$, $p < .001$, $\eta^2 = .46$. Figure 3 displays these results. Pair-wise comparisons suggested that training was particularly beneficial for those children who scored poorly on the pre-test but who did not need large numbers of prompts. Their accuracy in solving series completion items increased the most from pre- to post-test. Children who were already relatively accurate in solving the

Table 5. Basic statistics (means (M) and standard deviations (SD)) on prompts needed by ethnicity and session.

Training steps	Ethnicity	N	Training session 1		Training session 2	
			M	SD	M	SD
Prompts (total)	Indigenous	28	10.00	4.76	6.71	5.14
	Non-indigenous	28	11.04	4.36	9.04	4.83
	Total	56	10.52	4.56	7.87	5.08
Metacognitive	Indigenous	28	6.32	2.55	4.39	2.61
	Non-indigenous	28	6.71	2.20	5.46	2.48
	Total	56	6.52	2.37	4.93	2.58
Cognitive	Indigenous	28	3.64	2.48	2.32	2.73
	Non-indigenous	28	4.32	2.43	3.57	2.54
	Total	56	3.98	2.46	2.95	2.69

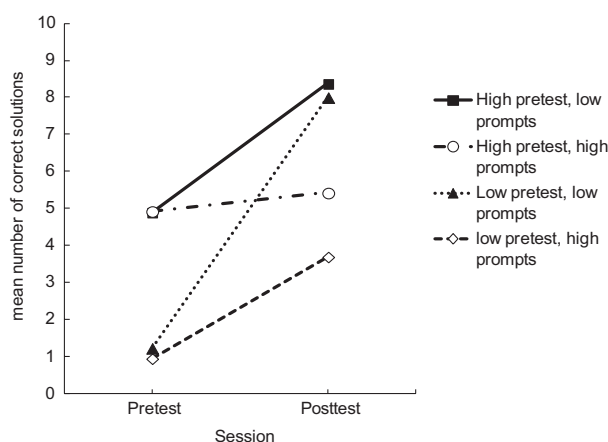


Figure 3. Mean number of correct solutions by Prompts & Pre-test category and Session.

pre-test items but nevertheless needed a considerable number of prompts did not show an increase in accuracy. However, these findings will need to be replicated in larger groups of participants.

Change in verbal and behavioural strategies

We also hypothesised that training would lead both ethnic groups to employ more sophisticated strategies. Thus, it was anticipated that their use of non-inductive verbal strategies would diminish, and the frequency of partial inductive and full inductive explanations would increase, although this trend would be less strong for the children from non-indigenous backgrounds. In a comparable way, it was expected that training would lead both ethnic groups to adopt more sophisticated behavioural strategies. A multivariate repeated measures ANOVA was performed with Session (pre-test/post-test) as within, and Condition (dynamic testing/control) and Ethnicity (indigenous/non-indigenous background) as between, factors, and the number of verbal explanations and behavioural strategies per strategy-category (inductive, partial inductive, non-inductive; analytical, partial analytical and non-analytical) as dependent variables. Multivariate effects were found for Session (Wilk's $\lambda = .327$, $F(5, 90) = 37.05$, $p < .001$, $\eta^2 = .67$), Condition (Wilk's $\lambda = .806$, $F(5, 90) = 4.34$, $p = .001$, $\eta^2 = .19$), Ethnicity (Wilk's $\lambda = .826$, $F(5, 90) = 3.79$, $p = .004$, $\eta^2 = .17$), Session $\times$ Ethnicity (Wilk's $\lambda = .772$, $F(5, 90) = 5.31$, $p < .001$, $\eta^2 = .23$), Session $\times$ Condition (Wilk's $\lambda = .698$, $F(5, 90) = 7.80$, $p < .001$, $\eta^2 = .30$), but not for Condition $\times$ Ethnicity or Session $\times$ Condition $\times$ Ethnicity. The results of these analyses are depicted in Figure 4 and 5. Univariate outcomes of this analysis are presented in the paragraph below.

For the *non-inductive verbal strategy-category*, the univariate analysis revealed a significant main effect for Session $F(5, 90) = 28.47$, $p < .001$, $\eta^2 = .23$, Condition $F(5, 90) = 9.44$, $p = .003$, $\eta^2 = .09$ and significant Session $\times$ Condition $F(5, 90) = 24.80$, $p < .001$, $\eta^2 = .21$ and Session $\times$ Ethnicity effects $F(5, 90) = 5.90$, $p = .017$, $\eta^2 = .06$. Children reported fewer non-inductive, non-advanced verbal strategies at the post-test session, and training appeared to differentially influence this change in the expected direction. Irrespective of training, indigenous children also used fewer non-inductive verbal strategies at post-test than did the non-indigenous children. For the *incomplete inductive verbal strategy-category*, no significant effects were found. The outcomes did not reveal any significant changes in the use of partial/incomplete inductive verbal strategies as a consequence of training. The analysis for the *full inductive verbal strategy-category* revealed significant main effects for Session $F(5, 90) = 37.30$, $p < .001$, $\eta^2 = .28$, Condition $F(5, 90) = 16.14$, $p < .001$, $\eta^2 = .15$ and Ethnicity $F(5, 90) = 4.05$, $p = .047$, $\eta^2 = .04$. Significant interactions were found for Session $\times$ Condition $F(5, 90) = 14.94$, $p < .001$, $\eta^2 = .14$, and Session $\times$ Ethnicity $F(5, 90) = 4.50$, $p = .036$, $\eta^2 = .05$ (sphericity assumed). Children apparently used more

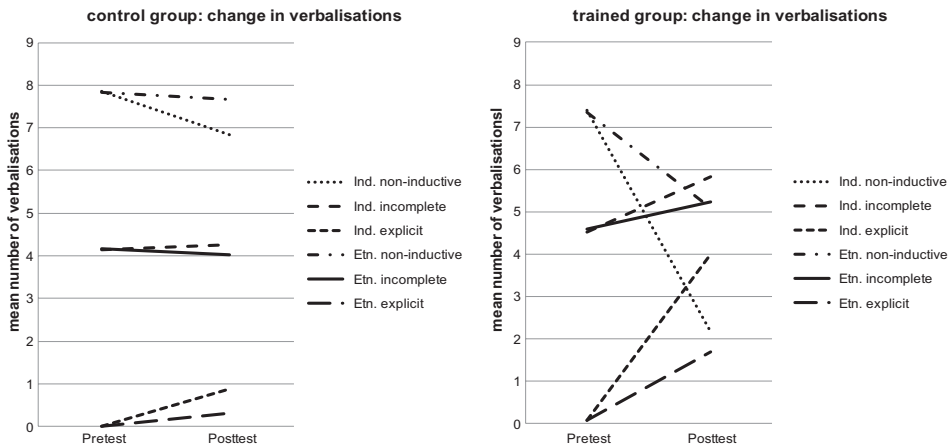


Figure 4. Progression paths of the use of different verbal strategies, displayed per condition (control versus trained group) and per ethnicity (Ind. = indigenous; Ethn. = non-indigenous group).

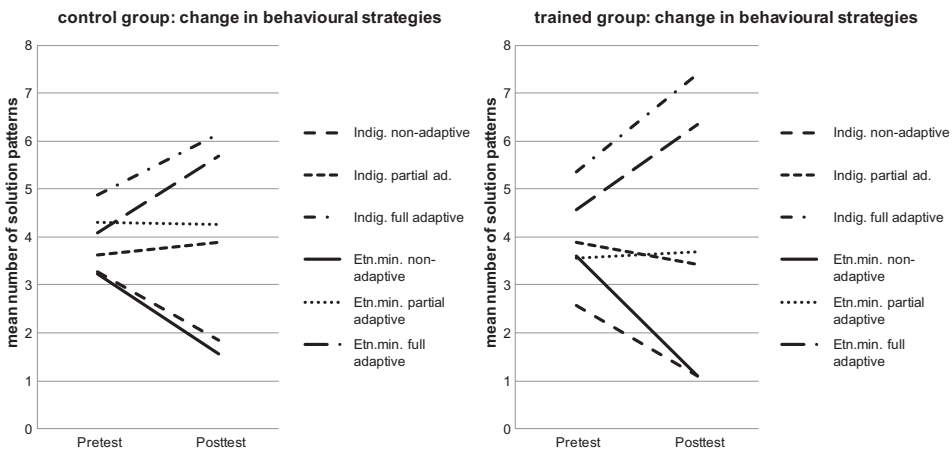


Figure 5. Progression paths of the use of different behavioural strategies, displayed per condition (control versus trained group) and per ethnicity (Ind. = indigenous; Ethn. = non-indigenous group).

advanced, inductive verbal strategies at the post-test session, and training appeared to significantly influence this shift towards more advanced verbal strategy-use. Indigenous children also used more inductive verbal strategies at post-test than did the children from non-indigenous backgrounds, and the training tended to foster this effect.

For both the non-analytical and the analytical behavioural strategy, large significant effects were found for Session, $F(5, 90) = 121.13, p < .001, \eta^2 = .56$ (non-analytical) and $F(5, 90) = 52.12, p < .001, \eta^2 = .36$ (analytical strategy), respectively. In contrast to our expectations, no significant interactions were found between Session and Ethnicity, Session and Condition, or Session $\times$ Ethnicity $\times$ Condition. Inspection of Figure 5 shows that non-analytical problem-solving behaviour diminishes from pre-to post-test, and the analytical strategy improves irrespective of condition and ethnicity. The effect sizes are large, indicating that trained and non-trained children showed considerable improvements in their performance. The results of the *partial analytical* strategy failed to reveal any significant main or interaction effect; this finding might indicate that the use of this behavioural strategy is in transition.

Table 6. Change in Verbal (upper part) and Behavioural (lower part) strategy-groups from pre- to post-test, by condition.

		1. Non-inductive	2. Mixed 1 and 3	3. Partial inductive	4. Mixed 3 and 5	5. Inductive	Missing	Total
<i>Verbal explanation – pre-test</i>								
Control	Frequency	41	3	4	2	6	2	58
	Percentage	70.7	5.2	6.9	3.4	10.3	3.4	100
Dynamic testing	Frequency	37	2	9	3	5	2	58
	Percentage	63.8	3.4	15.5	5.2	8.6	3.4	100
<i>Verbal explanation – post-test</i>								
Control	Frequency	38	1	2	2	13	2	58
	Percentage	65.5	1.7	3.4	3.4	22.4	3.4	100
Dynamic testing	Frequency	13	0	4	0	39	2	58
	Percentage	22.4	.0	6.9	.0	67.2	3.4	100
		1. Non-analytical	2. Mixed 1 and 3	3. Partial analytical	4. Mixed 3 and 5	5. Full analytical	Missing	Total
<i>Behavioural strategy – pre-test</i>								
Control	Frequency	6	2	10	7	24	9	58
	Percentage	10.3	3.4	17.2	12.1	41.4	15.5	100
Dynamic testing	Frequency	6	4	9	8	26	5	58
	Percentage	10.3	6.9	15.1	13.8	44.8	8.6	100
<i>Behavioural strategy – post-test</i>								
Control	Frequency	1	0	11	7	30	9	58
	Percentage	1.7	.0	19.0	12.1	51.7	15.5	100
Dynamic testing	Frequency	0	0	8	7	38	5	58
	Percentage	.0	.0	13.8	12.1	65.5	8.6	100

Change over time in verbal explanations and behavioural strategies

Finally, in order to examine the effects of dynamic testing upon strategy-use, the children were assigned to different strategy-groups, based upon their performance on the pre- and post-test. Crosstabs analyses (chi-square tests) were used to see how children changed their strategic behaviour over time. We sought to analyse the predicted shifts in *verbal strategy*-use by analysing the relationship between Condition (dynamic testing/control) and Verbal strategy Group (1) non-inductive; (2) mixed partial-non-inductive; (3) partial inductive; (4) mixed partial-full inductive; and (5) full inductive verbalisations (see 'Method' and Appendix 1 for more details). The pre-test results showed, as predicted, no significant association between condition and types of verbalisation ($\chi^2_{\text{pre-test}} = 2.62, p = .623$, 40% of the cells have expected count less than 5). In contrast, and as was predicted, we found a significant association between condition and verbal strategy-group for the post-test ($\chi^2_{\text{post-test}} = 28.92, p < .001$, 60% of the cells have expected count less than 5). As can be seen in Table 6 (upper part), children who were trained were more likely to transfer into the more advanced verbal strategy-groups than those who did not receive training, with a concomitant reduction in the variability of their strategy-use.

Comparable analyses were performed to examine changes in *behavioural strategy*. Again, we sought to analyse the predicted shifts in behavioural strategy-use by analysing the relationship between Condition and Behavioural strategy Group (1–5) (see 'Method' and Appendix 1 for more details). Results at pre-test and post-test showed no significant association between condition and behavioural strategies ($\chi^2_{\text{pre-test}} = .710, p = .950$; $\chi^2_{\text{post-test}} = 2.261, p = .520$) (20 and 25%, respectively, of the cells have expected count less than 5). The outcomes, presented in Table 6 (lower part), suggest that the trained children did not show a greater tendency to improve their strategic behaviour than those who were untrained.

Discussion

Dynamic assessment has long been held to be particularly valuable in assessing the cognitive abilities of children from different ethnic backgrounds (Feuerstein et al., 1979; Hessels, 2000; Tzuriel, 2013). It is

often argued that a lack of familiarity with tests and test procedures widely used in Western societies, together with a lack of self-confidence and communication problems, may lead to underperformance and correspondingly lower expectations of the child's potential (Ceci & Williams, 1997; Serpell, 2000). While dynamic testing approaches offer the promise of helping children from non-dominant cultures to improve their performance, it is likely that only a very intensive period of training will offer these children the opportunity to reach a comparable level of performance to that of their indigenous peers.

Several authors (e.g. Calero et al., 2013; Stevenson, 2012; Wiedl et al., 2014) have demonstrated that children from non-indigenous ethnic backgrounds can improve their performance following dynamic testing, although the limitations of a short intervention for children's long-term performance is recognised. The outcomes of the present study support these findings. For both non-indigenous and indigenous groups of children, our dynamic testing procedure resulted in greater accuracy, fewer corrections and reduced use of trial and error in the selection of puppet puzzle pieces than for the groups who received static testing and no training. These results parallel earlier findings in which only indigenous children acted as participants (Resing & Elliott, 2011). However, dynamic testing did not result in faster solution times. It is possible that as a result of the assistance received, children gave additional attention to the task at hand and, as a result, took more time before responding. In some contexts, taking additional time and care can be seen as a more sophisticated response to the challenges presented by complex problem-solving items.

As has been found before, progression paths for the different ethnic groups who had received training in our study did not interact but, rather, showed, as expected, equal slopes (Calero et al., 2013; Stevenson, 2012; Wiedl et al., 2014). The children from different ethnic-cultures did not fully catch up with their indigenous peers but, of course, this should not be expected on the basis of a rather short (30 min) training procedure. The finding that the progression paths of both groups had equal patterns that significantly differed from the groups of children not receiving training is valuable because it suggests that *learning processes* during dynamic testing do not differ for these groups. Children from different ethnic backgrounds did not, as a group, require a greater number of prompts although we noticed large individual differences.

Unlike the unstandardised, clinical approach to dynamic assessment advocated by Feuerstein and his followers, the graduated prompts approach offers a form of adaptive, scaffolded assistance that is readily suited to the requirement that assessments should follow a clear and consistent scientific procedure. However, a challenge for this approach has been that it can be difficult to avoid an overly mechanistic and inflexible response pattern. In particular, the assessor can struggle to respond in an adaptive fashion while simultaneously attempting to capture the many varied cognitive and behavioural responses of each child.

The tools and procedures used in the present study represent an attempt to provide an approach that uses graduated prompts in a manner that is considerably more sophisticated than was originally conceived in the seminal research of Brown et al. (Campione & Brown, 1987; Palincsar & Brown, 1984). Our approach is able to provide prompts as needed while also capturing a substantial data-set about the child's temporal and strategic responses at each stage of the assessment process. The development of sophisticated electronic equipment, such as that employed in the present study, can now enable us to systematically capture voluminous, complex data that hitherto, were largely inaccessible to clinicians and researchers.

Marked improvements in strategy-use are noted at both the group and individual level. After training, children from both the indigenous and non-indigenous groups showed progress, albeit differently, on the type of verbal strategy they employed. A clear and increased use of the more sophisticated strategy of full inductive reasoning was visible after training. The indigenous children, as expected, used more inductive verbal strategies at post-test than the group of children from non-indigenous ethnic backgrounds, and our training appeared to support this effect. At the same time, however, training did not lead the children to progress to the same extent on the behavioural strategy, mainly because all our groups of children showed comparable progression patterns. However, it is important to note that we found large individual differences in all groups.

Similar findings pertain when we consider the variability of strategy-use of individual children. Dynamically tested children, irrespective of their ethnic background, were far more likely to transfer to a more advanced verbal strategy group than those who did not undergo training, with the variability in verbal strategy also diminishing. As the results at the group level have already shown, however, the behavioural strategy measures failed to provide the same, clear picture. Some participants from each of the conditions, regardless of their ethnic background, showed considerable progression in their behavioural strategies. In attempting to explain this, we suggest that the children learned a lot from practice effects and from the generous instruction provided by the console. We also believe that the training provided by the console may need greater refinement; for example, it is possible that it yet over-emphasises verbal, at the expense of visual, cues.

A marked improvement in performance is also notable when the results are considered in the context of individual children's trajectories. Whereas only 5% of the control children demonstrated a significant change in verbalization strategies, 42% of the trained children demonstrated considerable progress towards advanced strategy-use from non-inductive to full inductive explanations of their strategies. Others were unable to provide a fully inductive explanation, most of the time providing partially explicit or implicit answers, without resorting to the use of non-inductive reasoning. These results therefore support our idea that, particularly in the case of the trained children, subgroups can be discerned that differ on the basis of their changing strategy-use. In combination with information regarding the number of prompts children need during training, and their progression in accuracy and trial-and-error, findings clearly show inter- and intra-variability in their use of strategies when solving series completion tasks. Some children in our study consistently employed advanced inductive problem-solving strategies across test sessions. Others started with multiple strategies at the first assessment session, but progressed after experiencing repeated practice. The various learning trajectories, we suggest, reflect individual differences regarding children's *potential* to deploy increasingly sophisticated ways of solving the complex inductive reasoning problems. Such differences are likely to apply equally to inductive reasoning in other academic domains such as math, or spelling (see e.g. Goswami, 1996; Klauer & Phye, 2008).

Findings from this study have provided us with clear information regarding accuracy progression in inductive reasoning and number and type of prompts individual children needed. Training appeared most beneficial for those children who obtained low scores on the pre-test, high on the post-test, but who did not require large numbers of prompts. Children from non-indigenous backgrounds, as expected, made as much progress as their indigenous peers, but did not (as earlier research in dynamic testing has suggested) make greater progress than their indigenous peers as a result of receiving assistance. This finding was relatively consistent over the different sources of data about the children's performance upon which we draw our conclusions. Our results therefore are in line with the conclusions of Brouwers et al. (2009) that cognitive variability of children from different cultural/ethnic backgrounds can best be thought in terms of differences and variability in the way that individuals approach and solve problems (see also Siegler, 1994, 2007). Our data support our prior expectation that dynamic test outcomes and information regarding changes in solving processes would provide us with insight into the underlying potential for learning of children from non-indigenous backgrounds.

It is hoped that greater understanding of individual children's problem-solving and inductive reasoning, that can be gained from relatively short and appealing interventions such as those provided in dynamic testing, may be used to inform more adaptive forms of classroom instruction for children, irrespective of their ethnic backgrounds. Further research into children's use of inductive reasoning strategies (involving different age groups) is likely to further our understanding of underlying processes and, consequently help to inform educational practice.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Alexander, P. A., Willson, V. L., White, C. S., & Fuqua, J. D. (1987). Analogical reasoning in young children. *Journal of Educational Psychology*, 79, 401–408.
- Backus, A. (2004). Turkish as an immigrant language in Europe. In T. K. Bhatia & W. C. Richie (Eds.), *The handbook of bilingualism* (pp. 689–724). Malden, MA: Blackwell.
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy, & cognition*. New York, NY: Cambridge University Press.
- Bleichrodt, N., Drenth, P. J. D., Zaal, J. N., & Resing, W. C. M. (1984). *Revisie Amsterdamse Kinder Intelligentie Test* [Revised Amsterdam child intelligence test]. Lisse: Swets & Zeitlinger.
- Bridgeman, B., & Buttram, J. (1975). Race differences on nonverbal analogy test performance as a function of verbal strategy training. *Journal of Educational Psychology*, 67, 586–590.
- Brouwers, S. A., Van de Vijver, F. J., & Van Hemert, D. A. (2009). Variation in Raven's Progressive Matrices scores across time and place. *Learning and Individual Differences*, 19, 330–338.
- Brown, R. P., & Day, E. A. (2006). The difference isn't black and white: Stereotype threat and the race gap on raven's advanced progressive matrices. *Journal of Applied Psychology*, 91, 979–985.
- Calero, M. D., Mata, S., Carles, R., Vives, C., López-rubio, S., Fernandez-parra, A., & Navarro, E. (2013). Learning potential assessment and adaptation the educational context: The usefulness of the ACFS for assessing migrant preschool children. *Psychology in the Schools*, 50, 705–721.
- Campione, J. C., & Brown, A. L. (1987). Linking dynamic assessment with school achievement. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 82–109). New York, NY: Guilford Press.
- Carlson, J. S., & Wiedl, K. H. (2013). Cognitive education: Constructivist perspectives on schooling, assessment, and clinical applications. *Journal of Cognitive Education and Psychology*, 12, 6–25.
- Cattell, R. B. (1979). Are culture fair intelligence tests possible and necessary? *Journal of Research and Development in Education*, 12, 3–13.
- Ceci, S. J., & Williams, W. M. (1997). Schooling, intelligence, and income. *American Psychologist*, 52, 1051–1058.
- Central Bureau for Statistics. (2007). *Jaarboek onderwijs in cijfers 2007* [Year book of education statistics, 2007]. Voorburg/Heerlen: Author.
- Fabio, R. A. (2005). Dynamic assessment of intelligence is a better reply to adaptive behavior and cognitive plasticity. *The Journal of General Psychology*, 132, 41–66.
- Ferrara, R. A., Brown, A. L., & Campione, J. C. (1986). Children's learning and transfer of inductive reasoning rules: Studies of proximal development. *Child Development*, 57, 1087–1099.
- Feuerstein, R., Rand, Y., & Hoffman, M. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device (LPAD)*. Baltimore, MD: University Park Press.
- Goswami, U. (1996). Analogical reasoning and cognitive development. *Advances in Child Development and Behavior*, 26, 91–138.
- Grigorenko, E. L. (2009). Dynamic assessment and response to intervention: Two sides of one coin. *Journal of Learning Disabilities*, 42, 111–132.
- Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications*. Cambridge: Cambridge University Press.
- Hessels, M. G. P. (2000). The Learning potential test for Ethnic Minorities: A tool for standardized assessment of children in kindergarten and the first years of primary school. In C. S. Lidz & J. G. Elliott (Eds.), *Advances in cognition and educational practice Vol 6. Dynamic assessment: Prevailing models and applications* (pp. 109–131). New York, NY: Elsevier.
- Holzman, T. G., Pellegrino, J. W., & Glaser, R. (1983). Cognitive variables in series completion. *Journal of Educational Psychology*, 75, 603–618.
- Jeltova, I., Birney, D., Fredine, N., Jarvin, L., Sternberg, R. J., & Grigorenko, E. L. (2011). Making instruction and assessment responsive to diverse students' progress: Group-administered dynamic assessment in teaching mathematics. *Journal of Learning Disabilities*, 44, 381–395.
- Klauer, K. J., & Phye, G. D. (2008). Inductive reasoning: A training approach. *Review of Educational Research*, 78, 85–123.
- Lidz, C. S., & Elliott, J. G. (Eds.). (2000). *Advances in cognition and educational practice Vol 6. Dynamic assessment: Prevailing models and applications*. Oxford: Elsevier.
- Millman, J., Bishop, C. H., & Ebel, R. (1965). Analysis of test wiseness in the cognitive domain. *Educational and Psychological Measurement*, 18, 787–790.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition & Instruction*, 1, 117–175.
- Qureshi, M. Y., & Seitz, R. (1993). Identical rules do not make letter and number series equivalent. *Intelligence*, 17, 399–405.
- Resing, W. C. M. (2000). Assessing the learning potential for inductive reasoning in young children. In C. S. Lidz & J. G. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications* (Vol. 6, pp. 229–262). Advances in cognition and educational practice (series editor: J.S. Carlson). Oxford: Elsevier.
- Resing, W. C. M., & Elliott, J. G. (2011). Dynamic testing with tangible electronics: Measuring children's change in strategy use with a series completion task. *British Journal of Educational Psychology*, 81, 579–605.

- Resing, W. C. M., Tunteler, E., De Jong, F. J., & Bosma, T. (2009). Dynamic testing in indigenous and ethnic minority children. *Learning and Individual Difference*, 19, 445–450.
- Resing, W. C. M., Xenidou-Dervou, I., Steijn, W. M. P., & Elliott, J. G. (2012). A “picture” of children’s potential for learning: Looking into strategy changes and working memory by dynamic testing. *Learning and Individual Differences*, 22, 144–150.
- Revelle, G., Zuckerman, O., Druin, A., & Bolas, M. (2005). Tangible user interfaces for children. In *Extended abstracts proceedings of the 2005 conference on human factors in computing systems. Special Interest Group*, (pp. 2051–2052). Portland, OR: CHI. doi:10.1145/1056808:1057095.
- Richland, L. E., Chan, T. K., Morrison, R. G., & Au, T. K. F. (2010). Young children’s analogical reasoning across cultures: Similarities and differences. *Journal of Experimental Child Psychology*, 105, 146–153.
- Serpell, R. (2000). Intelligence and culture. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 549–578). New York, NY: Cambridge University Press.
- Siegler, R. S. (1994). Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Science*, 3(1), 1–5.
- Siegler, R. S. (2007). Cognitive variability. *Developmental Science*, 10, 104–109.
- Siegler, R. S., & Svetina, M. (2002). A microgenetic/cross-sectional study of matrix completion: Comparing short-term and long-term change. *Child Development*, 73, 793–809.
- Siegler, R. S., & Svetina, M. (2006). What leads children to adopt new strategies? A microgenetic/cross-sectional study of class-inclusion. *Child Development*, 77, 997–1015.
- Simon, H. A., & Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review*, 70, 534–546.
- Stemler, S. E., Chamvu, F., Chart, H., Jarvin, L., Jere, J., Hart, L., & Grigorenko, E. L. (2009). Assessing competencies in reading and mathematics in Zambian children. In E. Grigorenko (Ed.), *Multicultural psychoeducational assessment* (pp. 157–185). New York, NY: Springer.
- Sternberg, R. J. (1985). *Beyond IQ. A triarchic theory of human intelligence*. New York, NY: Cambridge University Press.
- Sternberg, R. J. (2004). Culture and intelligence. *American Psychologist*, 59, 325–338.
- Sternberg, R. J., & Gardner, M. K. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology: General*, 112, 80–116.
- Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential*. Cambridge: Cambridge University Press.
- Sternberg, R. J., & Grigorenko, E. L. (2004). *Culture and competence: Contexts of life success*. Washington, DC: American Psychological Association.
- Sternberg, R. J., Grigorenko, E. L., Ngorosho, D., Tantufuye, E., Mbise, A., Nokes, C., ... Bundy, D. A. (2002). Assessing intellectual potential in rural Tanzanian school children. *Intelligence*, 30, 141–162.
- Stevenson, C. E. (2012). *Puzzling with potential: Dynamic testing of analogical reasoning in children*. Leiden: Division of Developmental and Educational Psychology, Department of Psychology, Faculty of Social and Behavioural Sciences, Leiden University.
- Tharp, R. G. (1989). Psychocultural variables and constants. *American Psychologist*, 44, 349–359.
- Tunteler, E., Pronk, C. M., & Resing, W. C. M. (2008). Inter- and intra-individual variability in the process of change in the use of analogical strategies to solve geometric tasks in children: A microgenetic analysis. *Learning and Individual Differences*, 18, 44–60.
- Tzuriel, D. (2013). Dynamic assessment of learning potential. In M. M. C. Mok (Ed.), *Self-directed learning oriented assessments in the Asia-Pacific* (pp. 235–255). Amsterdam: Springer.
- Van de Vijver, F. (2008). On the meaning of cross-cultural differences in simple cognitive measures. *Educational Research and Evaluation*, 14, 215–234.
- Verhaegh, J., Fontijn, W., Aarts, E., & Resing, W. C. M. (2013). In-game assessment and training of nonverbal cognitive skills using TagTiles. *Personal and Ubiquitous Computing*, 17, 1637–1646.
- Wiedl, K. H., Mata, S., Waldorf, M., & Calero, M. D. (2014). Dynamic testing with native and migrant preschool children in Germany and Spain, using the Application of Cognitive Functions Scale. *Learning and Individual Differences*, 35, 34–40.
- Williams, T. S. (1983). Some issues in the standardized testing of minority students. *The Journal of Education*, 165, 192–208.

Appendix 1

Categories of verbal and behavioural solving strategies, and their values (1 = non-advanced; 5 = most advanced) (upper part).
Strategy groups (verbal and behavioural strategies (1 = non-advanced; 5 = most advanced) (lower part).

Verbal category	Category description	
Non-inductive, non-serial (1)	No serial/inductive thinking is verbalised/explained or visible, for example, a 'failure to answer' or fantasies like 'I made this puppet here because this one is so lonely'	
Incomplete, partial inductive (2)	Only one or a few transformations in the item are mentioned inductively, for example, 'the heads are changing all the time from girl to boy to girl, then here it must be a boy' and at the same time no attention is given to the other parts of the puppet	
Inductive/serial (3)	Inductive description of the transformations in the series, for example, 'I see the legs changing from blue to green to blue, so the next one must have green legs' or more implicit description (with pointing/gestures) of the transformations in the series, for example, 'this becomes that and this becomes that as well'	

Behavioural category	Category description	
Heuristic non-analytical strategy, non-inductive & item-independent (1)	Idiosyncratic ways of solving; systematic solving of heads first or latest in the sequence; or fixed hierarchy (a first sign of structured solving of the problem). No sign of inductively solving the problem	
Partial inductive, analytical; item dependent (2)	Division of the problem-solving process by making two basic groups of pieces (pieces that change vs. pieces that remain the same). In the series the puppet's head is viewed as a separate part and is not taken into account. Or incomplete analytical solving by grouping puppet pieces only partially on the basis of Item characteristics or transformations available in the series	
Analytical, inductive strategy; item dependent (3)	Item solving by fully and adaptively grouping parts of the puppet into groups of pieces, based on item characteristics (colour, pattern, arms and body as groups) or transformations available in the series (pieces that transform in the same way grouped together)	

Strategy group	Strategy group	
Verbal explanations	Behaviour	Rules for classification
1. Non-inductive	1. Non-analytical	If non-inductive/non-analytical strategies were used the most in more than 33% of the items
2. Mixed 1 and 3: mixed partial and non-inductive strategy	2. Mixed 1 & 3: mixed partial and non-analytical strategy	If both types of strategies were seen in more than two times 33% of the items
3. Partial inductive	3. Partial analytical	If partial-inductive/partial-analytical strategies were used the most in more than 33% of the items
4. Mixed 3 and 5: mixed partial and inductive strategy	4. Mixed 3 and 5: mixed partial and full analytical strategy	If full analytical/inductive strategies and partial analytical/inductive strategies have been used in 2 times more than 33% of the items
5. Inductive	5. Full analytical	If full inductive or analytical strategies have been used the most in more than 33% of the items