

Bilingual education enhances creative fluency and flexibility over the first year of primary school

Agostini, Valeria; Apperly, Ian; Krott, Andrea

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version

Peer reviewed version

Citation for published version (Harvard):

Agostini, V, Apperly, I & Krott, A 2014, 'Bilingual education enhances creative fluency and flexibility over the first year of primary school', *Bilingualism: Language and Cognition*.

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Bilingual education enhances creative fluency and flexibility over the first year of primary school

Valeria Agostini, Ian Apperly, Andrea Krott

School of Psychology, University of Birmingham, United Kingdom
Centre for Developmental Science, University of Birmingham, United Kingdom

Corresponding author:

Valeria Agostini valeria.agostini4@gmail.com ORCID <https://orcid.org/0000-0003-0314-2998>

Co-authors:

Ian Apperly i.a.apperly@bham.ac.uk ORCID <https://orcid.org/0000-0001-9485-563X>

Andrea Krott a.krott@bham.ac.uk ORCID orcid.org/0000-0002-1667-3424

Data availability Statement

All data, analysis and materials for the experiment are available at the open science framework at https://osf.io/z7p23/?view_only=4fa2dbbb2623468b9b6a60a61c714829

Funding statement

This work was conducted as part of the PhD of the first author and financed by the Italian Government.

Conflict of Interest disclosure

The authors declare no conflicts of interest.

Ethics approval statement

The study protocol received full ethical approval from the Science, Technology, Engineering and Mathematics ethics Committee of the University of Birmingham, United Kingdom (ERN_17-0324)

Permission to reproduce material from other sources N/A

Acknowledgements: This work was conducted as part of the PhD of the first author and financed by the Italian Government. We would like to thank the part-taking schools, children and their parents. We are also grateful for the help of Fatima Mahroof and Hudaibia Mahroof with the scoring of the tasks.

Research Highlights

- Longitudinal study suggests that foreign language exposure in early formal education boosts creative thinking skills
- Children attending a bilingual school for six months significantly showed enhanced creative thinking skills compared to a monolingual control group without any foreign language provision
- Children exposed to weekly language lessons for six months increased creative thinking skills but not significantly compared to a monolingual control group
- Enhancements of creative thinking skills were unrelated to enhancements of cognitive control or cognitive flexibility skills

Abstract

Can exposure to a foreign language in the first year of school enhance divergent thinking skills? Ninety-nine monolingual children from predominantly White neighbourhoods ($M_{\text{Age}} = 57.7$ months, SD 1.2; 47 girls) attending bilingual schools, schools with weekly foreign language lessons, or schools without a foreign language provision (= controls) completed a divergent thinking and executive function tasks at the beginning of the school year and 24 weeks later. The groups did not differ on creativity measures at the beginning of the school year. Only bilingual school children and weekly language learners improved divergent thinking at the second testing point, with the former significantly outperforming controls on creative fluency and flexibility. Improvements could not be explained by executive function development. Therefore, a considerable amount of exposure to a foreign language in early formal education appears to boost creative thinking.

Keywords: bilingualism, second language learning, creativity, divergent thinking, bilingual education

Early bilingual experience, that is the experience of growing up with two languages from early childhood, has been shown to be beneficial for creative skills, most notably divergent thinking skills, which is the ability to produce a number of unique and original solutions to a problem¹ (e.g., Adi-Japha, Berberich-Artzi, & Libnawi, 2010; Leikin, 2013; Leikin & Tovli, 2014; Leikin, Tovli, & Malykh, 2014; review in Ricciardelli, 1992; van Dijk, Kroesbergen, Blom, & Leseman, 2019) (but see Booton, Hoicka, O'Grady, Chan, & Murphy, 2021; Lange, Hopman, Zemla, & Austerweil, 2020). There is some evidence that similar benefits are achieved via learning a second language (L2) in formal education, which is generally referred to as foreign language learning (e.g., Ghonsooly & Showqi, 2012; Landry, 1974). However, these studies have been cross-sectional, have not evaluated what “dose” of L2 might be required, have not controlled for potential group differences such as general intelligence (i.e. the ability to solve complex problems) or socio-economic status (SES), and have not examined whether improved creativity may be a secondary effect of L2 education on improved executive function. The present study addressed these shortcomings.

Enhancements in creative thinking through speaking a second language is important because higher creativity is positively related to children’s school performance (e.g., Hansenne & Legrand, 2012), and on a societal level, creativity is crucial for the creation of new products and jobs (Sternberg, 1999). Being able to speak another language might therefore benefit children’s academic success and society as a whole. For this reason, it is important to establish whether advancements in creative thinking are limited to individuals who learn to speak another language from very early on, or whether later exposure to a foreign language at school (i.e. “L2 exposure”) is similarly beneficial.

¹ While divergent thinking is only one aspect of creativity, namely the thought processes for creative thinking, we here use the terms creativity, creative thinking and divergent thinking interchangeably for the ability to produce multiple solutions to a problem (Guilford, 1967).

There is indeed some limited evidence that creativity might not only be boosted through learning an L2 in natural settings as in bilingualism. There are reports for creative benefits for individuals who are acquiring a second language through foreign language classes (Carringer, 1974; Ghonsooly & Showqi, 2012; Landry, 1973a, 1973b, 1974). What is unknown, though, is whether participants who showed advanced creative thinking skills in these studies already differed in these skills before they started learning a foreign language. Also, creativity in children is related to socio-economic status (Hendrie Kupczynszyn et al., 2024) and general intelligence (Krumm et al., 2018; Silvia 2008). Previous studies had not sufficiently controlled such factors. It is therefore important to conduct a longitudinal study that takes into account potential confounding variables.

While bilingualism has been found to positively impact creative thinking, it has been suggested that enhancement in creative thinking depends on the degree of bilingualism. It seems that a high degree of proficiency in both languages is needed (Creo, Mareque, & Pino-Juste, 2021; Kharkhurin, 2011; Kim, 2011; Lee & Kim, 2010, 2011; Leikin, Tovli, & Woldo, 2020; Ricciardelli, 1992; van Dijk et al., 2019) (but see Hommel, Colzato, Fischer, & Christoffels, 2011). This means that learning an L2 in formal education should have no effect on creative thinking until a high proficiency in L2 is reached. But enhancements in creative thinking have been reported also for those who were in the process of acquiring a second language through foreign language classes (Ghonsooly & Showqi, 2012; Landry, 1973a, 1973b, 1974). It is therefore unclear whether the amount of L2 exposure in educational settings would make a difference, that is whether only a high exposure leads to enhanced creativity.

Apart from the question whether learning an additional language in an educational setting impacts creative thinking skills, it is also still unclear how it might do so. Creativity has been related to basic executive functions (EFs), especially cognitive flexibility, that is the

ability to shift between rules or sets, and to attentional control (e.g., Chen, He, & Fan, 2022; Edl, Benedek, Papousek, Weiss, & Fink, 2014; Kharkhurin, 2011; Pasarín-Lavín et al., 2023; Sampedro & Peña, 2019; Zabelina & Robinson, 2010). Because bilingualism has been related to enhanced EF (review in Barac, Bialystok, Castro, & Sanchez, 2014), advantages in creative thinking skills in bilinguals have been argued to be due to enhanced EF. Sampedro and Peña (2019) found that cognitive flexibility was the mediating factor between level of bilingualism and creativity in preadolescent Basque-Spanish bilinguals, and Kharkhurin (2011) found that selective attention contributed to differences in creative abilities in a group of bilingual college students with various language backgrounds. However, Leikin et al. (2020) found no relationship between EF and creativity in a sample of bilingual adults. It has not been tested yet whether any advancement in creativity through L2 learning in the classroom is mediated by an accelerated development in EF.

Given these open questions, the present study addressed three aims. The first and main aim was to test whether creative abilities can be enhanced by exposure to a foreign language in the first year of formal education. More specifically, we tested children with and without a foreign language provision at school in their divergent thinking abilities. Given the shortcomings of previous studies, we recruited monolingual children from monolingual households and checked that participant groups did not differ in confounding variables such as SES, IQ, language skills or extra-curricular activities when they embarked on their L2 learning journey (Table 1). It is likely that learning a foreign language enhances children's expressive language skills. Since we were interested in children's creative thinking skills and not so much in their language skills, we tested children's divergent thinking abilities by administering a non-verbal task, namely a modified version of the "repeated figures" task of the Torrance Test of Creative Thinking (TTCT; Torrance, 1966). Applying a longitudinal design, we administered the test at the beginning of the first year of primary school and 24

weeks later. We expected children to perform very similarly at the beginning of primary school. If exposure to a foreign language enhances creative skills, we expected that children with a foreign language provision would score higher on divergent thinking skills than children without any foreign language provision at the second testing point.

The second aim of the study was to investigate the effect of the amount of second language exposure and therefore, indirectly, the effect of L2 proficiency. We therefore tested two groups of children with a foreign language provision. The first group attended bilingual schools (BilS) and therefore had a substantial exposure to a new language. The second group (L2) attended mainstream education and had a very limited L2 exposure through weekly short lessons. These groups were compared against a third group that attended schools without any L2 provision (NoL2). We expected that children exposed to a second language would show enhanced progress in divergent thinking measures, with the caveat that the very limited exposure to a second language in the weekly language learner group might not be sufficient to significantly boost children's creative development compared to no second language exposure.

The third aim of the study was to investigate whether any advanced creative skills development due to second language exposure would be related by accelerated development in EF skills. More specifically, we tested whether children's divergent thinking development could be explained by their development of two EF sub-skills, namely selective attention and general switching abilities. These were measured by means of the Attentional Network Task (ANT; Rueda et al., 2004) and the Dimensional Change Card Sort task (DCCS; Zelazo, 2006), respectively. We expected EF skills to increase between the two testing points. If accelerated development in divergent thinking skills in the two foreign language learning groups was related to enhanced development in EF skills, then the development in EF should mediate development in divergent thinking.

Method

Participants

We estimated that we would need at least 87 participants for a medium effect size $f = .25$, $\alpha = .05$, $\text{power} = .90$, 3 participant groups, 2 measurements, correlation among repeated measures = .2 and a nonsphericity correction $\epsilon = 1$. Participants were recruited by approaching parents via schools that fell into our three target categories (BiL2, L2, NoL2). We recruited 103 4-5 year-old children who attended the first year of primary school in predominantly White neighbourhoods in the United Kingdom and who were monolingual before entering school. Four of these children were removed from the analyses because they did not complete all tasks. The remaining 99 participants belonged to three different groups: One group was recruited from two bilingual schools (BiL2) ($n = 32$), located in Oxfordshire and South East London (United Kingdom). About 50% of their education was taught in a language other than English. One school taught core subjects (English, maths, and science) in English in the mornings and the remaining subjects (e.g., sports, arts) in French in the afternoons. The other school taught all subjects in English for half the week and in German, French, or Spanish for the other half. The second group of children (L2 learners, $n = 29$) attended one of four mainstream schools that provided weekly 30-60 minute foreign language lessons (German, French, or Spanish). The third group of children (NoL2, $n = 38$) attended one of two mainstream schools without a foreign language provision. All schools followed the national curriculum. The first year of primary school teaches children primarily through games and play, with activities in small groups and free play. Expressive arts and design is one of the seven areas of learning next to communication and language, personal, social and emotional development, physical development, literacy, mathematics and understanding the world. We are not aware of any engagement in expressive arts that went beyond the

curriculum in any of the schools. The majority of schools were state-funded, but 2 of the 4 schools with weekly language provision were private schools that charge a fee.

Table 1 lists the characteristics of the children in the three groups. All children and their parents were born in the United Kingdom or another English-speaking country. They had lived in the United Kingdom since birth, apart from two bilingual school children who had lived in Singapore in their first year of life (between 5 weeks and 15 months). All children were exposed exclusively to English at home. To ensure that the three groups of children did not differ in terms of cognitive abilities at the beginning of the study, we tested them on an English vocabulary test and a non-verbal IQ measure. We also compared them on age, gender, maternal and paternal background, SES, hours of extra-curricular activities, hours of computer usage, and number of (older) siblings. The three participant groups did not significantly differ with regards to any of these characteristics at the start of the study, apart from age, $F(2,96) = 7.2, p = .001, \eta p^2 = .131$. NoL2 children were significantly older than BilS children, $p = .001$, and marginally older than L2 learners, $p = .063$. Importantly, age did not correlate with any measures of divergent thinking or EF at either T1 or T2 but with improvement on the Dimensional Card Change Sort task (DCCS), Pearson $r(99) = -.252, p = .012$, and creative flexibility improvement, Pearson $r(99) = -.221, p = .028$, all other $ps > 0.05$. We thus partialled out any effect of age when analysing improvements.

Materials

Background Information Questionnaire

We administered a modified version of the Language and Social Background Questionnaire (Luk & Bialystok, 2013) to gather information on, for instance, children's gender, date of birth, extra-school activities, computer usage, maternal and paternal level of

education, annual family income, and languages spoken by participant. We only tested children who were monolingual and had no exposure to any non-English language.

Each child's socio-economic status (SES) was determined by averaging indices of parental education, occupational status and annual family income. Level of each parent's education was measured on a 6-point scale, with 1 being 'no formal educational qualification' and 6 'masters/doctoral degree, National Vocational Qualification level 5, or equivalent'. A parental index was obtained by averaging the scores of both parents. Occupational status was classified according to the Standard Occupational Classification Hierarchy redacted by the Office for National Statistic. An occupational index was calculated by averaging the scores of both parents (if applicable). Annual family income was obtained via a 7-point scale, with 1 being 'less than £15,000' and 7 '£65,000 or more'. To combine these indexes into a SES index, all three indices were converted onto a scale of 0-1.

General Intelligence

We measured children's intelligence with the means of the Raven's Coloured Progressive Matrices (Raven et al., 1990), a widely-used measure of non-verbal abstract reasoning. Participants are asked to complete visual patterns by choosing from 6 options of missing parts. The test was presented using Eprime (E-Studio 2.0), with participants pointing to the chosen missing part and the experimenter pressing a corresponding button on a keyboard. Each correct answer was given one score (maximum 36).

Vocabulary

We assessed children's language ability with the British Pictures Vocabulary Scale III (BPVS; Dunn et al., 2009), a standardised measure of receptive vocabulary for children aged 3 to 16 years. Children are asked to point to pictures (out of 4 options) that correspond

to words produced by the experimenter. The test contains 14 sets of 12 words. The test is terminated when a child makes 8 or more errors within a set. The test score is calculated as the total number of tables presented minus the total number of errors.

Divergent thinking

We tested divergent thinking skills with means of a modified version of the “Repeated figures” task from the figural form of the TTCT (Torrance, 1966), using different stimuli. This task assesses the ability to respond to a repeated stimulus with different ideas. At T1 children were presented with a landscape-oriented A4 sheet of paper with eighteen 28x28mm squares arranged in three rows of six squares. At T2, each square was replaced with two 3-cm-long horizontal parallel lines. Thus, children were given squares at T1 and lines at T2. Children were asked to “turn the squares (or the lines) into something else with their drawing”. They were given ten minutes to produce as many different responses as they could. They were told to think of something different every time and to think of something that only they could think of. The experimenter explained the task at both T1 and T2 by showing two example responses. At T1, the experimenter pointed to the first example response and said: “For example, I turned this square into a house” and, tracing over the drawing with a pencil, continued “I drew a roof here on top, a door, windows... and now it is not a square anymore, it is a house!”. Pointing to the second example, they continued “Then, I thought of something else and turned this other square into a window. You need to turn each square into something different every time. You have ten minutes to turn as many squares as you can into something else.” The same procedure was followed at T2. This time, the children were presented with two pairs of lines turned into a butterfly and a bag, respectively.

If children duplicated the examples given or replicated their own responses, the experimenter encouraged them to think of something different for the next square or pair of

lines. If children stopped drawing, the experimenter praised the responses that they had already given and encouraged them to think of something else to turn the squares/lines into. After each drawing children were asked to say what they drew, and, whenever it was not evident, they were also asked to explain what the square (or the lines) represented in their drawing. Children's answers were noted by the experimenter.

Responses were scored after all children had participated. A general rule for accepting or rejecting participants' responses was to determine whether the stimulus had been integrated into the drawing. Given the young age of the children, and their poor fine motor skills, it was sometimes not easy to establish the role of the stimulus (the square at T1 and the lines at T2) in their responses purely on the basis of what they drew. Therefore, we also took into account what children said about their drawings. As mentioned, whenever the role of the stimulus was not obvious, children were asked "What is the square/are the lines in your drawing?". If children's answers clearly showed that they meant to incorporate the stimulus but failed to do so because of practical inability, then the response was deemed valid. For example, sometimes children tried to trace over the square to integrate it into their drawing but, not being able to do so, drew a line around it instead. When asked what the square was, they answered that it was a part of the object that they depicted, for instance the seat of a chair. In such a case, the response was counted as valid. Conversely, if children showed that they used the stimulus only as an area to draw in and produced a drawing unrelated to the stimulus, then the response was deemed invalid. This happened especially at T1, when children tended to draw something inside the squares. For example, one child drew a snowman inside the square and said that the square was "snow around the snowman". This response was rejected.

As typical for the TTCT, responses were scored on three components of divergent thinking, namely fluency, flexibility, and originality. For the fluency score, one point was

given for each valid response. Plain replications of the examples presented to them or of their own previous responses were not considered as valid. However, variations of the examples given (e.g., “a haunted house”) or of their previous responses (e.g., “a doll’s face”, “a pirate’s face”, “a dog’s face”, etc.) were counted for fluency. A higher fluency score corresponded to superior creative fluency.

As for flexibility, the number of different response categories were counted. Following Beck et al. (2016), categories were determined across all responses of all participants at each testing point. This was done unanimously by the experimenter and a blind scorer. Categories are listed in Supplementary Table 1. Additional categories with single members were established when responses could not be assigned to any of these categories. Children were given one point for flexibility for each category they produced. A higher flexibility score corresponded to superior creative flexibility. The maximum possible score for both fluency and flexibility was 18 points.

For originality scores, we determined how many children gave a specific response. Thus, if 10 children had turned the square into a school, a ‘school response’ would be given 10 points. To determine an originality score for a particular child, we averaged the points for their responses and subtracted the score from the maximum score 99 (see 99 participating children), so that a higher originality score corresponded to superior originality (maximum possible score 98). We then divided this number by their fluency score. The latter was done to avoid that the originality score was conflated with fluency (Hocevar, 1979).

Selective Attention

We measured selective attention skills by implementing a common task, the Attention Network Test for children (ANT). We used a child-friendly version of the task by Rueda et al. (2004), in which participants were presented with rows of five identical fish. They were

asked to ‘feed’ the central fish by pressing a left button with a fish facing left or right button with a fish facing right, depending on the direction of the central fish. There were three conditions. In the neutral condition, a single fish was presented. In the congruent condition, the flanking fish faced into the same direction as the target fish, while in the incongruent condition, they faced into the opposite direction. A trial started with a fixation cross in the centre of a laptop screen for 400 ms, followed by a 150 ms attentional cue (an asterisk) and the stimulus which appeared after 450 ms. Attentional cues were presented either in the position of the fixation cross or above or below it, while stimuli were presented above or below the fixation cross. The maximum response time was 1700 ms.

In case of a correct response, the target fish opened its mouth and bubbles appeared, while a “woohoo” sound played. In case of incorrect or missing responses, a ‘wrong buzzer’ sound was played. The experiment was implemented in E-prime (E-Studio 2.0). Responses were recorded using a Cedrus RB-844 response pad. The experiment started with 24 practical trials, followed by two blocks of 48 trials each, with an equal number of all conditions and an equal number of right and left-facing target fish. Trials were presented randomly. We recorded both reaction times (RTs) and accuracy.

We followed the procedure by Rueda et al. (2004) to introduce the task to the participants. The experimenter placed a picture of a right-facing fish on the table above the button box. They told the participant that the fish was very hungry and that they need to feed him by pressing one of the two buttons on the button box. After pointing to the picture and the corresponding button, the experimenter explained “Sometimes the fish is facing this way. In this case, you need to press this button”. They then presented a picture with a target fish facing left and pointed to the left button, while explaining “but sometimes the fish is facing the other way. So, which button do you need to press if he is facing the other way?”. They showed them the correct button if the child did not respond. Next, they presented a picture

with a row of fish all facing left. While pointing to the central one, they said “Sometimes the fish is not alone on the screen, he is with other fish, but you need to look at him in the middle. Which button do you press to feed him?” The same was repeated with a row of fish facing to the right, asking “which button do you press if he is facing the other way?”. Finally, the experimenter presented a picture with incongruent flankers (either facing right or left). While pointing to the target fish, they explained: “I told you that you always have to look at the fish in the middle and not at the others because sometimes the other fish are tricky. They want to trick you and they go in the opposite direction. But you need to feed the fish in the middle and not the others. So which button do you press for the fish in the middle?”. The same was repeated for the other incongruent stimulus, asking “and which button do you press for this other one?”. In case of erroneous responses, correct responses were indicated and further explanations were given. The experimenter explained to the children that there would be a small cross on the screen and they had to keep their eyes on the cross as the target fish would occur above or below. Attentional cues were not mentioned. In addition, children were told to keep their index fingers on the two buttons and to press the buttons as quickly as possible when they see the fish.

The ANT provides indices for three skills: inhibitory (or cognitive) control, alertness and orientation. We calculated the Conflict index by subtracting performance on congruent trials from that on incongruent trials, and we did so for both accuracy and RT. These two subindices were standardised and summed, after reversing RTs so that higher values of both RT and accuracy represented better performance. For the alerting index, we subtracted the mean of the Double Cue condition from the mean of the No Cue condition so that a low value corresponded to a higher alertness for accuracy and a high value corresponded to higher alertness for RTs. We calculated an alerting index by adding standardised reversed accuracy scores and standardised RTs, so that higher values corresponded to better performance. For

the orienting index, we subtracted the mean of the Spatial Cue condition from the mean of the Central Cue condition. This way, a low value corresponded to a better orienting for accuracy and a high value corresponded with better orienting for RTs. We calculated the orienting index by adding standardised reversed accuracy scores and standardised RTs, so that higher values corresponded to higher performance.

Cognitive Flexibility

We assessed children's cognitive flexibility by means of the Dimensional Change Card Sort task (DCCS; Zelazo, 2006). Given the children's ages, we administered both the standard and the advanced versions, summing the two scores into a total score. Both versions of the task required children to sort 2 practice cards and 12 experimental cards (7x11 cm) according to either shape or colour of the object they depict. In the standard version, children were asked to sort the first half of cards by colour and the second half by shape. At the first testing point, 50% of the cards showed a red rabbit and 50% a blue boat. Children were asked to sort the cards into two transparent containers (16.8 cm x 11.6 x 4.5 cm). The containers were marked with one card each (i.e. a target card), attached to the back and clearly visible to the participants. One target card depicted a red boat, the other one a blue rabbit. This meant that the target cards did not show identical pictures to the ones that needed to be sorted.

Children were told that they were going to play a game with some cards. The experimenter explained: "The first game is called the *colour game*. In the colour game, you need to sort the cards by colour, so all the red cards go here [pointing to the container featuring the red target card] and all the blue cards go here [pointing to the container featuring the blue target card]". Next, the experimenter presented a red rabbit practice card, explaining: "Here's a red card. Where does it go?" After children pointed to the correct container, they were asked to place the card into it turning it facedown. This was repeated for

a blue boat practice card. If a child placed the card into the wrong container, the correct response was demonstrated. From then on, three red rabbit cards and three blue boat cards were presented in random order and children did not receive any feedback. After the completion of the 6 pre-switch trials, the experimenter introduced the switch by explaining: “Now we are changing game, we are playing another game called *picture game*. In the picture game, you need to sort cards by picture, so all the rabbits go here [pointing to the container with the target rabbit card] and all the boats go here [pointing to the container with the boat target card]. Here’s a rabbit, where does it go?”. The experimenter handed the card to the child who placed it without receiving feedback. Again, the child was asked to sort 3 red rabbit cards and 3 blue boat cards in pseudorandom order. Notably, in each trial of the standard version of the task, the experimenter mentioned the relevant feature of the card before handing it to the child. Thus, they would say “Here’s a red/blue card” for pre-switch trials and “Here’s a rabbit/boat card” for post-switch trials. Each correct response was awarded one point, with a maximum of 12 points.

After the completion of the standard version, the experimenter moved onto the advanced version of the task, in which the sorting rule could change from card to card and was depicted on the cards. Cards to be sorted by colour showed a rainbow, and cards to be sorted by shape showed black outlines of a rabbit and a boat. Again, children were asked to sort two practice cards and 12 testing cards, that is eight red rabbit cards and six blue boat cards. Half the cards indicated the colour sorting rule, the other half indicated the shape-sorting rule. Cards were given to the child in a pseudorandomised order. In half the trials, the rule stayed the same as the previous one, while in the other half, the rule switched.

The experimenter explained the advanced version of the task to the participant by saying: “Now we are going to play another game with this special set of cards. In this set, some cards have a rainbow and some cards have two small pictures. If the card has got a

rainbow [showing a rainbow-labelled rabbit card], it's a colour-game card and you have to play the colour game, but if the card has got two small pictures [showing an outlines-labelled rabbit card], it's a picture-game card and you need to play the picture game". The experimenter handed the rainbow card to the child, explaining: "This has got a rainbow, so it's a colour game card. Can you remember the rule for the colour game?" In case the child did not remember the rule, the experimenter said: "In the colour game, red cards go here and blue cards go here [pointing to the corresponding target cards]. This is a colour game card, where does it go?" In case the child did not place the card into the correct container, the experimenter showed them where it goes and explained the rule again. This was repeated with the second practice card, a picture-rule rabbit card. After the practice trials, 12 trials followed (3 colour-rule rabbit cards, 3 picture-rule rabbit cards, 3 picture-rule boat cards and 3 colour-rule boat cards), presented in a pseudo-random order. For all cards, the experimenter indicated the rule of the card when handing it to the child by saying, "Here's a colour-game/picture-game card". As for the standard version, each correct response was awarded one point, with a maximum of 12 points. When scoring responses, only first attempts were counted, ignoring if the child changed their mind. Taking the scores for the two versions of the task together, children could reach a maximum of 24 points.

The same procedure was repeated at the second testing point, but with different objects depicted on the cards and different colours (yellow flower and blue car).

Procedure

The parents of all children consented to the participation and filled in the Background Information questionnaire. Children were tested on a set of tasks at the beginning of the first year of primary school and repeated 24 weeks later. Testing at both time points took place in the same order on three days over the course of two weeks. On day 1, children completed the

Coloured Progressive Matrices (Raven, Court, & Raven, 1990), the modified version of the figural thinking task of the TTCT (Torrance, 1966), and the DCCS (Zelazo, 2006). On day 2, they took part in the ANT (Rueda et al., 2004) and the British Picture Vocabulary Scale III (BPVS; Dunn et al., 2009). On day 3, they took part in a Theory of Mind task, which is unrelated to the present research questions and will be reported elsewhere.

Results

In the following, we started by addressing our first two aims, namely whether creative abilities can be enhanced by exposure to a foreign language in the first year of formal education, and if so, whether this relationship is moderated by the amount of second language exposure. For that, we report the results of the divergent thinking task. Next we tested whether similar changes were observed for EF. Finally, we tested whether changes in EF mediated the effects of foreign language exposure on creative abilities.

Divergent thinking task

Responses were scored independently by the experimenter and a blind scorer. Both scorers were blind as to group membership of the children. The two agreed on the scores in 76% of all responses at T1 and in 78% at T2. Differences in scoring were discussed and a final score agreed. Similar number of responses were rejected in the three groups due to being invalid (BiL2: T1: mean 2.0 (SD 1.5), T2: mean 0.6 (SD 1.0); L2: T1: mean 2.2 (SD 1.7), T2: mean 1.6 (SD 1.6); NoL2: T1: mean 2.6 (SD 2.5), T2: mean 1.9 (SD 2.0)). Figure 1 shows the groups' valid creative fluency, flexibility and originality scores at T1 and T2.

We analyzed potential group differences in the multivariate pattern of the three divergent thinking measures (fluency, flexibility, and originality) across the two time points using a two-way multivariate analysis of variance (MANOVA) with the within-subjects

factor Time (T1 vs T2) and the between-subjects factor Group (BilS, L2, No-L2). In case of a Time x Group interaction, we report two series of posthoc MANOVAs. First, we report group comparisons at T1 and T2. For any significant group effect, we report post-hoc MANOVAs comparing pairs of groups. Second, we report development in divergent thinking between T1 and T2 for each group. In order to better understand the development of the groups in the submeasures of divergent thinking (fluency, flexibility and originality), we also report the results of univariate analyses.

The MANOVA for the multivariate pattern of fluency, flexibility and originality and age as covariate showed a significant effect of Time, *Wilks' Lambda* = .611, $F(3, 93) = 19.7$, $p < .001$, $\eta^2 = .389$, a significant effect of Group, *Wilks' Lambda* = .817, $F(6, 186) = 3.3$, $p = .004$, $\eta^2 = .096$, and a marginally significant Time x Group interaction, *Wilks' Lambda* = .880, $F(6, 186) = 2.0$, $p = .062$, $\eta^2 = .062$. Posthoc multivariate ANOVAs at each time point with age as covariate showed no effect of Group at T1, *Wilks' Lambda* = .952, $F(6, 186) = .8$, $p = .594$, $\eta^2 = .024$, but an effect of Group at T2, *Wilks' Lambda* = .836, $F(6, 186) = 2.9$, $p = .010$, $\eta^2 = .085$. Follow-up MANOVAs comparing pairs of groups on the multivariate pattern of the divergent thinking measures at T2 (with age as covariate) showed that BilS scored higher than NoL2 children, *Wilks' Lambda* = .840, $F(3, 65) = 4.1$, $p < .010$, $\eta^2 = .160$. L2 learners fell in between the other groups, not scoring significantly different from either, BilS: *Wilks' Lambda* = .904, $F(3, 56) = 2.0$, $p = .125$, $\eta^2 = .096$; NoL2 children: *Wilks' Lambda* = .934, $F(3, 62) = 1.5$, $p = .236$, $\eta^2 = .066$.

While the groups differed at T2, post-hoc MANOVAs for each group with age as covariate showed that all groups improved their divergent thinking across the two timepoints, BilS: *Wilks' Lambda* = .475, $F(3, 28) = 10.3$, $p < .001$, $\eta^2 = .525$; L2 learners: *Wilks' Lambda* = .599, $F(3, 25) = 5.6$, $p = .004$, $\eta^2 = .401$; NoL2 children: *Wilks' Lambda* = .763, $F(3, 34) = 33.5$, $p = .025$, $\eta^2 = .237$.

Figure 1 suggests that the groups' development somewhat differed for the three divergent thinking measures (fluency, flexibility, and originality). Univariate analyses of the measures showed a divergent pattern for originality, but results for creative fluency and flexibility confirm the overall MANOVA pattern (see Table 2 and Supplementary Material).

In sum, the three participant groups did not differ in their divergent thinking skills at T1. They improved their skills across the two timepoints to different degrees so that they differed at T2, particularly for creative fluency and flexibility. BilS children outperformed NoL2 children and L2 learners fell in between, not significantly differing from either of the other two groups.

Executive Function measures

Next, we tested whether the groups differed in their development of EF, conducting equivalent MANOVAs as for the divergent thinking measures above, but with DCCS and ANT indices (ANT measures inhibition, alertness, and orienting) as dependent variables. Again, we explored the development of the groups in the EF sub-measures with univariate analyses.

Some children had a high error rate in the ANT task. We therefore included only children who reached at least 60% accuracy at T1 (57 trials) in the analysis of executive functions, which corresponded to above chance performance (57 out of 96 total responses has a one-tailed probability of 0.041). This led to a reduction of the sample from 99 to 76 children (26 BilSs, 24 WL2s and 26 NoL2s). For the analysis, we removed all errors, response omissions, as well as responses < 200 ms (anticipatory responses) and > 2.5 SDs above the mean of each participant (19.5% of all trials).

Figures 2 and 3 show the development of the three groups in the EF measures across the two timepoints. A two-way MANOVA of the multivariate pattern of executive functions

(DCCS scores and the ANT measures of conflict index, alerting index, and orienting index) and age as covariate showed a significant effect of Time, *Wilks' Lambda* = .394, $F(4, 69) = 25.6$, $p < .001$, $\eta p^2 = .5606$, no significant effect of Group, *Wilks' Lambda* = .829, $F(8, 138) = 1.7$, $p = .104$, $\eta p^2 = .090$, and a significant Time x Group interaction, *Wilks' Lambda* = .720, $F(8, 138) = 3.1$, $p = .003$, $\eta p^2 = .151$. Posthoc MANOVAs at each timepoint (controlling for age) showed an effect of Group at both timepoints, T1: *Wilks' Lambda* = .666, $F(8, 138) = 2.5$, $p = .016$, $\eta p^2 = .125$; T2: *Wilks' Lambda* = .770, $F(8, 138) = 2.4$, $p = .018$, $\eta p^2 = .123$.

We followed up these results with MANOVAs comparing pairs of groups on the multivariate pattern of the executive function measures at each timepoint, controlling for age. At T1, L2 learners scored overall lower than both BilS children and NoL2 children, BilS: *Wilks' Lambda* = .768, $F(4, 46) = 3.3$, $p = .018$, $\eta p^2 = .232$; NoL2 children: *Wilks' Lambda* = .793, $F(4, 44) = 2.9$, $p = .034$, $\eta p^2 = .207$, while BilS children and NoL2 children did not significantly differ, *Wilks' Lambda* = .860, $F(4, 46) = 1.9$, $p = .131$, $\eta p^2 = .2140$. At T2, BilS children scored overall higher than NoL2 children: *Wilks' Lambda* = .691, $F(4, 46) = 5.1$, $p = .002$, $\eta p^2 = .309$, while L2 learners did not score significantly different from either group, BilS children: *Wilks' Lambda* = .882, $F(4, 44) = 1.5$, $p = .228$, $\eta p^2 = .118$; NoL2 children: *Wilks' Lambda* = .876, $F(4, 44) = 1.6$, $p = .202$, $\eta p^2 = .124$.

Figures 2 and 3 suggest that the groups' development in cognitive flexibility, measured by means of the DCCS, mirrors their development in creative fluency and flexibility, while the results of the ANT indices showed a very different pattern. Univariate analyses of the four measures confirm this impression (see Table 2 and Supplementary Material). It is striking that the ANT indices did not show any improvement over time. Only the orienting index showed a crossover Group x Time interaction, with no group differences at either time point. And BilS children had generally higher alerting scores than the other two groups.

We also tested how each group changed across the two timepoints. All groups significantly improved their EF skills, BILS: *Wilks' Lambda* = .187, $F(4, 21) = 22.9$, $p < .001$, $\eta^2 = .813$; L2 learners: *Wilks' Lambda* = .304, $F(4, 19) = 10.9$, $p < .001$, $\eta^2 = .696$; NoL2 children: *Wilks' Lambda* = .635, $F(4, 21) = 3.0$, $p = .041$, $\eta^2 = .365$.

In sum, the groups' development in EF was similar to that of their creative thinking development, but this was driven by the development in cognitive flexibility (DCCS).

Relationships between improvement in divergent thinking and EF measures

Since the groups differed in terms of both divergent thinking development and EF development, we conducted a mediation analysis to test whether children's creative thinking improvement might have been driven by EF improvements. For that, we created composite scores for both divergent thinking improvement and EF improvement for each participant by taking three steps. First, we standardized all scores and calculated improvements across the two timepoints for all measures. Second, we averaged these standardized improvement scores for each factor, that is a) improvement scores of fluency, flexibility, and originality to create a composite measure for divergent thinking and b) improvement scores of DCCS, Conflict index, Alerting index, and Orienting index to create a composite measure of EF. Third, we conducted a mediation analysis with Group as the independent variable, the composite divergent thinking development measure as dependent variable, and the composite EF development measure as mediator.

We next checked whether the following 4 conditions of mediation (Judd & Kenny, 1981) were met, basing the analysis on the reduced set of 76 participants (see explanation for ANT measures above). 1) The independent variable (Group) needed to significantly affect the outcome variable (critical thinking development) when the mediator (EF development) was not taken into account. 2) The independent variable (Group) had to significantly affect the

mediator (EF development). 3) The independent variable (Group) needed to affect the outcome variable (critical thinking development) to a lesser degree or not at all when the mediator (EF development) was controlled. 4) the mediator (EF development) had to affect the outcome variable (critical thinking development). While there were significant effects and trends for other relations, EF development did not predict critical thinking development, $b = 0.08$, $t(71) = 0.6$, $p = 0.554$, meaning condition (4) was not met. The same was the case for an exploratory mediation analysis that only took measures into account that had the same pattern (DCCS for EF development and creative fluency and flexibility for creative thinking development) and that was based on the full 99 participants, $b = 0.01$, $t(94) = 0.17$, $p = 0.867$. Thus, the data suggest that EF development did not mediate the effect of Group on divergent thinking development.

Discussion

We explored the impact of being exposed to a second language at school on divergent thinking skills, the effect of the amount of L2 exposure and whether any improvement in creative thinking could be explained by an improvement in EF. We found that exposure to a second language in formal settings seems to enhance **divergent thinking**. Also, while the three groups did not differ in **divergent thinking** at the first testing point, at the second testing point, bilingual school children outperformed children without a foreign language provision (control group). **Weekly language learners fell in between the two groups, not scoring significantly different from either group. A more fine-grained analyses of the three divergent thinking submeasures showed the same general pattern for fluency and flexibility as for the multivariate pattern as a whole, while the results for originality somewhat differed in that none of the groups significantly improved over time. Together, the data suggest that attending a bilingual school enhanced creative thinking. In contrast, they provide no firm evidence on**

the efficacy of weekly foreign language teaching, even though the trend in the data suggests an effect in the same direction, but of smaller magnitude.

Our findings are in line with a body of literature which found a bilingual advantage over monolinguals in creative fluency (Leikin, 2013; Leikin et al., 2014; Leikin et al., 2020) and flexibility (Adi-Japha et al., 2010; Carringer, 1974; Kharkhurin, 2017; Kharkhurin & Motalleebi, 2008; Leikin et al., 2014) as well as with studies which compared L2 learners at school to a group of children without a foreign language provision (Ghonsooly & Showqi, 2012; Landry, 1973a, 1973b, 1974). Unlike previous studies, though, we conducted a longitudinal study, checking divergent thinking ability at the start of foreign language exposure and carefully controlling for potential confounds. Thus, our results provide more controlled evidence that foreign language exposure in primary school seems to promote children's divergent thinking. As we observed advantages after only 6 months within the first year of formal education, our study also suggests that enhancement in divergent thinking can be achieved quite quickly. This stands in contrast to the findings by Landry (1973a, 1973b, 1974) who found a benefit only in higher grades of school and thus at more advanced stages of L2 acquisition. These differences might be due to our careful control of confounding factors.

Similar to the study by Leikin (2013), we found evidence of a benefit of a bilingual educational environment on creative originality. But the results for originality are less conclusive than those for fluency and flexibility. This is because originality scores for neither participant group significantly changed across the two testing points. Those for L2 learners and monolingual control children seemed to have rather decreased than increased, but the difference between testing points was very small. Notably, Leikin (2013) found a very similar pattern of results, using a different divergent thinking test. Just like in the present study, fluency and flexibility scores, but not originality scores, increased between the ages of ~46

and ~58 months, and this was the case for different types of children. In their case, these were monolingual children, children attending a bilingual kindergarten and children immersed in an L2 environment at kindergarten. And just like in our study, children attending a bilingual kindergarten scored higher on originality than monolingual children at the second testing point. The two studies together thus suggest that bilingual educational settings lead to higher creative originality, even though originality scores did not increase significantly. Instead, the difference between groups at time two appeared due to different trajectories of change in the three groups – something that requires confirmation in a sample tested over a longer period of time.

The question arises why originality develops differently from fluency and flexibility. The development of the latter two has been linked to children's increasing experiences and, related to that, their increasing longterm semantic and episodic memory (e.g., Bai et al., 2023). More experience and more associations in semantic memory can help children arrive at more solutions to a problem or task. While a similar argument could be made for originality – more experience can lead to more unique solutions - originality relies additionally on deep category exploration (Weiss & Wilhelm 2022; Nijstad et al., 2010). This is evidenced in the serial order effect of originality in divergent thinking tasks (Christensen et al., 1957): participants first respond with common ideas and produce increasingly original ideas with increasing number of responses. For instance, in our divergent thinking task, a child might have first turned a square into a simple house and later into a castle. Thus, they did not stop at an easy solution, but kept thinking beyond the prototype (i.e. a simple house). Due to the serial order effect, creative originality has been linked to persistence and effort (Weiss & Wilhelm 2022; Nijstad et al., 2010). Since originality is much more strongly related to effort than fluency and flexibility, it might be this additional aspect that causes a different development.

Our results suggest that one weekly session of L2 over six months is not sufficient to result in significant changes compared to no foreign language provision. However, it is important to note that weekly language learners did improve in both fluency and flexibility, while the control group did not. They also numerically fell in between the two other groups at T2, not differing as clearly from the bilingual school children as from the control group. A somewhat higher amount or a longer period of exposure might have potentially impacted their creative thinking more clearly.

We also tested whether any enhancement in creative thinking might have been due to enhancement in EF. **The mediation analyses suggested that this was not the case, despite the pattern of results for cognitive flexibility (measured by means of the DCCS) resembling that for creative fluency and flexibility.** Thus, while EFs might play a role in divergent thinking (Beaty, Silvia, Nusbaum, Jauk, & Benedek, 2014; Benedek, Franz, Heene, & Neubauer, 2012; Benedek, Jauk, Sommer, Arendasy, & Neubauer, 2014; Pan & Yu, 2018), our results suggest that enhanced EFs do not explain the bilingual advantage in creativity tasks as previously proposed (Hommel et al., 2011; Kharkhurin, 2011; Sampedro & Peña, 2019).

If improvements in divergent thinking due to foreign language exposure are not due to improvements in EF, what can be the mechanism? Bilingual enhanced divergent thinking ability has also been explained by richer and more elaborate conceptual and semantic networks (Adi-Japha et al., 2010; Kharkhurin, 2017). Kharkhurin (2017) showed that bilinguals' enhanced flexibility in creative production was linked to the fact that semantic networks in bilinguals connect words that are semantically unrelated for monolinguals. It is possible that our children learning a foreign language, especially those in bilingual education, had already enriched their networks of conceptual associations.

Alternatively, increased divergent thinking skills by children exposed to a foreign language might have been prompted not only by their encounter to a new language but also

by their exposure to a bicultural or multicultural environment. There is evidence that experience of culturally diverse environments (Gocłowska & Crisp, 2014) or cultural experience through living abroad (2021; Maddux & Galinsky, 2009) leads to creativity enhancements and that cultural diversity in school settings favours children's divergent thinking (review in Dunne, 2017; Vezzali, Gocłowska, Crisp, & Stathi, 2016). As we strictly limited the participation in our study to children from English monolingual households, it is very likely that especially bilingual schools, with staff and teachers that are native speakers of non-English languages, offered children a very first experience of a new culturally diverse context. This experience might have played a key role in their divergent thinking development.

Our study has both strengths and limitations. Its strengths are its longitudinal design and its control for various confounding variables. Not finding differences between the participant groups in terms of gender, parental background, SES, hours of extracurricular activities, computer use, number of siblings, English vocabulary and non-verbal IQ at the beginning of the studies ensured that any differences in creative thinking development was not due to participant differences. However, the fact that children were educated in different schools meant that their educational experiences (in addition to exposure to an L2) might not have been the same. Since all schools followed the national curriculum, core teaching is unlikely to have differed. But we do not know whether implementations of the curriculum differed or whether any additional enrichment activities might have affected their development of creative abilities. It is worth noting that any confounding effect of school differences would have somewhat be mitigated by the fact that each participant group was recruited from more than one schools. We recruited participants from 2 bilingual schools, 4 schools with L2 provision and 2 mainstream schools without L2 provision. It will be

important to control for any differences in curriculum implementation in future studies, especially those related to expressive art.

A further limitation of our study is that we did not test children's L2 proficiency at the end of the study. Given the potential link between vocabulary knowledge and creative thinking, it is possible that only those children who showed considerable L2 learning exhibited boosted creative thinking skills. It would be interesting to test in future studies in how far L2 language proficiency within a language learning group might be related to development in creative thinking. This would particularly be interesting to test in children who have limited L2 provision as in our weekly L2 group.

In conclusion, the results of our longitudinal study corroborate earlier findings of the impact of L2-learning in primary school settings on divergent thinking skills. Our findings clarify that significant advantages in creative thinking can be achieved in early stages of L2 learning, provided that children receive L2 instructions for a considerable amount of time. Our longitudinal design enabled us to investigate whether creative thinking changes were related to EF changes. We found that this was not the case. This speaks against an EF account of the bilingual creativity advantage. Instead, we speculate that bilingual school children might have developed an enriched lexical-semantic network with more elaborate semantic associations. This might derive from the acquisition of a new language and/or exposure to a new culturally diverse environment. Future studies should consider whether increasing the amount of L2 exposure in mainstream schools could positively impact children's divergent thinking skills.

References

- Adi-Japha, E., Berberich-Artzi, J., & Libnawi, A. (2010). Cognitive Flexibility in Drawings of Bilingual Children. *Child Development, 81*(5), 1356-1366. doi:10.1111/j.1467-8624.2010.01477.x
- Barac, R., Bialystok, E., Castro, D. C., & Sanchez, M. (2014). The Cognitive Development of Young Dual Language Learners: A Critical Review. *Early Childhood Research Quarterly, 29*(4), 699-714. doi:10.1016/j.ecresq.2014.02.003
- Beaty, R. E., Silvia, P. J., Nusbaum, E. C., Jauk, E., & Benedek, M. (2014). The roles of associative and executive processes in creative cognition. *Memory & Cognition, 42*(7), 1186-1197. doi:10.3758/s13421-014-0428-8
- Benedek, M., Franz, F., Heene, M., & Neubauer, A. C. (2012). Differential effects of cognitive inhibition and intelligence on creativity. *Personality and Individual Differences, 53*(4), 480-485. doi:10.1016/j.paid.2012.04.014
- Benedek, M., Jauk, E., Sommer, M., Arendasy, M., & Neubauer, A. C. (2014). Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity. *Intelligence, 46*, 73-83. doi:10.1016/j.intell.2014.05.007
- Booton, S. A., Hoicka, E., O'Grady, A. M., Chan, H. Y. N., & Murphy, V. A. (2021). Children's divergent thinking and bilingualism. *Thinking Skills and Creativity, 41*. doi:10.1016/j.tsc.2021.100918
- Carringer, D. C. (1974). Creative Thinking Abilities of Mexican Youth: The Relationship of Bilingualism. *Journal of Cross-Cultural Psychology, 5*(4), 492-504. doi:10.1177/002202217400500409
- Chen, X., He, J., & Fan, X. (2022). Relationships between openness to experience, cognitive flexibility, self-esteem, and creativity among bilingual college students in the U.S.

International Journal of Bilingual Education and Bilingualism, 25(1), 342-354.

doi:10.1080/13670050.2019.1688247

Creo, E. D., Mareque, M., & Pino-Juste, M. (2021). Deciphering the role of multilingualism in creativity at university: the influence of context. *Iral-International Review of Applied Linguistics in Language Teaching*. doi:10.1515/iral-2020-0170

Dunn, L. M., Dunn, D. M., Sewell, J., Styles, B., Brzyska, B., Shamsan, Y., & Burge, B. (2009). *British Picture Vocabulary Scale III*. London: GL Assessment Limited.

Dunne, C. (2017). Can Intercultural Experiences Foster Creativity? The Relevance, Theory and Evidence. *Journal of Intercultural Studies*, 38(2), 189-212.

doi:10.1080/07256868.2017.1291495

Edl, S., Benedek, M., Papousek, I., Weiss, E. M., & Fink, A. (2014). Creativity and the Stroop interference effect. *Personality and Individual Differences*, 69, 38-42.

doi:10.1016/j.paid.2014.05.009

Ghonsooly, B., & Showqi, S. (2012). The effects of foreign language learning on creativity. *English Language Teaching*, 5(4), 161-167.

Gocłowska, M. A., & Crisp, R. J. (2014). How Dual-Identity Processes Foster Creativity. *Review of General Psychology*, 18(3), 216-236. doi:10.1037/gpr0000008

Guildford, J. P. (1967). *The Nature of Human Intelligence*. Cambridge: Cambridge University Press.

Hansenne, M., & Legrand, J. (2012). Creativity, emotional intelligence, and school performance in children. *International Journal of Educational Research*, 53, 264-268.

doi:10.1016/j.ijer.2012.03.015

Hendrie Kupczynszyn, K., Arán Filippetti, V., & Oros, L. (2024). Socioeconomic status effects on children's creativity: The mediating role of executive functions. *Thinking*

Skills and Creativity, 51, 101437.

<https://doi.org/https://doi.org/10.1016/j.tsc.2023.101437>

Hocevar, D. (1979). A comparison of statistical infrequency and subjective judgment as criteria in the measurement of originality. *Journal of Personality Assessment*, 43(3), 297-299. doi:10.1207/s15327752jpa4303_13

Hommel, B., Colzato, L., Fischer, R., & Christoffels, I. (2011). Bilingualism and Creativity: Benefits in Convergent Thinking Come with Losses in Divergent Thinking. *Frontiers in Psychology*, 2(273). doi:10.3389/fpsyg.2011.00273

Judd, C. M., & Kenny, D. A. (1981). Process analysis: estimating mediation in treatment evaluations. *Evaluation Review*, 5(5), 602-619.
<https://doi.org/10.1177/0193841X8100500502>

Kharkhurin, A. V. (2011). The Role of Selective Attention in Bilingual Creativity. *Creativity Research Journal*, 23(3), 239-254. doi:10.1080/10400419.2011.595979

Kharkhurin, A. V. (2017). Language Mediated Concept Activation in Bilingual Memory Facilitates Cognitive Flexibility. *Frontiers in Psychology*, 8(1067).
doi:10.3389/fpsyg.2017.01067

Kharkhurin, A. V., & Motalleebi, S. N. S. (2008). The Impact of Culture on the Creative Potential of American, Russian, and Iranian College Students. *Creativity Research Journal*, 20(4), 404-411. doi:10.1080/10400410802391835

Kim, K. H. (2011). Proven Reliability and Validity of the Torrance Tests of Creative Thinking (TTCT). *Psychology of Aesthetics Creativity and the Arts*, 5(4), 314-315.
doi:10.1037/a0021916

Krumm, G., Filippetti, V. A., & Gutierrez, M. (2018). The contribution of executive functions to creativity in children: What is the role of crystallized and fluid

intelligence? *Thinking Skills and Creativity*, 29, 185-195.

<https://doi.org/10.1016/j.tsc.2018.07.006>

Landry, R. G. (1973a). The Enhancement of Figural Creativity Through Second Language Learning at the Elementary School Level. *Foreign Language Annals*, 7, 111-115.

Landry, R. G. (1973b). The relationship of second language learning and verbal creativity. *The Modern Language Journal*, 57(3), 110-113.

Landry, R. G. (1974). A Comparison of Second Language Learners and Monolinguals on Divergent Thinking Tasks at the Elementary School Level. *The Modern Language Journal*, 58(1/2), 10-15. doi:10.2307/323983

Lange, K. V., Hopman, E. W. M., Zemla, J. C., & Austerweil, J. L. (2020). Evidence against a relation between bilingualism and creativity. *PLoS One*, 15(6), e0234928. doi:10.1371/journal.pone.0234928

Lee, H., & Kim, K. H. (2010). Relationships Between Bilingualism and Adaptive Creative Style, Innovative Creative Style, and Creative Strengths Among Korean American Students. *Creativity Research Journal*, 22(4), 402-407. doi:10.1080/10400419.2010.523409

Lee, H., & Kim, K. H. (2011). Can speaking more languages enhance your creativity? Relationship between bilingualism and creative potential among Korean American students with multicultural link. *Personality and Individual Differences*, 50(8), 1186-1190. doi:10.1016/j.paid.2011.01.039

Leikin, M. (2013). The effect of bilingualism on creativity: Developmental and educational perspectives. *International Journal of Bilingualism*, 17(4), 431-447. doi:10.1177/1367006912438300

Leikin, M., & Tovli, E. (2014). Bilingualism and Creativity in Early Childhood. *Creativity Research Journal*, 26(4), 411-417. doi:10.1080/10400419.2014.961779

- Leikin, M., Tovli, E., & Malykh, S. (2014). Bilingualism and creative abilities in early childhood. *English Linguistics Research*, 3(2), 54-63. doi:10.5430/elr.v3n2p54
- Leikin, M., Tovli, E., & Woldo, A. (2020). The interplay of bilingualism, executive functions and creativity in problem solving among male university students. *Creative Studies*, 13(2), 308-324. Retrieved from <https://doi.org/10.3846/cs.2020.10397>
- Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, 25(5), 605-621. doi:10.1080/20445911.2013.795574
- Maddux, W., & Galinsky, A. (2009). Cultural Borders and Mental Barriers: The Relationship Between Living Abroad and Creativity. *J Pers Soc Psychol*, 96, 1047-1061. doi:10.1037/a0014861
- Pasarín-Lavín, T., Abín, A., García, T., & Rodríguez, C. (2023). Relationship between Executive Functions and Creativity in Children and Adolescents: A Systematic Review. *Children*, 10(6), 1002. <https://www.mdpi.com/2227-9067/10/6/1002>
- Pan, X., & Yu, H. H. (2018). Different Effects of Cognitive Shifting and Intelligence on Creativity. *Journal of Creative Behavior*, 52(3), 212-225. doi:10.1002/jocb.144
- Raven, J. C., Court, J. H., & Raven, J. C. (1990). *Manual for Raven's Progressive Matrices and Vocabulary Scales - Section 2: Coloured Progressive Matrices*. Oxford: Oxford Psychologists Press.
- Ricciardelli, L. A. (1992). Creativity and Bilingualism. *Journal of Creative Behavior*, 26(4), 242-254.
- Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, 42(8), 1029-1040. doi:10.1016/j.neuropsychologia.2003.12.012

- Sampedro, A., & Peña, J. (2019). Executive functions mediate the association between bilingualism and creativity in preadolescents. *Thinking Skills and Creativity*, 34, 100605. doi:10.1016/j.tsc.2019.100605
- Silvia, P. J. (2008). Creativity and Intelligence Revisited: A Latent Variable Analysis of Wallach and Kogan (1965). *Creativity Research Journal*, 20(1), 34-39. <https://doi.org/10.1080/10400410701841807>
- Sternberg, R. J. (1999). *Handbook of Creativity*. Cambridge: Cambridge University Press.
- Torrance, E. P. (1966). *The Torrance Test of Creative Thinking*. Lexington, MA: Personnel Press.
- van Dijk, M., Kroesbergen, E. H., Blom, E., & Leseman, P. P. M. (2019). Bilingualism and Creativity: Towards a Situated Cognition Approach. *Journal of Creative Behavior*, 53(2), 178-188. doi:10.1002/jocb.238
- Vezzali, L., Gocłowska, M. A., Crisp, R. J., & Stathi, S. (2016). On the relationship between cultural diversity and creativity in education: The moderating role of communal versus divisional mindset. *Thinking Skills and Creativity*, 21, 152-157. doi:10.1016/j.tsc.2016.07.001
- Zabelina, D. L., & Robinson, M. D. (2010). Creativity as Flexible Cognitive Control. *Psychology of Aesthetics Creativity and the Arts*, 4(3), 136-143. doi:10.1037/a0017379
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, 1(1), 297-301.

List of Tables

Table 1. Participant characteristics by groups

Table 2. Results of univariate analyses for all measures (Creative Thinking and EF)

Supplementary Table 1. Categories for flexibility scoring

List of Figures

Figure 1. Average creative fluency, flexibility and originality scores by groups at T1 and T2

Figure 2. DCCS scores by groups at T1 and T2

Figure 3. Mean conflict, orienting, and alerting indices of the ANT by groups at T1 and T2

Table 1. Participant characteristics by group.

	BilS (n=32)	L2 learners (n=29)	NoL2 (n=38)
<i>M</i> age in months (range)	56.0 (50-62)	57.2 (51-64)	59.5 (53-65)
Gender (f/m)	11/21	15/14	21/17
Maternal background UK ¹	96.9%	96.5%	100%
Paternal background UK ²	93.8%	100%	100%
<i>M</i> SES (0-1) (SD)	.78 (.19)	.71 (.21)	.76 (.16)
<i>M</i> hours of extra-curricular activities/week (SD)	1.2 (.3)	1.6 (.2)	1.4 (.2)
<i>M</i> hours of computer usage/week (SD)	2.2 (.5)	2.4 (.4)	2.6 (.6)
<i>Mdn</i> number of siblings (IQR)	1.0 (1-2)	1.0 (1-2)	1.0 (1-1.25)
<i>Mdn</i> number of older siblings (IQR)	.5 (0-.5)	1.0 (1-1.5)	1.0 (0-1)
<i>M</i> BPVS score (SD)	74.1 (12.6)	70.9 (13.7)	72.6 (14.0)
<i>M</i> IQ (Raven's)	15.4 (3.3)	14.8 (3.7)	15.1 (3.6)

Note.

1. Non-U.K. places of birth for mothers were USA (n=1) in the BilS group and Ireland (n=1) in the WL2 group.

2. Non-U.K. places of birth for fathers were New Zealand (n=1) and South Africa (n=1) in the BilS group.

Table 2. Results of univariate analyses for all measures (Creative Thinking and EF).

Measure	Time		Group		Time x Group	
	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
Creative Thinking						
Fluency	32.6	<.001*	.7	.515	5.4	.006*
T1			.3	.753		
T2			3.1	.049*		
Flexibility	29.4	<.001*	7.0	.001*	6.3	.003*
T1			.3	.721		
T2			9.6	<.001*		
Originality	.9	.334	3.1	.048*	2.6	.083(*)
T1			.8	.474		
T2			4.0	.022*		
Executive Functions (EF)						
DCCS	101.5	<.001*	2.0	.142	7.8	<.001*
T1			.3	.758		
T2			9.1	<.001*		
Conflict index	.003	.960	2.0	.139	1.2	.317
Orienting index	.008	.930	.184	.832	3.58	.033*
T1			2.22	.116		
T2			1.3	.270		
Alerting index	.004	.952	4.0	.023*	1.8	.173

Note. * marks significance at 0.05 level, (*) a trend for significance. Comparisons between the three groups for significant effects of Group are reported in the Supplementary Material.

Supplementary Table 1. Categories for flexibility scoring

Categories identified at T1	Categories identified at T2
Buildings (e.g., schools, churches, castles)	Buildings
Boxes and containers	Boxes and containers
Nests/homes/shelters for animals/people	Lines used as roads/paths/bases (with someone walking/standing on them or a vehicle driving on it)
Squares used as body of animals/people	Lines used as the body for animals/people/monsters
Squares used as fences/compounds/paddocks to delimitate an area (e.g. gardens, parks)	Electronic devices (e.g., Phones/tablets/laptops)
Stationary (e.g., cards, photos, papers, pictures)	Stationary
Furniture (e.g., chairs, tables, cupboards)	Furniture
Playground equipment (e.g., slides, swings)	
Vehicles (e.g., cars, lorries, vans)	Vehicles
Squares used as faces of animals/people/monsters	

Figure 1. Average creative fluency, flexibility and originality scores for the three participant groups (BilS = bilingual schools, WL2 = weekly second language learners, NoL2 = no second language provision) and testing points (T1 and T2). Error bars represent 95% confidence intervals.

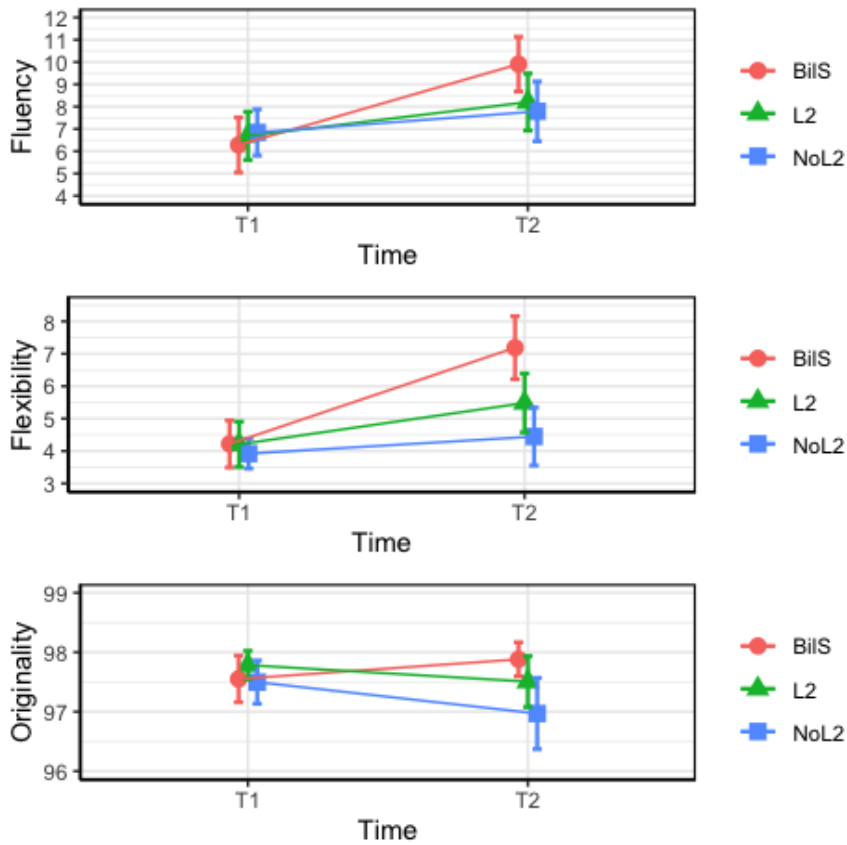


Figure 2. DCCS scores for the three participant groups (BiS = bilingual school children, L2 = L2 learners, NoL2 = children without L2 provision) at T1 and T2. Error bars represent 95% confidence intervals.

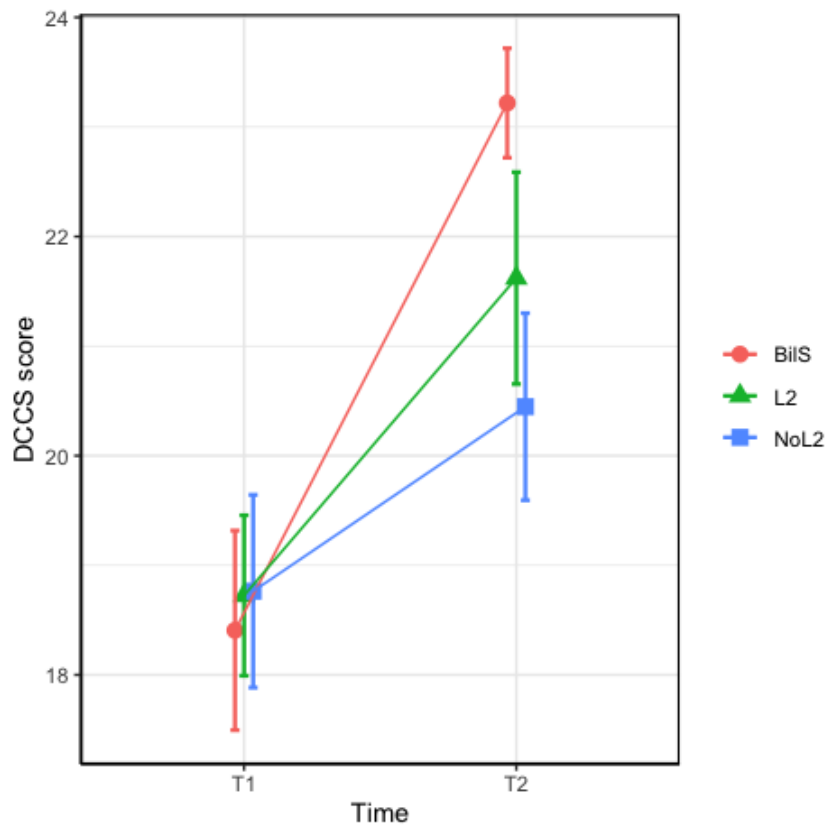
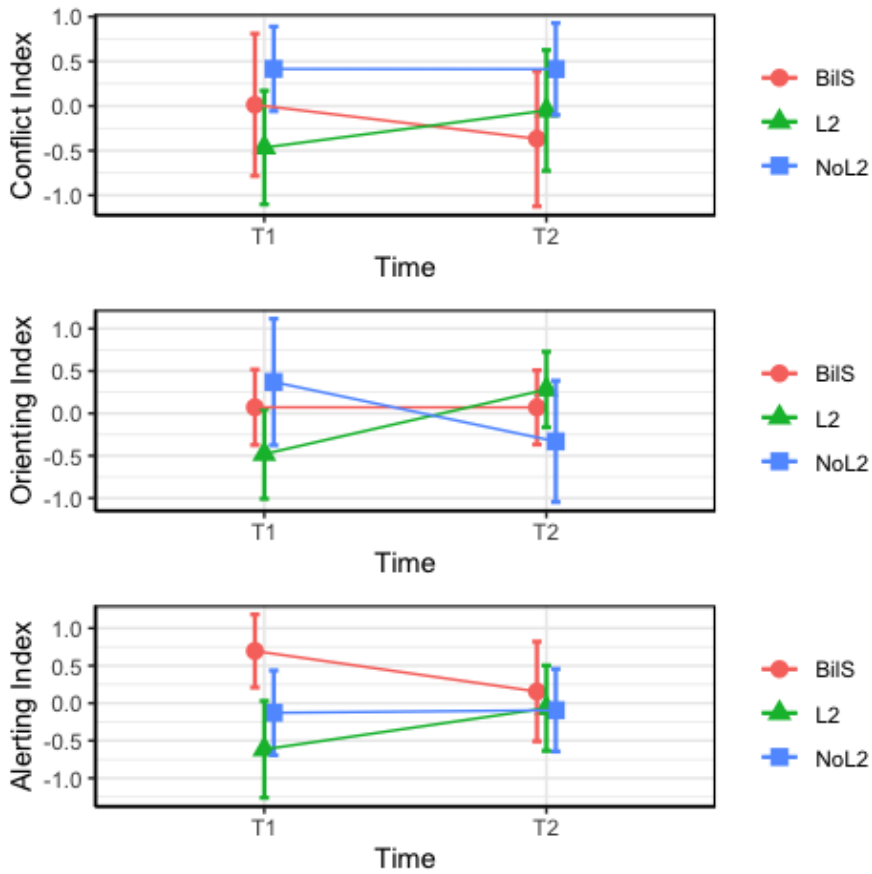


Figure 3. Mean conflict, orienting, and alerting indices of the ANT for the three participant groups (BiS = bilingual school children, L2 = L2 learners, NoL2 = children without L2 provision) at both testing points (T1 and T2). Error bars represent 95% confidence intervals.



Supplementary Material

Divergent thinking measures

Figure 1 shows the development of the three participant groups in terms of the three submeasures of divergent thinking.

Fluency

The results for a univariate analysis for fluency were in line with the MANOVA reported in the main text. We found a significant main effect of Time, $F(1, 96) = 32.6, p < .001, \eta p^2 = .253$, no effect of Group, $F(2, 96) = .7, p = .515, \eta p^2 = .014$, but a significant interaction between Time and Group, $F(2, 96) = 5.4, p = .006, \eta p^2 = .101$. Follow-up analyses revealed that at T1 the difference between the groups was not significant, $F(2, 96) = .3, p = .753, \eta p^2 = .006$, but at T2 it was, $F(2, 96) = 3.1, p = .049, \eta p^2 = .061$. Post-hoc comparisons for T2 revealed that BilS children had significantly higher fluency scores than NoL2 children, $p = .018$, and marginally higher fluency scores than L2 learners, $p = .074$, while the fluency scores of L2 learners and NoL2 children did not differ, $p = .645$.

Furthermore, fluency significantly improved from T1 to T2 for BilS children by an average of 3.6 (SD 3.8), $F(1, 31) = 29.9, p < .001$, and for L2 learners by an average of 1.5 (SD 3.4), $F(1, 28) = 5.6, p = .025$. In contrast, the average fluency improvement of 0.9 (SD 3.4) for NoL2 children was not significant, $F(1, 37) = 3.0, p = .091$.

Flexibility

Results for a univariate analysis of creative flexibility again followed the general pattern of the MANOVA, albeit with a clearer difference between BilS and L2 learners. We found main effects of Time, $F(1, 96) = 29.4, p < .001, \eta p^2 = .234$, and of Group, $F(2, 96) = 7.0, p = .001, \eta p^2 = .128$, as well as a significant interaction between Time and Group, $F(2,$

96) = 6.3, $p = .003$, $\eta p^2 = .116$. Follow-up analyses showed again no effect of Group at T1, $F(2, 99) = .3$, $p = .721$, $\eta p^2 = .007$, but at T2, $F(2, 99) = 9.6$, $p < .001$, $\eta p^2 = .167$.

Comparisons for T2 showed that BilSs had significantly higher flexibility scores than both L2 learners, $p = .013$, and NoL2s, $p < .001$, while the scores of L2 learners and NoL2 children did not differ, $p = .112$. And similar to fluency, flexibility improved from T1 to T2 for BilS children by an average of 3.0 (SD 3.1), $F(1, 31) = 28.7$, $p < .001$, and for L2 learners by an average of 1.3 (SD 2.4), $F(1, 28) = 8.0$, $p = .008$. Again, the average improvement of .5 (SD 3.0) for NoL2 children was not significant, $F(1, 37) = 1.1$, two-sided $p = .291$.

Originality

The results of the univariate analysis of originality scores were somewhat different from those for fluency and flexibility measures. We found no main effect of Time, $F(1, 96) = .9$, $p = .334$, $\eta p^2 = .010$, but of Group, $F(2, 96) = 3.1$, $p = .048$, $\eta p^2 = .061$, as well as a trend for an interaction between Time and Group, $F(2, 96) = 2.6$, $p = .083$, $\eta p^2 = .051$. Follow-up analyses showed no effect of Group at T1, $F(2, 98) = .8$, $p = .474$, $\eta p^2 = .015$, but at T2, $F(2, 96) = 4.0$, $p = .022$, $\eta p^2 = .077$. BilS children outperformed NoL2 children at T2, $p = .006$, while the scores of L2 learners fell in between those of the other two groups but did not significantly differ from either BilS children, $p = .284$, or NoL2s, $p = .113$. Notably, compared to fluency and flexibility scores, originality scores did not significantly change over time. Comparing values over time shows that there was no significant change for any of the groups, BilS: $F(1, 31) = 2.2$; $p = .152$; L2: $F(1, 28) = 1.4$; $p = .254$; NoL2s: $F(1, 37) = 2.6$; $p = .117$.

In sum, the groups did not differ at any measure of divergent thinking at T1. Across the two time points, only BilS children and L2 learners improved and only in terms of fluency and flexibility. This meant that the BilS children outperformed NoL2 children in

terms of fluency and flexibility. L2 learners fell in between the two groups. In terms of flexibility they were not significant from either group. In terms of fluency, they scored significantly lower than BilS children and more similarly to NoL2 children.

Executive Function measures

Dimensional Change Card Sort task (DCCS)

Figure 2 shows the result of the DCCS. We found significant main effects of Time, $F(1, 73) = 101.5, p < .001, \eta p^2 = .582$, and Group, $F(2, 73) = 2.0, p = .142, \eta p^2 = .052$, as well as a Time x Group interaction, $F(2, 73) = 7.8, p < .001, \eta p^2 = .175$. Groups did not differ at T1, $F(2, 73) = .3, p = .758, \eta p^2 = .008$, but they did so at T2, $F(2, 73) = 9.1, p < .001, \eta p^2 = .199$. Post-hoc tests for T2 showed that BilS children had higher DCCS scores compared to L2 learners, $p = .030$, and NoL2 children, $p < .001$, while L2 learners scored marginally higher than NoL2 learners, $p = .054$. Furthermore, DCCS scores improved from T1 to T2 for all groups, BilS children: average improvement of 4.8 (SD 2.3), $t(31) = 12.0$, two-sided $p < .001$, L2 learners: average improvement of 2.9 (SD 2.9), $t(28) = 5.4$, two-sided $p < .001$, NoL2 children: average improvement of 1.7 (SD 3.0), $t(37) = 3.4$, two-sided $p = .002$.²

Attentional Network Task (ANT)

As indicated, the ANT provides measures for conflict, alerting and orienting. Figure 3 shows the results of the three measures. As evident in the following tests, the results for the ANT were very different from the other tests. First, children did not improve from T1 to T2. Second, BilS children scored higher on the alerting index.

Conflict

² Note that an analysis with all 99 participants showed the same result pattern.

For the conflict index, there was no significant effect of Time, $F(1, 73) = .003, p = .960, \eta p^2 < .001$, or Group, $F(2, 73) = 2.0, p = .139, \eta p^2 = .053$, or a Time x Group interaction, $F(2, 73) = 1.2, p = .317, \eta p^2 = .031$.

Orienting

For the orienting index, we found again no effect of Time, $F(1, 73) = .008, p = .930, \eta p^2 < .001$, or of Group, $F(2, 73) = .184, p = .832, \eta p^2 = .005$, but a significant Time x Group interaction, $F(2, 73) = 3.58, p = .033, \eta p^2 = .090$. However, groups did not differ at T1, $F(2, 75) = 2.22, p = .116, \eta p^2 = .057$, or at T2, $F(2, 73) = 1.3, p = .270$. Furthermore, only L2 learners showed an improvement in orienting from T1 to T2, while other groups did not show any significant change, BilS children: $t(25) < -.01$, two-sided $p = .994$, L2 learners: $t(23) = 2.1$, two-sided $p = .043$, NoL2 children: $t(25) = -1.4$, two-sided $p = .185$.

Alerting

For the alerting index, there was no significant effect of Time, $F(1, 73) = .004, p = .952, \eta p^2 < .001$, or a Time x Group interaction, $F(2, 73) = 1.8, p = .173, \eta p^2 = .047$, but a significant main effect of Group, $F(2, 73) = 4.0, p = .023, \eta p^2 = .098$. Posthoc tests (Bonferroni corrected) revealed that the alerting index was significantly higher for BilS children than for L2 learners, $p = .008$, and marginally higher than for NoL2 children, $p = 0.055$. L2 learners and NoL2 children did not differ from each other, $p = .416$.