Thinking Skills and Creativity 41 (2021) 100918

Contents lists available at ScienceDirect

Thinking Skills and Creativity

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

Children's divergent thinking and bilingualism

Sophie A. Booton[a,*], Elena Hoicka[b], Aneyn M. O'Grady[a], Hiu Ying Nicole Chan[a], Victoria A. Murphy[a]

[a] Department of Education, University of Oxford, 15 Norham Gardens, Oxford, OX2 6PY, United States
[b] School of Education, University of Bristol, 35 Berkeley Square, Bristol BS8 1JA, United States

ARTICLE INFO

Keywords:
Creativity
Divergent thinking
Bilingualism
Cognitive development

ABSTRACT

Most children in the world grow up bilingual, and bilingualism has been linked to a range of linguistic and other cognitive skills. One such skill is creativity, which is thought could be increased in bilinguals due to enhanced executive functions or more diverse cultural experiences. However, extant literature with children has produced mixed results, perhaps due to methodological limitations. In this study, bilingual and monolingual children ($N = 111$, 60% bilingual) sampled from the same British schools completed three measures of divergent thinking, alongside measures of nonverbal intelligence, vocabulary, and exposure to English. No differences were found between monolingual and bilingual children across any of the divergent thinking tasks or measures, either before or after controlling for possible confounds, and effect sizes were negligible to small. This well-powered, pre-registered study provides no evidence for a bilingual advantage in the divergent thinking component of creativity amongst children, suggesting that previous mixed results may have been due to a high prevalence of false positives. Thus, while bilingualism has many benefits for children, divergent thinking is not one of them.

1. Introduction

Many, if not most, children around the world are growing up bilingual (Department for Education, 2020; Qi, 2016; Ryan, 2013; UNESCO, 2016). Acquiring more than one language to a degree of fluency, and regularly using and switching between these multiple languages brings several direct benefits for children, including improved language skills across both first and second languages (Murphy et al., 2015; Yelland et al., 1993), the ability to communicate with more people (Eberhard, Simons, & Fennig, 2021), and potential for greater vocational and economic prospects (Agirdag, 2014). Researchers have sought to identify cognitive benefits of bilingualism beyond the domain of language, and bilingualism is claimed to impact a broad range of cognitive skills in adults, including learning, memory, metacognition, theory of mind, and attention (Adesope et al., 2010; Barac et al., 2014; Escudero, Mulak, Fu, & Singh, 2016). Therefore, it is critical that we understand the impact of the bilingual experience on children's broader cognitive development too.

The impact of bilingualism on creativity has attracted significant research attention. Creativity is defined as the ability of an individual or group to generate ideas or products that are both novel and useful within a social context (Amabile & Pratt, 2016) and is considered a critical 21st century skill for children to develop to successfully participate in contemporary societies and economies (Ananiadou & Claro, 2009). For children, examples of creativity could include making novel and meaningful personal interpretations

* Corresponding author at: University of Oxford, Department of Education, 15 Norham Gardens, Oxford, Oxfordshire OX2 6PY, United Kingdom.
E-mail address: sophie.booton@education.ox.ac.uk (S.A. Booton).

https://doi.org/10.1016/j.tsc.2021.100918
Received 24 May 2021; Received in revised form 2 August 2021; Accepted 3 August 2021
Available online 8 August 2021

during learning, such as thinking of their own ideas for why Pluto should or should not be considered a planet, or participating in everyday creative activities, such as painting a picture or inventing a joke (Kaufman & Beghetto, 2009). Creativity involves the processes of both divergent thinking, that is the capacity to generate multiple solutions to a problem, and convergent thinking, or the process of selecting ideas to produce a correct or preferred solution (Guilford, 1967).

Children's creativity has been assessed in several ways (Plucker & Makel, 2019), including divergent thinking tests (Guilford, 1967; Torrance, 1966), insight tasks (Duncker & Lees, 1945; Lee et al., 2014; Mednick, 1962), and consensual assessment of creative products (Amabile, 1982). The most commonly used of these types of tests are divergent thinking measures (Plucker & Makel, 2019), such as the Torrance Test of Creative Thinking or TTCT (Kim, 2011; Torrance, 1966). Such measures require participants to generate multiple ideas in response to a prompt, which can involve ideational tasks (e.g., generating uses for an object, Guilford, 1956), figural tasks (such as making pictures by completing figures, Torrance, 1966), linguistic tasks (e.g., generating definitions of polysemous words, Jacobs & Pierce, 1966), or motor tasks (e.g., demonstrating different ways to move across a room, Torrance, 1981). Scoring produces a number of outcome variables, typically consisting of evaluating the number of valid responses (fluency), the number of different categories of response generated (flexibility), the rareness of each response amongst the sample or a standard population (originality), and the degree of detail of each response (elaboration). However, other specific features have also been assessed, such as abstractness of titles (producing expressive rather than literal labelling titles for figures), resistance to premature closure (using less direct means to close an incomplete figure), and a creative strengths score based on a number of other traits such as humour, emotional expressiveness, and richness of imagery (Torrance & Ball, 1984). Divergent thinking tasks have been found to be reliable and valid across a number of studies with both adults and children (Kim, 2006; Runco et al., 2010; Zachopoulou et al., 2009).

With regard to how bilingualism may affect creativity, two specific hypotheses have been put forward. Firstly, via executive functions (Sampedro & Peña, 2019): being bilingual has been associated with improved executive functions developed through switching between languages (Bialystok, 2015), and in turn executive functions have been correlated with divergent thinking skills, which are thought to help to inhibit common responses and switch categories more easily (Benedek et al., 2012; Kubota et al., 2020; Lee & Therriault, 2013). However, recent research has also challenged the link between bilingualism and executive function (Lehtonen et al., 2018; Nichols et al., 2020) and executive function and some aspects of creativity (Beck et al., 2016). A second explanation for the putative link between bilingualism and creativity is the cultural diversity hypothesis (e.g., Kharkhurin, 2008). This account proposes that bilinguals may be more creative because they have more diverse life experiences (Ritter et al., 2012) such as moving between countries and participating in multiple cultures. Indeed, in research on bilingualism and creativity, samples are often from immigrant populations (Kharkhurin, 2008) and in countries like the UK, most bilingual children speak languages which originate from other countries (Department for Education, 2012). Research suggests that multicultural experiences enhance creativity using several indices, including divergent thinking, for adults (Leung & Chiu, 2010; Leung et al., 2008) and adolescents (Chang et al., 2014). Thus, it may be differences in cultural experiences that drive potential differences in creativity between monolingual and bilingual individuals.

However, the evidence pertaining to the impact of bilingualism on creativity with children is far from consistent across studies. Existing studies have generated conflicting findings regarding the impact of bilingualism on children's creativity, with most using divergent thinking measures. Some studies have suggested that bilingual children perform better than monolingual children in terms of verbal creativity (Leikin, 2013; Leikin & Tovli, 2014), whilst others have found no effect (Jacobs & Pierce, 1966) or suggested a bilingual advantage only in figural and not verbal measures (Okoh, 1980; Ricciardelli, 1992; Sampedro & Peña, 2019). Other studies found different results depending not only on the general mode of presentation (verbal/figural), but also the specific task domain (e.g. mathematical vs. practical problem-solving) (Leikin, 2013). In addition, most studies have found effects of bilingualism on some combinations of specific outcome measures but not others from the same task (Carringer, 1974; Lee & Kim, 2011; Leikin, 2013; Ricciardelli, 1992; Torrance et al., 1970), such as originality but not fluency or flexibility (Leikin, 2013); imagination score but not fluency (Ricciardelli, 1992); or title abstractness and creative strengths, but not fluency, elaboration, originality, or resistance to premature closure (Lee & Kim, 2011). In one study for example, bilingual children performed poorer than monolinguals on a figural creativity task in terms of fluency and flexibility, but better than monolinguals in terms of elaboration (Torrance et al., 1970).

The mixed findings within and between studies could be explained by several factors. Firstly, most studies have small samples with some having as few as 10 to 15 participants per subgroup (Leikin, 2013; Leikin & Tovli, 2014; Ricciardelli, 1992), which do not make for reliable comparisons (Button et al., 2013). Secondly, because divergent thinking tasks generate multiple scores (e.g., fluency, flexibility, and originality), there is a risk of false positive results where studies have not *a priori* limited their variables of interest and/or controlled for multiple comparisons (Wicherts et al., 2016). For example, Leikin, (2013) calculated fluency, flexibility and originality scores for two tasks (pictorial multiple solution and creating an equal number task) and found an effect of bilingualism on only one of these six measures (originality on the creating an equal number task).

Furthermore, existing research does little to rule out important confounds that could explain differences between groups. Very few studies have controlled for other potentially confounding variables between the language groups compared, such as intelligence. Intelligence, while distinct from creativity, is known to predict creativity scores (Kim, 2005). Intelligence could feasibly differ between language groups in many studies, especially where the children in different language groups are sampled from different schools or areas (Jacobs & Pierce, 1966; Leikin, 2013; Torrance et al., 1970). Furthermore, many studies compare proficient, balanced bilinguals - that is those who have good proficiency in both their first and second languages – to monolingual or lower proficiency bilingual groups (Carringer, 1974; Lee & Kim, 2011; Leikin & Tovli, 2014; Sampedro & Peña, 2019). Balanced bilinguals have been shown to have superior language skills and may have superior general cognitive skills than other bilingual groups (Woumans, Ameloot, Keuleers, & Van Assche, 2019). One exception to this was a study which found that bilinguals performed better than monolinguals in terms of verbal but not figural originality and that this effect was maintained after controlling for verbal intelligence (Okoh, 1980). However, the intelligence measures used differed between groups and the full statistics for the analysis were not reported.

Given the mixed evidence on this question, the methodological limitations of previous studies, and the recognized importance of replication of findings within psychology (Open Science Collaboration, 2015), the aim of this study was to examine whether bilingual and monolingual children differ in terms of their creativity by implementing a more robust study design. By including a larger sample, pre-registering our analyses, and controlling for potential confounds, the present study will allow us to better distinguish whether the effect of bilingualism on divergent thinking in children is reliable, rather than a product of other differences between groups (e.g., in intelligence). Thus, in this pre-registered study, 111 British monolingual and bilingual children recruited from the same schools completed three measures of divergent thinking: a word meanings task, an alternate uses task, and a drawing task. Participants also completed measures of their English vocabulary and non-verbal intelligence. The research question was whether bilingual and monolingual children differed in terms of their creative performance and whether these differences remained after controlling for their verbal and non-verbal intelligence.

2. Method

2.1. Design

A mixed design was used, with a between-subjects factor of language group (bilingual or monolingual) and within-subjects factors of divergent thinking measures. All participants completed all divergent thinking tasks in the same fixed order. The methods and analyses were pre-registered on the Open Science Framework (https://osf.io/swq49/) and any exploratory deviations from this are noted.

2.2. Participants

Participants were 111 children (59 female), recruited through five state schools in southern England. One further child could not complete any of the creativity tests because they could not follow the instructions (their English vocabulary was very low - BPVS raw score of 12). Post-hoc power analysis shows that with a minimum group size of 46, the power to detect the large effect size ($F = 0.4$) shown in previous studies was 97% and a medium effect size ($F = 0.3$) was 81%.

Children were from school year groups 1 and 4 (aged 5 to 6 years and 8 to 9 years). Details of the participants in each group are shown in Table 1. All children were English speakers and used English daily at school. Bilingual children (n=67) spoke English and another language more than once a week. The bilingual children used a variety of languages: 18 different languages were used in total, with the most common being Polish (n=12), Punjabi (n=8), Tetum (n=5), Malayalam (n=5), Arabic (n=4), and Nepalese (n=4). In terms of language dominance, the bilingual children were on average balanced with a slight bias towards English (with a score of 9 being balanced, see Language Background section below; $M = 8.20$, $SD = 3.21$, $Min = 1.00$, $Max = 15.00$). The monolingual and bilingual groups did not differ in age, gender, or non-verbal intelligence (see Table 1). They did differ in English vocabulary and years learning English, in favour of the monolingual group.

2.3. Measures

2.3.1. Divergent Thinking

Children's divergent thinking was measured using three tasks.

2.3.1.1. Linguistic Task: Word Meanings Test. This test measures children's ability to produce many different senses of a word and is adapted from the Word Meanings Test in work which examined children's linguistic creativity (Jacobs & Pierce, 1966; Okoh, 1980). In this task, children were given a word which has many senses, and asked to provide as many different uses of that word as possible. The five stimuli, presented in a fixed order, were: set, head, hand, turn, and cross. An example was given (roll) and a practice item with feedback (pound) prior to the target items. Children were given 45 seconds for each word. The scores for this test were fluency (the number of responses generated per item, excluding repetitions and unclassifiable responses), flexibility (the number of different senses generated per item), and originality (the reciprocal of the average proportional frequency of response categories). Higher scores indicate more creative responses.

The instructions participants were given were as follows: "*I'm going to tell you a word. I want you to tell me as many different meanings*

Table 1
Details of participants in the sample.

	Monolingual	Bilingual	t / χ^2	*df*	*p*
N	44	67			
Age (SD)	7.21 (1.52)	7.65 (1.50)	-1.52	109	.132
Female (%)	23 (52%)	30 (45%)	0.60	1,111	.439
Non-verbal IQ raw (SD)	11.32 (3.80)	12.93 (4.67)	-1.90	109	.060
English vocabulary raw (SD)	99.80 (22.20)	93.04 (22.88)	1.58	109	.118
English vocabulary standard (SD)	97.28 (13.11)	87.45 (13.00)	3.88	109	<.001
Years of English (SD)	7.10 (1.63)	5.39 (2.49)	3.57	109	.001

for the word as you can." They were given the following example: *"For example, if I said the word 'roll' you could say roll means a bread roll; to roll a ball; to do a forward roll; rock and roll music".* Children then completed a practice item and received feedback on their responses. They then completed the 5 items. If the child didn't respond for 10 seconds, they were given a neutral prompt: "What else can [word] mean?"

Definitions of each of the 5 target words were sourced from WordNet (Princeton University, 2010). Some highly similar definitions were grouped (e.g., the verb and adjective forms of set as in to harden or be hardened) to reduce ambiguity when coding. In addition to the valid definitions, responses could be coded as Invalid (not a valid definition, 8.3%), Ambiguous (ambiguous between 2 valid meanings, 2.1%), or Other (another meaning not listed, 0.1%). All responses were coded by an independent rater who was blind to the language status of the participants and hypotheses. Another rater coded 10% of the data and inter-rater agreement was substantial (kappa=0.845). Responses rated as Other were checked and a new definition added if required.

2.3.1.2. Figural task: Circles test. The circles task is a figural measure of divergent thinking (Beck et al., 2016) based on the Torrance Test of Creative Thinking (Torrance, 1966). In this task, children were presented with an A4 sheet of paper containing 24 line-drawings of circles. The children were instructed to use the circles to make as many different drawings as possible. They were advised that drawings did not need to be neat and that they had limited time to complete the task. The first two circles were completed as examples, having been made into a pig and a clock. Children had 2 minutes to complete the task, and if they did not draw for a 10 second period, they were given a neutral prompt: "Can you do any more drawings?" As the child finished each drawing, the experimenter asked: "What is that picture of?" and wrote down the child's response. The scores for this test were fluency (the number of non-repeat drawings completed which did not copy the examples), flexibility (the number of categories drawings fit into), originality (the reciprocal of the average proportional frequency of response categories). Higher scores indicate more creative responses.

Coding of responses followed the procedures reported in Beck et al. (2016). First, a coding scheme for responses was created. Two independent coders looked at all drawing labels with a frequency of 2 or more and created categories based on these labels. For the 51 categories initially generated, agreement was 55%. Through discussion, a consensus was reached on 26 categories (for example, animals, faces, food etc.). After creating the coding scheme, all responses were coded by two independent raters. Raters were blind to the language status of the participants. The raters had access to children's drawings to inform their categorisation. Inter-rater agreement was near perfect (kappa=0.957). Discrepancies were resolved through discussion.

2.3.1.3. Ideational task: Object uses test. The object uses task (Beck et al., 2016) is a verbal measure of divergent thinking adapted from the Alternative Uses Test (Guilford, 1967) . In this task, children were instructed to generate as many different uses for a common object as they could. Children were presented with two trials. In each trial, they saw a photographic image of a common object (either a blanket or a brick). They were given 1 minute to generate uses for the object, and if they did not provide a response for a 10 second period, they were given a neutral prompt: "What else could you do with it?". The scores for this test were fluency (the number of non-repeat suggestions generated, excluding purely descriptive comments, per item), flexibility (the number of categories suggestions fit into, per item), and originality (the reciprocal of the average proportional frequency of response categories). Higher scores thus indicate more creative responses.

Coding of responses followed the procedures presented in Beck et al. (2016). First, two coding schemes were created, for responses on each of the two trials (blanket and brick). Two independent coders looked at all responses with a frequency of 2 or more and created categories based on these labels. For the categories initially generated (29 for brick and 39 for blanket), agreement was 71% for brick and 77% for blanket. Through discussion, a consensus was reached on 20 categories for brick (for example, building, stacking, lifting) and 23 categories for blanket (for example, sleeping, picnics, folding). After creating the coding schemes, all responses were coded by two independent raters. Raters were blind to the language status of the participants. Inter-rater agreement was substantial (kappa=0.831). All discrepancies were resolved through discussion.

2.3.2. Child language background

To determine the language background of the participants, a short questionnaire was completed by their parents. The questionnaire contained six questions, two of which were relevant to the present data. The first covered the child's first language(s), and other language(s) and the age at which the child began learning English. The second question asked about the frequency of bilingual children's use of their other language on a 5-point scale from 'Daily' to 'Less than once a month'.

Bilingual children also self-reported on the dominance of their use of English versus their home language. They were asked how often they spoke or were spoken to in English or their home language with parents; other family members; and friends. The scale for each question was from 0 (completely in English) to 3 (completely in home language), and thus a total score ran from 0 (English completely dominant) through 9 (balanced) to 18 (home language completely dominant).

2.3.3. English vocabulary

To measure receptive vocabulary, the British Picture Vocabulary Scale version 3 (BPVS3; Dunn & Dunn, 2009) was used. In this task, children hear a word and are required to point to a picture in an array of four that depicts the meaning of the word. Raw scores, which reflect the number of items completed or likely to be completed correctly, were taken as the key outcome measure.

2.3.4. Non-verbal intelligence

To assess non-verbal intelligence, the Matrix Reasoning subtest from the Weschler Intelligence Scale for Children, fifth edition

(WISC-V; Weschler, 2014) was used. In this task, children see an array (a sequence or matrix) of shapes with one element missing and select the shape that matches the array from five response options. Raw scores (i.e. the number of items answered correctly) were taken as the outcome measure, as the test did not have standardised scores for children aged 5 years.

2.4. Procedure

The data was collected as part of a wider study on multilingual children's cognitive capacities (Booton et al., 2021). Testing was conducted in a quiet area in the child's school. After receiving parent consent and child assent, children completed a battery of tasks across three sessions, each taking 20 to 30 minutes. The BPVS was completed at the end of the first session; the WISC matrix reading subtest in the middle of the third session; and the divergent thinking tasks at the end of the third session, with alternate uses first, followed by the circles test, then the word meanings test. All tasks were completed in English.

3. Results

3.1. Data analysis and cleaning

Missing data were as follows: one monolingual participant was missing data on all 3 measures of the Word Meanings Test (due to experimenter error); two participants (one monolingual, one bilingual) were missing data for originality on the circles task (because they only produced drawings that were copies of the examples); six bilingual participants were missing data on their years learning English and frequency of language use (due to lack of parent response); and two bilingual participants were missing data on self-reported language dominance (due to experimenter error). In total, 0.8% of data values were missing (15.3% cases had some data missing). Missing data was imputed using multiple imputation (Hayati Rezvan, Lee, & Simpson, 2015) with SPSS version 26. All reported numeric demographic and test outcome variables were included in the imputation. Predictive mean matching was used with 10 imputations. The key analyses were run with and without imputation and the pattern of results was identical, so imputed results are reported. Some variables were skewed: positively skewed variables were log transformed (fluency scores for all tests, flexibility for object uses and circles tests) and negatively skewed variables were cubed (originality in object uses) or squared (years of English). Raw scores are reported in descriptives.

The analyses will be reported in line with those that were pre-registered. The pre-registered analyses comprised six ANOVAs of language status (monolingual/bilingual) for flexibility and originality for figural, ideational and linguistic divergent thinking.

3.2. Descriptive statistics

Descriptive statistics for divergent thinking test performance for bilingual and monolingual children are show in Table 2.

Correlations between divergent thinking test performance, non-verbal intelligence, English vocabulary, and age are shown in Table 3. Correlations between measures within the same creativity task were positive and significant after partialling out age (except for between word meanings fluency and originality). Correlations between different creativity tasks were positive and mostly significant but generally smaller after partialling out age. Vocabulary was significantly positively correlated with most creativity measures, as was non-verbal intelligence, but the latter became non-significant after partialling out age.

The key analyses were six ANOVAs conducted to examine the effect of bilingualism on divergent thinking, in terms of flexibility and originality in the three test types (linguistic, figural, and ideational). Due to group differences in years of speaking English, and trends for differences in non-verbal intelligence, age, and English vocabulary, six ANCOVAs were also conducted with these variables as covariates. The results are shown in Table 4. None of the tests showed a significant difference between bilingual and monolingual groups. Partial eta squared values also indicate negligible to small effect sizes.

4. Discussion

The present study investigated the effect of bilingualism on children's divergent thinking. Our results do not suggest there is an

Table 2
Descriptive statistics for bilinguals and monolinguals in divergent thinking tasks.

Test	Score	Monolinguals				Bilinguals			
		Mean	SD	Min	Max	Mean	SD	Min	Max
Linguistic (Word meanings)	Fluency	2.85	1.04	1.00	5.60	2.63	0.91	1.00	4.60
	Flexibility	1.90	0.64	0.20	2.80	1.88	0.65	0.60	3.80
	Originality	0.66	0.077	0.39	0.79	0.65	0.075	0.51	0.82
Figural (Circles task)	Fluency	4.43	3.26	1.00	20.00	4.69	2.85	1.00	14.00
	Flexibility	3.09	2.21	1.00	14.00	2.79	1.66	1.00	7.00
	Originality	0.84	0.10	0.68	1.00	0.82	0.096	0.68	0.99
Ideational (Alternate uses task)	Fluency	4.82	1.64	1.50	7.00	4.81	2.52	1.50	15.50
	Flexibility	3.76	1.43	1.50	6.50	3.59	1.88	1.00	11.00
	Originality	0.90	0.037	0.74	0.94	0.88	0.043	0.80	0.97

Table 3
Correlations between scores in divergent thinking tasks and verbal and non-verbal intelligence.

	Word Meanings			Circles Task			Object Uses						
	1 Fluency	2 Flexibility	3 Originality	4 Fluency	5 Flexibility	6 Originality	7 Fluency	8 Flexibility	9 Originality	10 Non-verbal intelligence	11 English vocabulary	12 Years of English	13 Age
1		.674*	.374*	.372*	.440*	.324*	.489*	.479*	.361*	.333*	.518*	.485*	.453*
2	.578*		.664*	.201*	.341*	.238*	.349*	.366**	.252*	.443*	.608*	.455*	.508*
3	.212*	.565*		.131	.268*	.218*	.290*	.331**	.237*	.427*	.638*	.407*	.452*
4	.307*	.100	.030		.695*	.303*	.379*	.308**	.225*	.222*	.193*	.186	.232*
5	.386*	.267*	.188*	0.678		.698*	.230*	.194*	.129	.126	.327*	.294*	.232*
6	.316*	.221*	.197*	.290*	.697*		.165	.145	.109	.052	.274*	.261*	.096
7	.425*	.252*	.193*	.337*	.177	.145		.916*	.757*	.285*	.324*	.278*	.275*
8	.418*	.278*	.245*	.263*	.141	.125	.909*		.811*	.276*	.382*	.338*	.265*
9	.351*	.229*	.211*	.207*	.107	.100	.761*	.816*		.112	.225*	.213*	.110
10	.055	.171	.194*	.096	-.035	-.014	.145	.141	.054		.530*	.219*	.653*
11	.320*	.416*	.506*	.050	.237*	.287*	.194*	.285*	.207*	.153		.658*	.682*
12	.305*	.228*	.198*	.065	.201*	.253*	0.151	.235*	0.185	-.260*	.440*		.582*

Note: Correlations above the diagonal bivariate, below the diagonal partial controlling for age. * $p < .05$.3.3 Comparing Bilingual and Monolingual Groups

Table 4
ANOVA and ANCOVA results for comparison of monolingual and bilingual children.

Test	Score	ANOVA			ANCOVA		
		F	p	η_p^2	F	p	η_p^2
Ideational (Object uses task)	Flexibility	0.03	.859	.000	0.47	.493	.004
	Originality	0.19	.662	.002	0.25	.616	.002
Figural (Circles task)	Flexibility	0.58	.450	.005	0.03	.864	.000
	Originality	2.04	.156	.018	0.00	.985	.000
Linguistic (Word meanings)	Flexibility	0.90	.346	.008	0.01	.935	.000
	Originality	2.65	.107	.024	0.61	.437	.006

effect of bilingualism on creativity in terms of divergent thinking, at least as it was defined and measured in our study. This was the case for both flexibility and originality scores across tests of divergent thinking which used both verbal and non-verbal response modes and three different types of stimuli (words, ideas, and figures).

The lack of evidence for an effect of bilingualism identified here is not entirely inconsistent with previous findings. The present results do diverge from previous findings by suggesting that there is no evidence of an effect of bilingualism in the same sample across three different divergent thinking tests and all outcome measures, either before or after controlling for potential confounds, specifically non-verbal intelligence, vocabulary, language exposure, and age. Previous studies found an effect of bilingualism on at least one outcome measure in verbal (e.g. Leikin, 2013) or figural (Sampedro & Peña, 2019) divergent thinking tasks, and usually – though not always (Torrance et al., 1970) – with an advantage for bilinguals. However, all existing studies of the effect of bilingualism on children's creativity that we identified have also generated null effects for some domains and outcome measures of divergent thinking (Carringer, 1974; Jacobs & Pierce, 1966; Lee & Kim, 2011; Leikin, 2013; Leikin & Tovli, 2014; Okoh, 1980; Ricciardelli, 1992; Sampedro & Peña, 2019; Torrance et al., 1970). Thus, the findings dispute the conclusion from the inconsistent results of previous studies that bilingual children have superior performance in creativity tasks.

The null results reported here appear robust, due to the adequate sample size and reliable measures, and thus rule out some possible alternative explanations of the null results. Firstly, sample size is unlikely to be an issue as the sample size in the present study was sufficient to detect large effect sizes and showed reasonable power to detect medium effect sizes. It was also larger than average for previous studies. Indeed, the effect sizes for the null results presented here are also small: this suggests that if effects exist but are undetected, they are likely to be small and therefore of limited practical significance. Secondly, the null results were not likely due to measurement error, as the tasks used have been shown to be reliable and valid in previous research (Beck et al., 2016; Dewing, 1970; Jacobs & Pierce, 1966; Okoh, 1980; Runco, 1985; Runco & Acar, 2010; Torrance, 1966). Indeed the inter-rater reliabilities in the present study were substantial to near perfect (Conger, 2017) and performance on tests correlated with age, vocabulary, and nonverbal IQ, as would be expected from previous research (Haavold, 2018; Lee & Therriault, 2013; Nusbaum & Silvia, 2011; Wallach & Kogan, 1965; Wu et al., 2005).

This failed attempt to conceptually replicate the previous finding of a positive effect of bilingualism on divergent thinking in children raises the possibility that results from previous research were exaggerated in size, not reliable, or due to methodological limitations. The use of multiple outcome measures in this field increases the risk of false positive results – or Type I error – when significance adjustments are not made (Wicherts et al., 2016). Some studies included small sample sizes (Carringer, 1974; Leikin, 2013; Leikin & Tovli, 2014) which reduces the likelihood that statistically significant results reflect a genuine effect (Button et al., 2013). Furthermore, many studies sampled monolingual and bilingual children from different populations (Jacobs & Pierce, 1966; Leikin, 2013; Leikin & Tovli, 2014; Sampedro & Peña, 2019; Torrance et al., 1970), which increase the likelihood of confounding differences between samples (e.g., in SES). This is compounded by a lack of control for correlates of creative task performance (e.g., language proficiency and non-verbal IQ) in existing studies. Therefore, it seems plausible that previous differences identified might have been exaggerated in size and could be explained by confounding differences between groups rather than an effect of bilingualism itself.

The lack of benefit of bilingualism to children's creativity suggested here contributes to the wider debate on the cognitive benefits of bilingualism (Bialystok, 2001). In terms of the theoretical position on the causes of cognitive development, the data here suggest that developing linguistic proficiency in two or more languages does not seem to contribute to children's divergent thinking development. However, the findings also speak to a field which is keen to demonstrate the cognitive advantages of bilingualism beyond linguistic skills (Fox et al., 2019), sometimes for the honourable purpose of justifying the need for more and better quality foreign or second language teaching (Mehmedbegovic, 2018). This study does not show any such bilingual advantage for creativity in children. However, regardless of this and similar findings, the value of bilingualism and language learning should be recognised as an ends in their own right, and should not need to be justified on the basis of other cognitive benefits beyond language and communication skills (Eberhard, Simons, & Fennig, 2021; Fox et al., 2019; Murphy et al., 2015; Yelland et al., 1993).

Despite the improved methodology of this study compared to previous research in the area, there are some limitations to mention. One possible limitation is that in this sample, bilinguals had less exposure to English (the language of testing) than their monolingual peers. Whilst English exposure and vocabulary were controlled in analyses, and whilst the sample used here are representative of many bilingual children who learn their languages sequentially, it may be that the results are not representative of simultaneous bilinguals. This study, like most previous ones, used divergent thinking measures which assess the novelty but not the usefulness of creative ideas, and thus the findings cannot be generalised to all facets of the creative process. A further limitation is that this study did not measure constructs related to the putative mechanistic accounts for how bilingualism might impact creativity (executive functions or multicultural experiences). For example, we do not know whether bilingual children in our sample had more experience of living in other countries or participating in different cultures than monolingual children, only that they spoke a diversity of languages originating from other countries. Thus, it is still possible that multicultural experiences impact divergent thinking independently of bilingual status.

Further research in this area would be beneficial to identify contributors to children's creativity. A systematic review and meta-analysis of previous findings in the area of bilingualism and creativity, and including published and unpublished data in case of publication biases (Kühberger et al., 2014), would be beneficial in establishing whether the weight of the evidence is in favour of a significant effect or not. Furthermore, future research should more systematically consider the individual differences that may mediate an effect of bilingualism on creativity: in particular, factors such as language proficiency, general intelligence, multicultural experiences, and executive functions. The features of the specific creative task should also be considered: for instance, second language learners are less likely to produce common multi-word phrases (e.g., pay attention, Smith & Murphy, 2015), so tasks which score novel linguistic productions may demonstrate an effect of bilingualism.

In conclusion, no evidence was found for an effect of bilingualism on divergent thinking in this well-powered, pre-registered study. There was a consistent lack of effect across three measures of divergent thinking used in previous studies. This raises the possibility that any differences between bilinguals and monolinguals are due to other between-group confounds such as general intelligence. It also suggests that future research should more carefully explore other candidate factors that contribute to children's divergent thinking, such as cultural diversity.

CRediT author statement

Sophie A. Booton: Conceptualization, Methodology, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Project Administration. **Elena Hoicka**: Writing – Review & Editing. **Aneyn M O'Grady**: Data curation. **Hiu Ying Nicole Chan:** Data curation. **Victoria A. Murphy:** Supervision, Funding Acquisition, Writing – Review & Editing.

Funding

This work was supported by Ferrero International [grant number R52124/CN001].

Declarations of Competing Interest

None.

Acknowledgements

The authors would like to thank the children, teachers and parents who participated in this research. We would also like to thank Alex Hodgkiss for his assistance with coding.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*(2), 207–245. https://doi.org/10.3102/0034654310368803.

Agirdag, O. (2014). The long-term effects of bilingualism on children of immigration: Student bilingualism and future earnings. *International Journal of Bilingual Education and Bilingualism, 17*(4), 449–464. https://doi.org/10.1080/13670050.2013.816264.

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology, 43*(5), 997–1013. https://doi.org/10.1037/0022-3514.43.5.997.

Amabile, T. M., & Pratt, M. G. (2016). The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior, 36*, 157–183. https://doi.org/10.1016/j.riob.2016.10.001.

Ananiadou, K., & Claro, M. (2009). 21st century skills and competences for new millennium learners in OECD countries. In *OECD Education Working Papers*. Retrieved from https://doi.org/10.1787/218525261154.

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. https://doi.org/10.1016/j.ecresq.2014.02.003.

Beck, S. R., Williams, C., Cutting, N., Apperly, I. A., & Chappell, J. (2016). Individual differences in children's innovative problem-solving are not predicted by divergent thinking or executive functions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, (20150190), 371. https://doi.org/10.1098/rstb.2015.0190.

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. https://doi.org/10.1016/j.paid.2012.04.014.

Bialystok, E. (2001). *Bilingualism in development : Language, literacy, and cognition*. Cambridge, UK: Cambridge University Press.

Bialystok, E. (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives, 9*(2), 117–121. https://doi.org/10.1111/cdep.12116.

Booton, S. A., Hodgkiss, A., Mathers, S., & Murphy, V. A. (2021). Measuring knowledge of multiple word meanings in children with English as a first and an additional language and the relationship to reading comprehension. *Journal of Child Language, First View*, 1–33. https://doi.org/10.1017/S0305000921000052.

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*(5), 365–376. https://doi.org/10.1038/nrn3475.

Carringer, D. C. (1974). Creative thinking abilities of mexican youth: The relationship of bilingualism. *Journal of Cross-Cultural Psychology, 5*(4), 492–504. https://doi.org/10.1177/002202217400500409.

Chang, J. H., Hsu, C. C., Shih, N. H., & Chen, H. C. (2014). Multicultural families and creative children. *Journal of Cross-Cultural Psychology, 45*(8), 1288–1296. https://doi.org/10.1177/0022022114537556.

Conger, A. J. (2017). Kappa and rater accuracy: Paradigms and parameters. *Educational and Psychological Measurement, 77*(6), 1019–1047. https://doi.org/10.1177/0013164416663277.

Department for Education. (2012). *EAL pupils by school*. Retrieved from https://www.naldic.org.uk/research-and-information/eal-statistics/eal-pupils/.

Department for Education. (2020). *Schools, pupils and their characteristics*. Retrieved from https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics.

Dewing, B. Y. K. (1970). The reliability and validity of selected tests of creative thinking in a sample of seventh-grade west Australian children. *British Journal of Educational Psychology, 40*(1), 35–42. https://doi.org/10.1111/j.2044-8279.1970.tb02096.x.

Duncker, K., & Lees, L. S. (1945). On problem-solving. *Psychological Monographs, 58*(5), i–113. https://doi.org/10.1037/h0093599.

Dunn, D. M., & Dunn, L. M. (2009). BPVS III. *British picture vocabulary scales* (3rd Ed.). London: GL Assessment.

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (2021). *Ethnologue: Languages of the world* (24th edition). Dallas, Texas: SIL International. Retrieved from. http://www.ethnologue.com.

Escudero, P., Mulak, K. E., Fu, C. S. L., & Singh, L. (2016). More limitations to monolingualism: Bilinguals outperform monolinguals in implicit word learning. *Frontiers in Psychology, 7*, 1–13. https://doi.org/10.3389/fpsyg.2016.01218.

Fox, R., Corretjer, O., & Webb, K. (2019). Benefits of foreign language learning and bilingualism: An analysis of published empirical research 2012–2019. *Foreign Language Annals, 52*(4), 699–726. https://doi.org/10.1111/flan.12424.

Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin, 53*(4), 267–293. https://doi.org/10.1037/h0040755.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Haavold, P.Ø. (2018). An Investigation of the Relationship between Age, Achievement, and Creativity in Mathematics. *Journal of Creative Behavior, 54*, 555–566. https://doi.org/10.1002/jocb.390.

Hayati Rezvan, P., Lee, K. J., & Simpson, J. A. (2015). The rise of multiple imputation: A review of the reporting and implementation of the method in medical research. *BMC Medical Research Methodology, 15*(1), 1–14. https://doi.org/10.1186/s12874-015-0022-1.

Jacobs, J. F., & Pierce, M. L. (1966). Bilingualism and creativity. *Journal of Elementary English, 43*, 499–503. http://doi.org.ezproxy.usq.edu.au/10.4135/9781412971959.n43.

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond Big and Little: The Four C Model of Creativity. *Review of General Psychology, 13*(1), 1–12. https://doi.org/10.1037/a0013688.

Kharkhurin, A. V. (2008). The effect of linguistic proficiency, age of second language acquisition, and length of exposure to a new cultural environment on bilinguals' divergent thinking. *Bilingualism: Language and Cognition, 11*(2), 225–243. https://doi.org/10.1017/S1366728908003398.

Kim, K. H. (2005). Can only intelligent people be creative? A meta-analysis. *The Journal of Secondary Gifted Education, 16*(2/3), 57–66. https://doi.org/10.4219/jsge-2005-473.

Kim, K. H. (2006). Can We Trust Creativity Tests? A Review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal, 18*(1), 87–98. https://doi.org/10.1207/s15326934crj1801.

Kim, K. H. (2011). The APA 2009 Division 10 debate: Are the Torrance Tests of Creative Thinking still relevant in the 21st century? *Psychology of Aesthetics, Creativity, and the Arts, 5*(4), 302–308. https://doi.org/10.1037/a0021917.

Kubota, M., Chevalier, N., & Sorace, A. (2020). How bilingual experience and executive control influence development in language control among bilingual children. *Developmental Science, 23*(1), 1–13. https://doi.org/10.1111/desc.12865.

Kühberger, A., Fritz, A., & Scherndl, T. (2014). Publication bias in psychology: A diagnosis based on the correlation between effect size and sample size. *PLoS ONE, 9*(9), 1–8. https://doi.org/10.1371/journal.pone.0105825.

Lee, C. S., Huggins, A. C., & Therriault, D. J. (2014). A measure of creativity or intelligence?: Examining internal and external structure validity evidence of the remote associates test. *Psychology of Aesthetics, Creativity, and the Arts, 8*(4), 446–460. https://doi.org/10.1037/a0036773.

Lee, C. S., & Therriault, D. J. (2013). The cognitive underpinnings of creative thought: A latent variable analysis exploring the roles of intelligence and working memory in three creative thinking processes. *Intelligence, 41*(5), 306–320. https://doi.org/10.1016/j.intell.2013.04.008.

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. https://doi.org/10.1016/j.paid.2011.01.039.

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin, 144*(4), 394–425. https://doi.org/10.1037/bul0000142.

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

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

Leung, A. K. Y., & Chiu, C. Y. (2010). Multicultural Experience, Idea Receptiveness, and Creativity. *Journal of Cross-Cultural Psychology, 41*(5), 723–741. https://doi.org/10.1177/0022022110361707.

Leung, A. K. Y., Maddux, W. W., Galinsky, A. D., & Chiu, C. Y. (2008). Multicultural Experience Enhances Creativity: The When and How. *American Psychologist, 63*(3), 169–181. https://doi.org/10.1037/0003-066X.63.3.169.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review, 69*(3), 220–232. https://doi.org/10.1037/h0048850.

Mehmedbegovic, D. (2018). What Every Policymaker Needs to Know About the Cognitive Benefits of Bilingualism. In M. Kelly (Ed.), *Languages after Brexit* (pp. 109–124). https://doi.org/10.1007/978-3-319-65169-9_10.

Murphy, V. A., Macaro, E., Alba, S., & Cipolla, C. (2015). The influence of learning a second language in primary school on developing first language literacy skills. *Applied Psycholinguistics, 36*(5), 1133–1153. https://doi.org/10.1017/S0142716414000095.

Nichols, E. S., Wild, C. J., Stojanoski, B., Battista, M. E., & Owen, A. M. (2020). Bilingualism affords no general cognitive advantages: A population study of executive function in 11,000 people. *Psychological Science, 31*(5), 548–567. https://doi.org/10.1177/0956797620903113.

Nusbaum, E. C., & Silvia, P. J. (2011). Are intelligence and creativity really so different? Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence, 39*(1), 36–45. https://doi.org/10.1016/j.intell.2010.11.002.

Okoh, N. (1980). Bilingualism and divergent thinking among Nigerian and Welsh school children. *The Journal of Social Psychology, 110*, 163–170. https://doi.org/10.1080/00224545.1980.9924242.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251). https://doi.org/10.1126/science.aac4716.

Plucker, J. A., & Makel, M. C. (2019). Assessment of creativity. In J. C. Kaufman, & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (pp. 48–73). https://doi.org/10.1017/cbo9780511763205.005.

Qi, G. Y. (2016). The importance of English in primary school education in China: Perceptions of students. *Multilingual Education, 6*(1), 1–18. https://doi.org/10.1186/s13616-016-0026-0.

Ricciardelli, L. A. (1992). Bilingualism and cognitive development in relation to threshold theory. *Journal of Psycholinguistic Research, 21*(4), 301–316. https://doi.org/10.1007/BF01067515.

Ritter, S. M., Damian, R. I., Simonton, D. K., van Baaren, R. B., Strick, M., Derks, J., & Dijksterhuis, A. (2012). Diversifying experiences enhance cognitive flexibility. *Journal of Experimental Social Psychology, 48*(4), 961–964. https://doi.org/10.1016/j.jesp.2012.02.009.

Runco, M. A. (1985). Reliability and convergent validity of ideational flexibility as a function of academic achievement. *Perceptual and Motor Skills, 61*(3_suppl), 1075–1081. https://doi.org/10.2466/pms.1985.61.3f.1075.

Runco, M. A., & Acar, S. (2010). Do tests of divergent thinking have an experiential bias? *Psychology of Aesthetics, Creativity, and the Arts, 4*(3), 144–148. https://doi.org/10.1037/a0018969.

Runco, M. A., Millar, G., Acar, S., & Cramond, B. (2010). Torrance tests of creative thinking as predictors of personal and public achievement: A fifty-year follow-up. *Creativity Research Journal, 22*(4), 361–368. https://doi.org/10.1080/10400419.2010.523393.

Ryan, C. (2013). *American Community Survey Reports: Language Use in the United States*. Retrieved from https://www2.census.gov/library/publications/2013/acs/acs-22/acs-22.pdf.

Sampedro, A., & Peña, J. (2019). Executive functions mediate the association between bilingualism and creativity in preadolescents. *Thinking Skills and Creativity, 34* (100605), 1–11. https://doi.org/10.1016/j.tsc.2019.100605.
Smith, S. A., & Murphy, V. A. (2015). Measuring productive elements of multi-word phrase vocabulary knowledge among children with English as an additional or only language. *Reading and Writing, 28*(3), 347–369. https://doi.org/10.1007/s11145-014-9527-y.
Torrance, E. P. (1966). *The Torrance tests of creative thinking-norms-technical manual research edition-verbal tests, forms A and B- figural tests, forms A and B*. Princeton, NJ: Personnel Press.
Torrance, E. P. (1981). *Thinking creatively in action and movement*. Benesville, IL: Scholastic Testing Service.
Torrance, E. P., & Ball, O. E. (1984). *The Torrance tests of creative thinking streamlined (revised) manual, figural A and B*. Bensenville, IL: Scholastic Testing Service.
Torrance, E. P., Gowan, J. C., Wu, J. J., & Aliotti, N. C. (1970). Creative functioning of monolingual and bilingual children in Singapore. *Journal of Educational Psychology, 61*(1), 72–75. https://doi.org/10.1037/h0028767.
Wallach, M. A., & Kogan, N. (1965). A new look at the creativity-intelligence distinction. *Journal of Personality, 33*(3), 348–369. https://doi.org/10.1111/j.1467-6494.1965.tb01391.x.
Weschler, D. (2014). *WISC-V: Administration and scoring manual* (5th Ed.). Bloomington, Minnesota: PsychCorp.
Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid P-hacking. *Frontiers in Psychology, 7*, 1–12. https://doi.org/10.3389/fpsyg.2016.01832.
Woumans, E., Ameloot, S., Keuleers, E., & Van Assche, E. (2019). The relationship between second language acquisition and non-verbal cognitive abilities. *Journal of Experimental Psychology: General, 148*(7), 1169–1177. https://doi.org/10.1037/xge0000536.
Wu, C. H., Cheng, Y., Ip, H. M., & McBride-Chang, C. (2005). Age differences in creativity : Task structure and knowledge base. *Creativity Research Journal, 17*(4), 321–326. https://doi.org/10.1207/s15326934crj1704_3.
Yelland, G. W., Pollard, J., & Mercuri, A. (1993). The metalinguistic benefits of limited contact with a second language. *Applied Psycholinguistics, 14*(4), 423–444. https://doi.org/10.1017/S0142716400010687.
Zachopoulou, E., Makri, A., & Pollatou, E. (2009). Evaluation of children's creativity: Psychometric properties of Torrance's "thinking creatively in action and movement" test. *Early Child Development and Care, 179*(3), 317–328. https://doi.org/10.1080/03004430601078669.
UNESCO (2016). If you don't understand, how can you learn? Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000243713.
http://wordnetweb.princeton.edu/perl/webwn, (2010–. (Accessed 1 March 2019).