The Montessori Model and Creativity

David J. Fleming, Furman University
Brooke Culclasure, Furman University
Daniel Zhang, Georgetown University

Keywords: *Montessori, creativity, creative potential, divergent thinking, convergent thinking, school choice*

Abstract: Prior research has demonstrated that the characteristics of school environments can impact the development of creativity in children. Thus, we explored the construct of creativity in the context of a Montessori environment. We used the Evaluation of Potential Creativity to measure creativity in children during one academic year. The study sample comprised 77 third-grade students at a Montessori public school in the southeastern United States and 71 demographically similar students at a traditional public school. Results show that Montessori students performed somewhat better on the Evaluation of Potential Creativity assessment than similar non-Montessori students did. Subgroup analyses indicate that male Montessori students demonstrated higher creativity than did male non-Montessori students. The findings of this study augment the body of research supporting creative development in Montessori children and suggest that researchers should continue to focus on the measurement of creativity in studies related to the efficacy of the Montessori model.

Considered one of the most important skills for childhood development (Kaufman & Sternberg, 2010; Runco, 2004), creativity contributes to an individual's problem-solving and innovative abilities, which play a crucial role in personal growth and development (Besançon & Lubart, 2008). Creativity is commonly acknowledged as the ability to produce original works (Nijstad & Paulus, 2003; Runco & Jaeger, 2012). Most scholars agree that creativity, often referred to as a *habit of mind*, "involves invention, problem-solving, and adaptation" (Cossentino & Brown, 2014–2015, p. 229). Plucker, Beghetto, and Dow (2004) defined creativity as "the interaction among aptitude, process, and environment by which an individual or group produces a perceptible product that is both novel and useful as defined within a social context" (p. 90). This latter definition suggests that creativity is not an intrinsic characteristic but an ability that can be influenced by contextual factors. Several studies on creative ability also have demonstrated the impact of educational context (Besançon & Lubart, 2008; Besançon, Lubart, & Barbot, 2013). These studies show that characteristics of the school environment—such as instructions from teachers, tasks and exercises, and classroom space—can either foster or suppress creativity development (Besançon & Lubart, 2008; Besançon et al., 2013).

Operationalizing Creativity

While creativity is a difficult construct to measure, many scholars believe it can be "identified, described, and measured" (Cossentino & Brown, 2014–2015, p. 229). A number of researchers have developed assessments that examine different aspects of creativity. One of the most frequently used methods to assess creativity is the psychometric approach (Kaufman & Sternberg, 2010). Such an approach uses various creativity tests to measure an individual's creative potential. Tests of creative potential usually fall into two categories: those that evaluate creative expression, such as verbal responses or drawing, and those that evaluate creative thinking (Barbot, Besançon, & Lubart, 2015; Lau, Cheung, Lubart, Tong, & Chu, 2013). The most widely used creativity tests, such as the Torrance Tests of Creative Thinking, the Wallach–Kogan Creativity Tests, Guilford's Alternate Uses, and the Test for Creative Thinking-Drawing Product, belong to the latter. These tests emphasize subjects' divergent thinking, meaning the extent to which the participant can expand the range of creative problem-solving. Thus, these creativity tests require students to develop multiple alternative concepts based on original ideas (Lau et al., 2013). However, these traditional creativity tests have come under criticism because they fail to evaluate *convergent thinking*, another critical part of creative thinking identified by scholars (Kaufman & Sternberg, 2010). Convergent thinking refers to the process of combining elements and then presenting them in new ways.

To measure creative potential through examining both divergent and convergent thinking, this study employed the Evaluation of Potential Creativity (EPoC)[1], a validated assessment developed by Barbot, Besançon, and Lubart (2011). The EPoC requires participants to generate new ideas based on a stimulus in the divergent task and asks participants to integrate various items into a new product in the convergent task. The EPoC offers multiple test forms to examine different dimensions of creative potential, including verbal–literary, graphic–artistic, and social problem-solving. In this study, the graphic–artistic test is used. The reliability and validity of the EPoC was determined by a confirmatory factor analysis, which demonstrated an acceptable adjustment of the data to the theoretical model for multiple test forms. External validity was confirmed by a comparison of EPoC scores with IQ measurements, as well as a demonstrated correlation between EPoC scores and personality-relevant dimensions. There also were correlations between EPoC scores and the classic subtest of divergent thinking derived from the Torrance Tests of Creative Thinking, indicating both convergent and divergent validity (Barbot et al., 2011).

Evidence of Creativity and Montessori

There are many reasons why Montessori education may affect students' creativity: the independence and freedom offered to students, the structure of the Montessori classroom, the flexibility of space and time, and the emphasis on intrinsic motivation and collaboration. Introduced in the early 20th century, the Montessori pedagogy emphasizes the freedom of children and building an environment that supports each child's development (Gutek & Gutek, 2016; Lillard, 2005). Students are encouraged to learn through doing versus being instructed by teachers (Lillard, 2005). The role of the teacher is as a facilitator of learning, acting to meet students' individual needs through observations (Humphryes, 1998). Children in mixed-aged classrooms are free to choose where to work, who to work with, and which of the specially developed Montessori materials to use at their own pace (Lillard, 2005). Cossentino and Brown (2014–2015) further presented the Montessori classroom as a place where creativity is cultivated: "The Montessori classroom is explicitly designed to enable the acquisition of specific bodies of knowledge alongside the cultivation of cognitive flexibility, risk-taking, and tolerance of ambiguity" (p. 230). These and other scholars believe the Montessori model to be a holistic educational approach that nurtures students' creative development.

Although limited in number, several studies have evaluated the relationship between Montessori education and creativity. Lillard and Else-Quest (2006) conducted a study comparing Montessori and non-Montessori students after both primary and elementary school. The authors examined creativity, in addition to other measures of academic and social development, in a 12-year-old cohort. Students were asked to complete a story within 5 minutes that began "_____ had the best/worst day at school." Researchers found that Montessori

[1] Evaluation du Potentiel Créatif (EPoC) was initially developed and validated in a sample of French students. It is translated as the *Evaluation of Potential Creativity* in the United States but is still commonly referred to using the French acronym.

students produced stories that were significantly more creative than non-Montessori students' stories. A study by Heise, Böhme, and Körner (2010), which examined the development of intelligence and creativity of pupils of Montessori and traditional teaching methods, found that Montessori students showed higher levels of creativity and better performance in geometry. While these studies provided support for the notion that Montessori increases creativity more than traditional education does, a recent evaluation came to a different conclusion. Lillard et al. (2017) conducted a 3-year longitudinal study with a cohort of students, beginning in preschool. In the study, researchers measured numerous aspects of academic, social, and cognitive development, including creativity. Lillard et al. used Guilford's Alternative Uses to measure creativity. Results demonstrated no statistically significant differences between Montessori and non-Montessori students on the measure of creativity across the years of the study. In light of the conflicting conclusions regarding the effect of Montessori on creativity, the study presented in this paper provides additional insight into this debate.

In addition to studies that directly explore Montessori and creativity, the various elements of Montessori have also been shown to benefit creative development in children. For example, several studies have found that educational environments in which children view themselves to have some level of control and that allow for free choice in activities and collaborative learning have been shown to produce higher levels of creativity (Amabile & Gitomer, 1984; Ryan & Grolnick, 1986). In addition, focusing on intrinsic motivation rather than extrinsic rewards, as is the case in Montessori education, has been shown to affect creativity. A number of studies investigated the influence of extrinsic rewards on creativity. According to Lillard (2005), the use of rewards was shown to reduce intrinsic motivation to learn and think creatively, leading students to learn only material on which they expected to be evaluated and rewarded. Lepper, Greene, and Nisbett (1973) found that children who self-selected to draw, but were later prompted to draw with the knowledge that an award would be given, showed lower levels of creativity compared to children who were never presented with the possibility of receiving an award. Another study produced similar findings when researchers asked elementary school students to take two photographs and then create a line of text to go with each picture (Amabile, Hennessey, & Grossman, 1986). Students who were led to believe that the photography task was an advance reward produced fewer creative lines. Moreover, Amabile (1979) found that undergraduate students' awareness that a work would be evaluated, without knowing the specific criteria for evaluation, reduced the originality and creativity of the work.

While creativity may differ by school context, student characteristics also play an important role. For example, several studies have focused on gender in the development of creative potential. A study by Sayed and Mohamed (2013) explored gender differences in divergent thinking in approximately 900 Egyptian children from kindergarten through grade 6. The students' divergent thinking was assessed using the Test for Creative Thinking-Drawing Production. Results of the study indicated no consistent gender differences in divergent thinking. Additionally, a longitudinal study by Lau and Cheung (2015) that used the Wallach–Kogan Creativity Tests to measure creativity in nearly 2,500 junior high students in a Chinese school showed some patterns of gender differences in scores, depending on the grade levels; however, gender differences narrowed by grade 9. Another study investigated gender differences in creativity among 985 schoolchildren using the Test for Creative Thinking-Drawing Production (Wu-jing & Wong, 2011). Results showed complex patterns of gender differences and no consistent advantages for either boys or girls (Wu-jing & Wong, 2011). The relatively few studies focusing on gender and creative development warrant additional research in this area.

Method

Participants

In this study, we examined how the performance of Montessori and non-Montessori students differed on the graphic–artistic section of the EPoC assessment. The sample comprised 148 third-grade students at two public schools during the 2015–2016 academic year. Of these, 77 attended a Montessori public school and 71 attended a traditional public school. Previous evaluations of Montessori programs have noted varying levels of fidelity to the Montessori model in Montessori schools (e.g., Lillard, 2012). Proper implementation of Montessori may be particularly difficult in public-school settings, as some standards and accountability requirements may prove incompatible with high-fidelity Montessori education.

To ensure that the public Montessori school participating in this study was of high fidelity, trained Montessori professionals observed the classrooms and interviewed Montessori teachers in the school. Based on the findings of Montessori observers, the research team was confident that this school implemented the Montessori model with high fidelity.

The Montessori school included in this study was selected because it was a no-choice situation regarding participation in a public Montessori program; in other words, all students in the district enrolled in preschool were placed in a public Montessori program. Thus, the third-grade students in the sample began Montessori education in the district at age 3 or 4. This was important because, whereas Montessori education is a parental choice in most public schools that offer a Montessori program, this school offered only Montessori classes to children aged 3 and 4, thus helping to mitigate some of the issues related to selection bias.

The Montessori school in this study did not have a waiting list of students, so a randomized control trial was not possible. When selecting the non-Montessori sample of students, the research team considered traditional public schools that were similar to the Montessori school in the study in a number of important dimensions, including school size, grade configuration, location, and student demographics. The Montessori school and the traditional comparison school used in this study are both in rural areas of the same state in the southeastern United States. While there were demographic differences between the samples in this study, we were primarily concerned with ensuring that the samples were similar in the percentage of low-income students. This factor was emphasized in the comparison-school selection process because of the important effect that parental income can have on student outcomes.

Design and Procedure

The research team merged the EPoC results with a state database, which provided demographic information on the study participants. This allowed for a demographic comparison between the Montessori and non-Montessori students in this study. Further, these demographic variables (i.e., race, free/reduced meal eligibility, gender, English language learners, special education status) may affect student creativity and, therefore, were taken into account when examining the relationship between Montessori participation and creativity. These demographic variables, as well as the Montessori indicator variable, were dummy coded (1 = yes, 0 = no) for inclusion in the multivariate analyses described below.

Trained researchers administered the EPoC assessment to both Montessori and non-Montessori students in comparable school-day settings. The EPoC assessment was standardized (i.e., same task materials, same time allotted, same instructions, same scoring method) and required students to produce work (i.e., drawings) based on a specific set of stimuli. The researchers asked students to complete one divergent-exploratory task and one convergent-integrative task during the first session and then another divergent-exploratory task and convergent-integrative task in a second session approximately two weeks later. This allowed each child to show his or her creative potential on two occasions with two slightly different tasks for divergence and convergence.

For the divergent-exploratory tasks, we showed students a picture of an abstract shape (see Figure 1) and asked them to complete as many drawings as they could that incorporated the object. Students also completed a similar task using a picture of a concrete object, such as a carrot (see Figure 2). The more drawings they completed, the higher score they achieved on the divergent task. The divergent-exploratory task score equaled the number of legitimate ideas (i.e., drawings) produced.

For the convergent-integrative tasks, we showed students one image of a set of eight different and unrelated abstract shapes and one image of a set of eight different

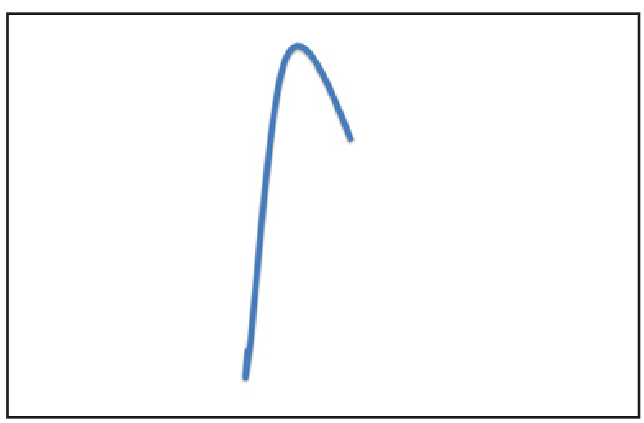

Figure 1. Example of divergent-exploratory (abstract) object from Evaluation of Potential Creativity assessment.

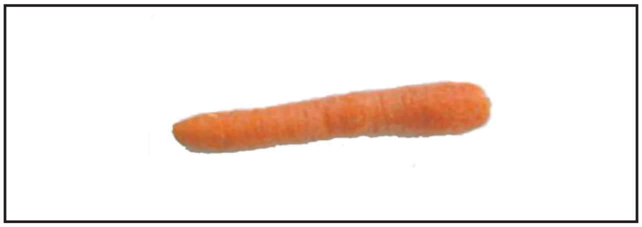

Figure 2. Example of divergent-exploratory (concrete) object from Evaluation of Potential Creativity assessment.

and unrelated concrete objects. Each child was asked to complete two drawings: one drawing using the eight abstract shapes and one drawing using the eight concrete shapes. We then asked the students to tell the story behind each drawing, and scored the students on a 7-point scale for each drawing based on detailed EPoC guidelines, with 1 the lowest score and 7 the highest. See Figure 3 for an example of a drawing that received the lowest score of 1 on the convergent-integrative task using concrete objects. See Figure 4 for an example of a drawing that received a 7, the highest possible score, on the same task.

Scoring for the convergent tasks accounted for a number of elements, including whether the participant

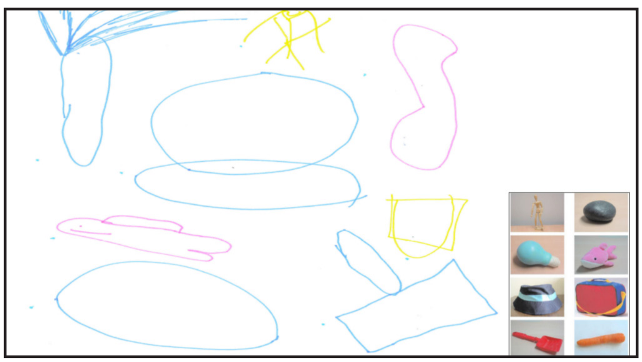

Figure 3. Example of low-scoring convergent-integrative (concrete) task from Evaluation of Potential Creativity assessment.

Figure 4. Example of high-scoring convergent-integrative (concrete) task from Evaluation of Potential Creativity assessment.

used all eight elements, the ways in which elements were combined in new and creative ways, whether the drawing was meaningful, and the originality of the idea being expressed. Participants' divergent-exploratory and convergent-integrative scores were not based on their craftsmanship or technical drawing ability. Two trained evaluators blindly scored each of the 148 convergent-integrative abstract drawings and 124 convergent-integrative concrete drawings. The average of the two scores for both the abstract and concrete drawings constituted the students' total scores on these two assessments. Using a weighted interrater-reliability procedure, we found that the ratings from the two coders produced a kappa statistic of .65 on the abstract drawings. This was considered a *substantial* level of agreement (Landis & Koch, 1977). For the concrete drawings, kappa was .58, a *moderate* level of agreement (Landis & Koch, 1977). A summary of the four parts of the EPoC assessment is presented in Table 1.[2]

Following the EPoC scoring guidelines, we used the abstract and concrete divergent-exploratory measures to create a single divergent-exploratory score. The final EPoC score, which is the focus of this study, is a combination of the divergent-exploratory score, the concrete convergent-integrative score, and the abstract convergent-integrative score, as specified by the EPoC guidelines. In addition to these outcomes, coders on the research team also measured technical drawing ability by assessing each student's ability to create a meaningful and visually appealing drawing by incorporating a variety of skills and abilities, including perspective, proportion, texture, differential shapes, and size.

Statistical Analysis Approach

The main analyses proceeded as follows. First, we compared the demographic characteristics of the Montessori and non-Montessori students. Then, the relationship between Montessori participation and creativity was examined. We used difference-in-means

[2] Despite the best efforts of the research team, some students did not complete all four assessments. Twenty-four students were missing the convergent-integrative concrete score. Following guidance from the creators of the EPoC, we used these students' convergent-integrative abstract scores in place of their concrete scores when computing the total score. This maximizes the sample sizes for the analyses. When the analyses are limited to students who have complete data ($n = 124$), the results are substantively similar to what is presented in the results section.

Table 1
Four Tasks Constituting the Evaluation of Potential Creativity Assessment

	Dimension	
Stimulus	Divergent-exploratory	Convergent-integrative
Abstract	Creating a number of unique drawings using an abstract stimulus	Combining eight abstract shapes into one meaningful drawing
Concrete	Creating a number of unique drawings using a concrete stimulus	Combining eight concrete objects into one meaningful drawing

t tests to investigate the bivariate relationships between Montessori status and creativity. Then, we employed ordinary least squares (OLS) regression to estimate differences in final EPoC scores between Montessori and non-Montessori students. Given that the component and final EPoC scores were continuous, researchers employed OLS regression to isolate the relationship between Montessori status and the dependent variable while accounting for differences in student demographics.

After finding that Montessori students exhibited higher levels of creativity than did non-Montessori students, we examined the different components of the EPoC assessment to identify areas in which Montessori students outscored their counterparts. Further, we explored whether Montessori education increased creative potential for some groups of students more than for others. To examine this possibility, we estimated multivariate OLS regressions that interacted the Montessori indicator variable with various subgroups. In the first model, we interacted the Montessori indicator with student gender, allowing examination of whether the Montessori effect was different for male and female students. In the second model, we estimated the interaction between Montessori participation and free-reduced meal eligibility. In the third model, we investigated the interaction between Montessori participation and student race. Because of the small sample sizes, the analysis was limited to an examination of White, non-Hispanic students and non-White students. Besides the main effects and interaction terms of these variables, the analyses also controlled for the other demographic variables that were included in the other regressions. Finally, we performed robustness checks that examined the extent to which selection bias may explain the results found in these analyses.

Results

Preliminary Analyses

We first established that the two groups of participating third-grade students in this study were similar according to demographics. Significantly, the two samples were found to be very similar in terms of the proportion of students who were eligible for free or reduced meals, our proxy for low-income status, as seen in Table 2. Montessori students were more likely to be male, non-Hispanic White or Black, and not using special education services, when compared to non-Montessori students. However, these differences were not statistically significant. There were statistically significant differences between the two groups in terms of the proportion of Hispanic students and those deemed English language learners. Approximately 10% of Montessori students were Hispanic compared to 23% of traditional students in the sample, and 7% of Montessori students were English language learners compared to 17% of traditional students. While the samples were found to be similar overall, we performed additional statistical procedures, described below, to account for the differences found.

To get a sense of the differences in the total raw EPoC scores and the different EPoC components individually, we performed multiple bivariate, difference-in-means tests to examine the relationship between school type and students' scores before adjusting for student demographic factors. The results are in Table 3, which shows that, before adjusting for demographics, Montessori students' final EPoC scores were higher than those of non-Montessori students. This difference was significant at the $p < .10$ level (two-tailed). Montessori students also outscored non-Montessori students on the divergent-exploratory tasks. On average, Montessori students

Table 2
Demographics of Montessori and Non-Montessori Study Participants

Characteristic	Montessori ($n = 77$)	Non-Montessori ($n = 71$)
Female	36 46.8%	38 53.5%
White	57 74.0%	46 64.8%
Black	12 15.6%	7 9.9%
Hispanic	8** 10.4%	16 22.5%
Poverty status	57 74.0%	51 71.8%
Special education status	9 11.7%	9 12.7%
English language learner status	5** 6.5%	12 16.9%

*** $p < .01$.** $p < .05$. * $p < .10$.

Table 3
Raw Evaluation of Potential Creativity Scores of Montessori and Non-Montessori Participants

Outcome	Montessori	Non-Montessori	Difference	
Final EPoC score	22.97	19.65	2.31*	(1.29)
Divergent-exploratory	15.05	12.55	2.50**	(1.25)
Convergent-integrative (abstract)	3.40	3.67	-0.26	(0.24)
Convergent-integrative (concrete)	3.45	3.43	0.02	(0.21)
Technical drawing ability	3.32	3.46	-0.14	(0.27)

Note. Standard errors in parentheses.
*** $p < .01$.** $p < .05$. * $p < .10$.

created 2.5 more drawings than did their non-Montessori counterparts, a statistically significant difference. The differences for technical drawing ability and both of the convergent-integrative outcomes, however, were small and not statistically significant.

Main Effects of Montessori Education on Creativity

While these results suggested a Montessori advantage on two outcomes, there were demographic differences between the two groups to consider, as demonstrated in Table 2. Thus, a multivariate analysis was used to examine these scores. To examine whether demographic factors accounted for Montessori students' performance on the EPoC test, we estimated a linear regression predicting students' final EPoC scores, the main outcome of this analysis. Table 4 presents the regression coefficients and robust standard errors. After controlling for race, poverty status, gender, student disability, and English language learner status, Montessori students scored 2.28 points higher on the EPoC than did non-Montessori students.

This is a marginally statistically significant difference with a *p* value of .077. To get a sense of the magnitude of this difference, we reestimated the regression, using standardized scores as the dependent variable, by converting the final EPoC score to a *z* score, which had a mean of 0 and a standard deviation of 1. The regression coefficients were now in standard deviation units, similar to Cohen's *d*, a popular measure of effect size. Other education evaluations have implemented this approach as well (e.g., Center for Research on Education Outcomes, 2013; Jenkins et al., 2018). Montessori students scored 0.28 standard deviations (SE = 0.16) higher than non-Montessori students on the EPoC assessment. This is a substantively large difference. Meta-analyses that examine the effect sizes of various education interventions provide a benchmark for the .28 effect size presented here. Cheung and Slavin (2016) found that the mean Cohen's *d* effect size for analyses with sample sizes similar to those in this study was 0.26, and the mean effect size across 449 quasiexperimental education studies was 0.23. The effect size of Montessori education on creativity in this study was similar to the average effect size of comparable studies. Special education status is the only other covariate that was statistically significantly related to final EPoC scores; students who received special education services scored about 3 points lower than students who did not.[3]

Because Montessori education may enhance different aspects of creative potential, we also examined the differences between Montessori and non-Montessori students in terms of the constituent parts of the EPoC, as well as the students' technical drawing ability, after adjusting for student demographics. Table 5 indicates that Montessori students scored 2.63 points higher on the divergent-exploratory score than did non-Montessori students. This result means that Montessori students, on average, drew 2.63 more pictures than non-Montessori students did, incorporating the abstract and concrete

Table 4
Predicting Final Evaluation of Potential Creativity Score—Multivariate Ordinary Least Squares Regression Results

Characteristic	Final EPoC score
Montessori	2.28* (1.28)
Poverty status	-0.37 (1.69)
Special education status	-2.96* (1.67)
English language learner status	0.37 (2.93)
Female	-0.19 (1.32)
Black	-1.46 (2.11)
Hispanic	-1.46 (2.33)
Other race	2.19 (4.94)
Constant	20.75*** (1.54)
Observations	148
F statistic	0.93
R^2	0.05

Note. Robust standard errors in parentheses.
*** $p < .01$. ** $p < .05$. * $p < .10$.

stimuli. While this outcome had direct interpretability, it also was converted into a standardized score with a mean of 0 and standard deviation of 1, as was done with total EPoC score. When these standardized scores were used as the dependent variable, Montessori students scored 0.34 standard deviations higher than non-Montessori students on the divergent-exploratory score. While Montessori students achieved lower scores on the two convergent tasks and technical drawing ability than non-Montessori students did, none of the results were statistically significant.

Interaction Effects

The above analyses suggest that Montessori education is associated with higher levels of creativity. However,

[3] Our analyses are sensitive to the presence of outliers on the dependent variable. When these observations are excluded from the regression in Table 4, Montessori students still outscore non-Montessori by 0.24 standard deviations, but this difference is no longer statistically significant (*p* value = .106) at the $p < .10$ level. Given our role in data collection, we believe that the outlier cases do not reflect errors in measurement, but rather simply high scores on the EPoC tests. Therefore, we included all cases in the analyses we present here. However, we acknowledge that the presence of outliers in conjunction with the small sample sizes are a limitation of this study.

Table 5
Examining Components of the Evaluation of Potential Creativity Assessment

	Divergent-exploratory	Convergent-integrative (abstract)	Convergent-integrative (concrete)	Technical drawing ability
Montessori	2.63**	-0.37	-0.02	-0.14
	(1.25)	(0.24)	(0.21)	(0.28)
Poverty status	-0.36	0.22	-0.33	-0.16
	(1.62)	(0.29)	(0.26)	(0.34)
Special education status	-2.11	-0.43	-0.19	0.09
	(1.52)	(0.38)	(0.27)	(0.40)
English as a second language status	2.51	-0.91	-0.84*	-1.23**
	(2.99)	(0.59)	(0.50)	(0.48)
Female	-0.86	0.25	0.41*	0.79***
	(1.27)	(0.24)	(0.22)	(0.28)
Black	-1.13	-0.13	-0.10	-0.47
	(1.97)	(0.37)	(0.27)	(0.37)
Hispanic	-2.13	0.02	0.42	0.62
	(2.37)	(0.58)	(0.45)	(0.41)
Other race	4.61	-1.24***	-1.15***	-1.21
	(5.13)	(0.37)	(0.23)	(0.84)
Constant	13.58***	3.63***	3.57***	3.29***
	(1.56)	(0.27)	(0.26)	(0.35)
Observations	148	148	124	148
F statistic	1.08	2.81***	13.57***	2.80***
R^2	0.05	0.07	0.08	0.10

Note. Robust standard errors in parentheses.
***$p < .01$. **$p < .05$. *$p < .10$.

whether attending a Montessori school may be particularly effective for certain subgroups of students remains an open question. To examine this possibility, we estimated three interaction models, which examined differences in the effect of Montessori education by gender, free/reduced meal eligibility, and race. These models included controls for all student demographic factors used in Table 4.

After estimating these regressions, we determined the predicted final scores for these subgroups of students.[4] Figure 5 displays the predicted final EPoC score from the separate regressions for gender, income, and race.[5] The other variables in the model were held at their observed values (Williams, 2012). According to Figure 5, male Montessori students had a predicted final EPoC score of 23, while male non-Montessori students had a score of 18. This marginal effect was statistically significant at the $p < .05$ level, meaning that Montessori participation was associated with greater levels of creativity for male

[4] Full regression results are available from the authors upon request.

[5] The predicted outcomes reflect average marginal effects estimated using the Margins command in Stata. The significance tests used in Figure 5 correspond to the relationship between Montessori participation and final EPoC score within each subgroup (e.g., Montessori female students vs. non-Montessori female students, White Montessori students vs. White non-Montessori students).

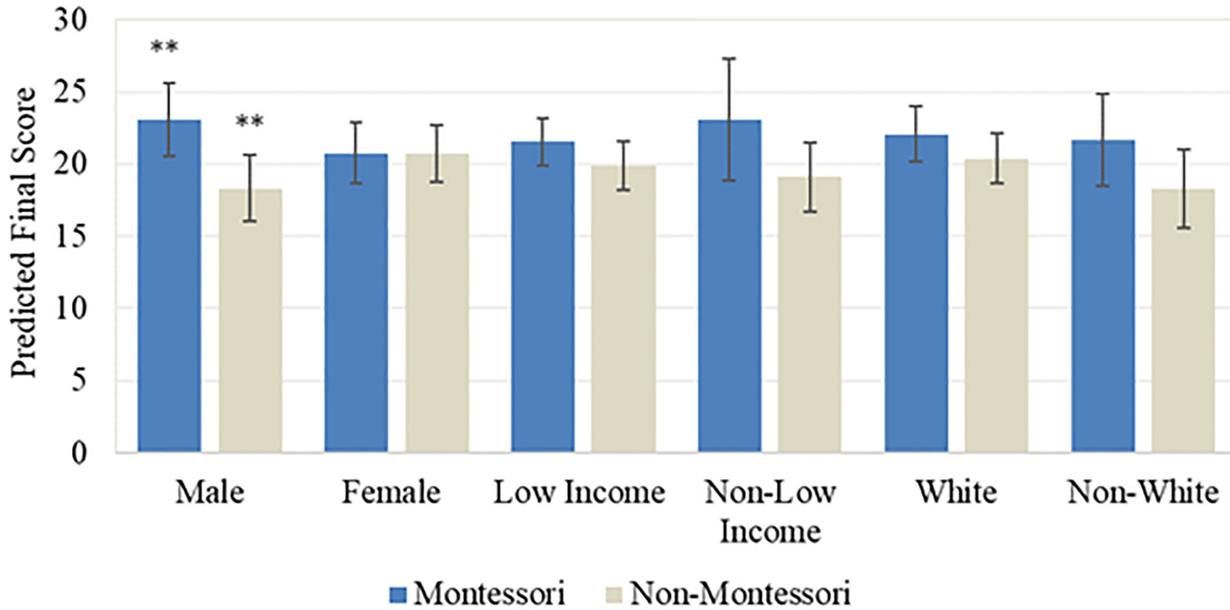

Figure 5. Predicted final scores for interaction models. Figure depicts predicted final score with 90% confidence intervals. ***$p < .01$, **$p < .05$, *$p < .10$.

students. We then examined the differences between male students when using the standardized EPoC final score as the outcome variable. Male Montessori students scored 0.59 standard deviations above male non-Montessori students. When examining the other subgroup analyses for income and race, Montessori students consistently scored higher than non-Montessori students. However, these differences were not statistically significant. The small sample size of this study was particularly limiting for these subgroup analyses, as there was not enough power to detect small differences between Montessori and non-Montessori students.

Robustness Analyses

A major challenge in evaluations of this type is selection bias. In the case examined here, selection bias may occur if important factors led some parents to choose public Montessori education for their children and if these factors were related to student creativity. For example, more-involved parents may have been more likely to send their children to Montessori programs. These parents also may have been more active with their children at home and may have encouraged creative problem-solving. If selection bias is not accounted for, the higher levels of creativity exhibited by Montessori students on the final EPoC score may simply be because their parents were more involved, not because they participated in Montessori education. Unlike some other studies that examined the effects of Montessori education (Lillard & Else-Quest, 2006; Lillard et al., 2017), we were not able to use a randomized lottery to account for selection bias. Rather, we hoped to decrease the chances of selection bias by selecting a school district that did not allow parents to choose between public Montessori education and traditional preschool: Montessori education was the only public option. Further, the analyses accounted for a number of important student characteristics in the form of control variables in the OLS regressions.

Because selection bias may be unobserved, we were unable to estimate how selection bias affected the results presented here. However, methods exist that allow researchers to examine the percentage of the estimated effect that must be caused by bias to invalidate the inference that there is a difference between the scores of Montessori and non-Montessori students (i.e., to no longer have a statistically significant result). We applied such a procedure (Rosenberg, Xu, & Frank, 2018) to the main result of this analysis. Using the Montessori coefficient (2.28) and standard error (1.28) from the regression in Table 4, which predicted the final EPoC score, we estimated that 7% of the Montessori effect would have to be caused by selection bias to infer that there was no statistically significant relationship between Montessori status and final EPoC score.

Frank, Maroulis, Duong, and Kelcey (2013) examined a number of evaluations on these dimensions and found that the bias necessary to invalidate the results ranges from 2% to 60% for these education studies. The result here of 7% was on the low end of that distribution, but it was higher or equivalent to the bias needed to invalidate inferences related to a tutoring program (Miller & Connolly, 2013) and a counseling program to encourage college enrollment (Stephan & Rosenbaum, 2013).

Another way to consider the threat of selection bias is to examine the impact threshold for confounding variables (Crosnoe & Cooper, 2010; Frank et al., 2008). This value quantifies how powerful an unknown confounder must be correlated to both Montessori participation and creativity to negate the relationship found in this analysis or to make the relationship between Montessori participation and final EPoC scores no longer statistically significant at $p < .10$. For example, how correlated must parental involvement be with Montessori participation and creativity to make the Montessori estimate in Table 4 no longer statistically significant? Using the technique presented by Rosenberg et al. (2018), the impact threshold for confounding variables was estimated to be .012. This estimate meant that, for the Montessori coefficient to no longer be statistically significant, the confounding variable must be correlated at .11 with Montessori participation and with the final EPoC score at .11, conditional on the covariates included in the model. For comparison purposes, we examined the other covariates in the regression model. None of the covariates was correlated with both Montessori participation and creativity at the impact threshold for confounding variables level. This result meant that the omitted confounder would need to be more strongly related to Montessori participation and the final EPoC score than are free or reduced meal eligibility, race, English language learner status, gender, and special education status in the data.

Discussion

Prior research shows that creativity, which is critical for children to develop as they move toward adulthood, can be be affected by the educational context in which students learn. The question then centers on which educational environments are most conducive to the development of creativity in children. This study explored the potential of Montessori education to affect creativity in children. Several of its key elements make it likely to affect this construct, particularly the independence and freedom of choice given to children and the lack of extrinsic rewards to motivate them, both of which provide an environment for children to develop creative skills. Past research has supported the notion that components of Montessori education could increase students' creativity.

This study suggests that experience with Montessori education may be related to greater levels of creative potential. This relationship was particularly pronounced for male students, as male Montessori students scored significantly higher on the final EPoC than male non-Montessori students. We were unable to identify why Montessori education may be particularly effective for male students. Future studies should more closely examine the mechanisms through which Montessori education may enhance the creativity of male students.

The findings of this evaluation should not be overstated. While this study provides some evidence that Montessori education may enhance creativity, the analyses suggest no statistically significant differences between Montessori and traditional students for most of the results presented here. For example, we found insignificant interaction effects between Montessori participation and race and poverty status. Further, because the positive relationship between the main effect of Montessori participation and final EPoC score was statistically significant at the $p < .10$ level (two-tailed), readers should interpret this association with caution. The marginal statistical significance can be partially explained by the small sample size, as the estimated effect size of Montessori participation was near the mean effect size of education evaluations of a similar type. The Montessori advantage on the divergent-exploratory component of the EPoC and the higher final EPoC score for male Montessori students compared to that of male traditional students was significant at the $p < .05$ level (two-tailed), and the effect sizes were substantively important.

Like all evaluations of this type, this study has limitations. First, creativity, by its very nature, is difficult to study and measure. The use of the EPoC in this study allowed us to examine both divergent and convergent creativity with a validated instrument, but it was limited to graphic–artistic tests. Other researchers may prefer different measures of creativity than what was analyzed here. Additionally, only one public Montessori program in a rural school district in the southeastern United States, with 77 third-grade Montessori students, participated in this study. It is unclear if the results of this study would apply to Montessori students in different grades, different types of Montessori schools (e.g., private vs. public), and different locations (e.g., urban vs. rural). Further, the small sample size limited the power of the study and made subgroup analyses particularly challenging. The non-experimental nature of this study is the final significant limitation. As noted above, unobserved selection bias in the form of omitted variables threatens the internal validity of this study. Because a randomized control trial was not feasible, the research team chose a comparison school that was demographically similar to the Montessori school; the team also used important covariates in a multivariate analysis to try to mitigate the problem of selection bias. Acknowledging that selection bias may exist despite these efforts, we provided robustness checks, which quantified how large selection bias would need to be to invalidate the result.

In reviewing means of measuring the efficacy of models in pre-K–12 education, such as Montessori education, many researchers have begun to realize the importance of including measurements of social-emotional skills, such as creativity, in any comprehensive study. Although the importance of these types of skills is recognized intuitively, longitudinal research also has confirmed that such qualities predict academic, social, psychological, and physical wellbeing (Duckworth & Yeager, 2015). The findings of this study, which suggest that Montessori students perform better on an assessment of creativity, add to the body of research supporting creative development in Montessori children and suggest that researchers should continue to focus on the measurement of creativity in studies related to the efficacy of Montessori education.

Author Information

†Corresponding Author
David J. Fleming is an associate professor in politics and international affairs and a senior researcher with the Richard W. Riley Institute at Furman University.

Brooke Culclasure† is the Riley Institute at Furman University's research director. She can be reached at Furman University, 3300 Poinsett Highway, Greenville, SC, 29613. Phone: 864-294-3236 or brooke.culclasure@furman.edu.

Daniel Zhang is a master's candidate in the Security Studies Program at Georgetown University's Edmund A. Walsh School of Foreign Service.

References

Amabile, T. M. (1979). Effects of external evaluation on artistic creativity. *Journal of Personality and Social Psychology, 37*, 221–233. doi:10.1037//0022-3514.37.2.221

Amabile, T. M., & Gitomer, J. (1984). Children's artistic creativity. *Personality and Social Psychology Bulletin, 10*, 209–215. doi:10.1177/0146167284102006

Amabile, T. M., Hennessey, B. A., & Grossman, B. S. (1986). Social influences on creativity: The effects of contracted-for reward. *Journal of Personality and Social Psychology, 50*(1), 14–23. doi:10.1037//0022-3514.50.1.14

Barbot, B., Besançon, M., & Lubart, T. (2011). Assessing creativity in the classroom. *The Open Education Journal, 4*, 58–66. doi:10.2174/1874920801104010058.

Barbot, B., Besançon, M., & Lubart, T. (2015). Creative potential in educational settings: Its nature, measure, and nurture. *Education 3-13, 43*, 371–381. doi:10.1080/03004279.2015.1020643

Besançon, M., & Lubart, T. (2008). Differences in the development of creative competencies in children schooled in diverse learning environments. *Learning and Individual Differences, 18*, 381–389. doi:10.1016/j.lindif.2007.11.009

Besançon, M., Lubart, T., & Barbot, B. (2013). Creative giftedness and educational opportunities. *Educational and Child Psychology*, *30*(2), 79–88.

Center for Research on Education Outcomes. (2013). *National charter school study*. Stanford, CA: Stanford University. Retrieved from https://credo.stanford.edu/documents/NCSS%202013%20Final%20Draft.pdf

Cheung, A. C. K., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher, 45*, 283–292. doi:10.3102/0013189X16656615

Cossentino, J., & Brown, K. (2014–2015). Assessing creativity and critical thinking in schools: Montessori as a holistic intervention. *AMI Journal, 2014–2015*, 229–232. Retrieved from https://www.public-montessori.org/wp-content/uploads/2016/10/Assessing-Creativity-and-Critical-Thinking-in-Schools-Cossentino-and-Brown.pdf

Crosnoe, R., & Cooper, C. E. (2010). Economically disadvantaged children's transitions into elementary school: Linking family processes, school contexts, and educational policy. *American Educational Research Journal, 47*, 258–291. doi:10.3102/0002831209351564

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher, 44*, 237–251. doi:10.3102/0013189x15584327

Frank, K. A., Maroulis, S. J., Duong, M. Q., & Kelcey, B. M. (2013). What would it take to change an inference? Using Rubin's causal model to interpret the robustness of causal inferences. *Educational Evaluation and Policy Analysis, 35*, 437–460. doi:10.3102/0162373713493129.

Frank, K. A., Sykes, G., Anagnostopoulos, D., Cannata, M., Chard, L., Krause, A., & McCrory, R. (2008). Does NBPTS certification affect the number of colleagues a teacher helps with instructional matters? *Educational Evaluation and Policy Analysis, 30*(1), 3–30. doi:10.3102/0162373707313781

Gutek, G. L., & Gutek, P. (2016). *Bringing Montessori to America: S. S. McClure, Maria Montessori, and the campaign to publicize Montessori education*. Tuscaloosa: The University of Alabama Press.

Heise, E., Böhme, E., & Körner, S. (2010). Montessori-orientierter und traditioneller Grundschulunterricht: Ein Vergleich der Entwicklung von Rechtschreibung, Rechnen, Intelligenz und Kreativität. *Psychologie in Erziehung und Unterricht, 57*, 273–289. http://dx.doi.org/10.2378/peu2010.art19d

Humphryes, J. (1998). The developmental appropriateness of high-quality Montessori programs. *Young Children, 53*(4), 4–16.

Jenkins, J. M., Duncan, G. J., Auger, A., Bitler, M., Domina, T., & Burchinal, M. (2018). Boosting school readiness: Should preschool teachers target skills or the whole child? *Economics of Education Review, 65*, 107–125. https://doi.org/10.1016/j.econedurev.2018.05.001

Kaufman, J. C., & Sternberg, R. J. (2010). *The Cambridge handbook of creativity*. New York, NY: Cambridge University Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Lau, S., & Cheung, P. C. (2015). A gender-fair look at variability in creativity: Growth in variability over a period versus gender comparison at a time point. *Creativity Research Journal, 27*(1), 87–95. doi:10.1080/10400419.2015.992685

Lau, S., Cheung, P. C., Lubart, T., Tong, T. M., & Chu, D. H. (2013). Bicultural effects on the creative potential of Chinese and French children. *Creativity Research Journal, 25*(1), 109–118. doi:10.1080/10400419.2013.752281

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology, 28*(1), 129–137. doi:10.1037/h0035519

Lillard, A. S. (2005). *Montessori: The science behind the genius*. New York, NY: Oxford University Press.

Lillard, A. S. (2012). Preschool children's development in classic Montessori, supplemented Montessori, and conventional programs. *Journal of School Psychology, 50*, 379–401. doi:10.1016/j.jsp.2012.01.001.

Lillard, A. S., & Else-Quest, N. (2006). Evaluating Montessori education. *Science, 313*, 1893–1894. doi:10.1126/science.1132362

Lillard, A. S., Heise, M. J., Richey, E. M., Tong, X., Hart, A., & Bray, P. M. (2017). Montessori preschool elevates and equalizes child outcomes: A longitudinal study. *Frontiers in Psychology, 8*. http://dx.doi.org/10.3389/fpsyg.2017.01783

Miller, S., & Connolly, P. (2013). A randomized controlled trial evaluation of time to read, a volunteer tutoring program for 8- to 9-year-olds. *Educational Evaluation and Policy Analysis, 35*, 23–37. doi:10.3102/0162373712452628

Nijstad, B. A., & Paulus, P. B. (2003). *Group creativity: Innovation through collaboration*. Oxford, England: Oxford University Press.

Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist, 39*(2), 83–96. doi:10.1207/s15326985ep3902_1

Rosenberg, J. M., Xu, R., & Frank, K. A. (2018). Konfound-It! Quantify the robustness of causal inferences [Computer software]. Retrieved from http://konfound-it.com

Runco, M. A. (2004). Creativity. *Annual Review of Psychology, 55*(1), 657–687. doi:10.1146/annurev.psych.55.090902.141502

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal, 24*(1), 92–96. doi:10.1080/10400419.2012.650092

Ryan, R. M., & Grolnick, W. S. (1986). Origins and pawns in the classroom: Self-report and projective assessments of individual differences in children's perceptions. *Journal of Personality and Social Psychology, 50,* 550–558. doi:10.1037/0022-3514.50.3.550

Sayed, E. M., & Mohamed, A. H. H. (2013). Gender differences in divergent thinking: Use of the Test of Creative Thinking-Drawing Production on an Egyptian sample. *Creativity Research Journal, 25,* 222–227. doi:10.1080/10400419.2013.783760

Stephan, J. L., & Rosenbaum, J. E. (2013). Can high schools reduce college enrollment gaps with a new counseling model? *Educational Evaluation and Policy Analysis, 35,* 200–219. doi:10.3102/0162373712462624

Williams, R. (2012). Using the margins command to estimate and interpret adjusted predictions and marginal effects. *Stata Journal, 12,* 308–331.

Wu-jing, H., & Wong, W. (2011). Gender differences in creative thinking revisited: Findings from analysis of variability. *Personality and Individual Differences, 51,* 807–811. doi:10.1016/j.paid.2011.06.027