Journal of Montessori Research
2016, Volume 2, Issue 1

Removing Supplementary Materials from Montessori Classrooms Changed Child Outcomes

Angeline S. Lillard† and Megan J. Heise

University of Virginia

Acknowledgements

Support was provided by grants to the first author from the Brady Education Foundation and the Sir John Templeton Foundation. We thank the children, parents, and schools who participated in the research, the research assistants who helped to conduct the studies, and, especially, Alyssa Hart, who oversaw data collection and coding.

Keywords*: cognitive development, executive function, reading, fidelity of implementation*

Abstract. Montessori classrooms vary in the degree to which they adhere to Maria Montessori's model, including in the provision of materials. Specifically, some classrooms use only Montessori materials, whereas others supplement the Montessori materials with commercially available materials like puzzles and games. A prior study suggested such supplementation might explain observed differences across studies and classrooms (Lillard, 2012), but an experimental study is the best test of this possibility. Here we present such an experiment, with 52 children in three Montessori classrooms, all of which included supplementary materials at the start of the study. Non-Montessori materials were then removed from two of the classrooms, and all children were given six pretests as a baseline. Four months later, children were retested to see how much they had changed across that period. Children in the classrooms from which the non-Montessori materials were removed advanced significantly more in early reading and executive function and advanced to some degree more in early math than children in the other two classrooms. There were no differences across the classroom types in amount of change on the tests of vocabulary, social knowledge, or social problem-solving skills.

Although virtually all Montessori classrooms share some features, like giving children some choices about what work they do and when they do it, classrooms can also vary a great deal. One of the dimensions along which they vary is the materials offered, with some classrooms including toys and puzzles in addition to the Montessori materials. Lillard (2011), in an effort to establish an agreed-upon core set of Primary classroom materials, created a teacher questionnaire that listed dozens of materials seen in Primary Montessori classrooms. The survey asked 29 Primary-level teacher trainers—59% at American Montessori Society (AMS) training centers and the rest at Association Montessori Internationale (AMI) training centers—to describe each material as *necessary*, *desirable*, *acceptable*, or *best avoided* in a Primary-level classroom.

Although there was not 100% agreement (even within training types), at least 85% of all the teacher-trainers agreed a large core set belongs in a Montessori Primary classroom. At issue is whether

Published online May 2016

child outcomes differ when classrooms include materials beyond this core set.

One reason this study is of interest is that using only Montessori materials could reflect fidelity of implementation, and fidelity might explain inconsistent results observed across studies of Montessori outcomes. Some studies have found that children attending Montessori schools have better outcomes than other children. For example, Dohrmann and colleagues looked at standardized test scores and grades for Milwaukee high school students who, years earlier, had attended public Montessori or other schools (Dohrmann, Nishida, Gartner, Lipsky, & Grimm, 2007). They found that the Montessori students scored significantly higher in math and science, and slightly (but not significantly) higher in social studies and English compared to age-matched peers. Lillard and Else-Quest (2006) compared children who had lost a lottery to attend a Montessori school with children who gained admission; they found several significant differences at ages 5 and 12, all favoring children in Montessori. Other studies have also shown positive results for Montessori children (Besançon & Lubart, 2008; Brown & Steele, 2015; Miller & Bizzell, 1984; Rathunde & Csikszentmihalyi, 2005a, 2005b; Rodriguez, Irby, Brown, Lara-Alecio, & Galloway, 2005). However, some studies have not shown better outcomes for children in Montessori programs. For example, Lopata and colleagues found that Montessori children scored lower in reading in eighth grade than did children in other programs (Lopata, Wallace, & Finn, 2005). Krafft and Berk (1998) found less private speech (a self-regulatory activity) in Montessori children compared to children in a play-oriented preschool program (see also Cox & Rowlands [2000] and early results from the 1970s Head Start). However, the Montessori implementation in these latter cases appeared to be of low fidelity. For example, Krafft and Berk described children's work occurring over a single 45-minute period at work stations; the work period was thus too short, and work stations are not part of Montessori education. Another study, which found better outcomes for Latino students in Montessori versus traditional schools and equal outcomes for African-American students, had single-aged classrooms, and other features of implementation fidelity were not well addressed ("We believe [the programs were] rigorous," Ansari & Winsler, 2014, p. 5). In sum, a possible explanation for different outcomes is that the research showing less positive outcomes was conducted at schools in which Montessori implementation was of low fidelity.

Fidelity in Montessori can be measured in many ways, and currently there is no single, accepted measure. As mentioned above, the materials a program offers and uses can be one index. One study compared Primary children in three classic Montessori classrooms (offering exclusively Montessori materials as determined by Lillard [2011]) with children in nine supplemented Montessori classrooms (offering a variety of other materials, such as worksheets, commercial puzzles, and crafts, in addition to Montessori materials). In fall and spring, children were given a wide variety of tests that assessed early academic and socioemotional competence, allowing measurement of change across the school year. In the classic classrooms, children were engaged with Montessori materials almost 100% of the time, whereas in the supplemented classrooms, engagement with Montessori materials ranged from 38%–56% of the time. The gain from fall to spring was higher among classic Montessori children on every variable tested—significantly so for most variables. The Head–Toes–Knees–Shoulders (HTKS) test of executive function, for example, is a Simon Says-type game in which children must do the opposite of what the experimenter tells them to do; for instance, children must touch their toes when the experimenter says, "Touch your head!" From fall to spring, children in classic Montessori gained on average almost 14 points (equivalent to following an additional seven out of 40 commands correctly compared to their fall performance), whereas supplemented Montessori children gained an average of just 7 points, or 3.5 commands. The Letter–Word test of early reading, a subscale of the Woodcock–Johnson Tests of Achievement III (WJ III; Woodcock, McGrew & Mather, 2001) also showed particularly strong gains for classic Montessori children, as did the Picture Vocabulary test.

At issue is whether the presence of only Montessori materials caused these different levels of gain, or whether some other "third variable" was responsible for the differences. Lillard (2012) proposed materials as an index of fidelity; yet, in and of themselves, the materials might not be important. Perhaps teachers who choose to have only Montessori materials in their classrooms also adhere more tightly to other aspects of the Method, and it is those aspects, rather than the materials, that led to the larger gains.

To examine this issue, in the present study a Head of School at a school with three Primary classrooms removed all non-Montessori materials from two of those classrooms over a weekend. Researchers tested children immediately after this change was made and again 4 months later. Changes in children's scores across the 4 months in the two classrooms from which the non-Montessori materials were removed were compared with changes in the one classroom in which the non-Montessori materials remained.

Method

Participants

Fifty-five children completed pretest assessments in the first 2 weeks after they returned from the winter holiday; 53 of these children took posttests 4 months later, in the final weeks of the school year (the remaining two children were absent during the retest period). In addition, one child performed much worse in the spring on four of five tests, suggesting error (child not trying or incapacitated) given that children typically improve on these tasks with age; this child was also excluded from all analyses. The final sample of 52 children had a mean age in January of 57.4 months (SD = 13.2, range 31–83 months, 27 boys). Of these, 45% were 3-year-olds, 36% were 4-year-olds, and 18% were 5-year-olds. Breaking down this demographic by subsample, there were 35 children in the two classrooms from which materials were removed (M_{age} = 56.9 months, SD = 14.0, range = 31–83 months, 16 boys). In the unchanged classroom, there were 17 children (M_{age} = 58.2 months, SD = 11.7, range = 40–77 months, 11 boys). The percentages of children and mean ages at each age level in each type of class were about the same. Ethnicity data were not collected, but the school's demographic representation is similar to that of the local community: about 70% white, 20% African American, 5% Asian, and 5% multiracial or other races.

Materials and Procedure

Participants were children from three Primary classrooms (3 to 6 years) at a Montessori school that used non-Montessori materials in all classrooms. The Head of School volunteered to test the hypothesis that removing materials would influence child outcomes. Prior to removing materials and then again well into the spring semester, four classroom "snapshots" were taken, in which an observer noted what each child in each classroom was doing; the percentage of children engaged with Montessori materials and the percentage engaged with supplementary materials were calculated.

Two teachers agreed to have supplementary materials removed from their classrooms. Children in the different classrooms performed similarly on all pretests, and key teachers in both types of classrooms had taught Primary children at the school for more than 10 years. Parents received a cover letter from the Head of School, a letter from the researchers describing the study, and an informed consent form. All children with parental consent participated (see prior section); participation rates ranged from 65% to 75% across the three classrooms. The testers were blind to the intervention and the study hypotheses, except that the first author conducted one participant's pretest as part of the research assistant training. Testing occurred in January and May of a single school year.

Setting. Each classroom in the school had two trained Montessori teachers and 24 to 27 children in each classroom. All teachers were certified by a major Montessori organization (AMI or AMS). The school implemented the Montessori program with some deviations. The primary deviations from the program described in Montessori's books were (a) the use of two trained teachers, rather than one teacher and one untrained assistant; (b) the replacement of work periods with specials—art, music, and Spanish—three times a week (out of 10 work periods); and (c) the supplementary non-Montessori materials, removed from two classrooms for the experiment. Examples of these supplementary materials include a basket of small, plastic ladybugs intended for counting; cassette players and head phones for listening to

stories while looking at books; commercial puzzles; commercial building blocks; a plastic baby doll with a washtub; and worksheets of Montessori materials for coloring.

Measures. Six measures were given in a fixed order at both pretest and posttest to assess social cognition (theory of mind), social problem solving, executive function, reading, vocabulary, and math.

Theory of mind. Social cognition or theory of mind was assessed using the theory of mind scale (Wellman & Liu, 2004). This set of tasks is designed to measure an understanding of others' minds and emotions. Researchers administered four of the five tasks in the scale: knowledge access, contents false belief, not-own belief, and real–apparent emotion. We omitted the first task, not-own desire, because all children of the ages tested were expected to pass it. For the knowledge access task, the experimenter first showed the child a toy drawer and asked what he or she thought was in the drawer. The experimenter revealed the true contents of the drawer (a small toy dog at pretest and a toy frog at posttest) and then placed the item back in the drawer. The experimenter then presented the child with a small doll and said the doll had never seen inside the drawer before. The test questions asked whether the doll knew what was in the drawer and whether the doll had seen inside the drawer; children received 1 point for a *no* answer to each question. For the contents false belief task, the experimenter showed the child a box of Band-Aid adhesive bandages that contained a toy pig at pretest, and a box of crayons that contained a brush at posttest. After the child was shown the contents, the box was closed again. The experimenter told the child that a doll had never seen inside the box before and asked the child what the doll thought was inside (1 point) and whether the doll had seen inside (1 point). In the not-own belief task, the experimenter asked the child if he or she believed that a cat would be hiding in a garage or in some bushes (each scenario was shown in a picture). The experimenter told the child that a doll believed the opposite of what the child believed; the child then was asked where the doll would search for the cat (1 point). For the real–apparent emotion task, the experimenter presented three simple faces that were labeled *happy*, *sad*, and *okay*, based on mouth appearance. At pretest, children were told that a boy's aunt

> promised that she would buy Matt a toy car. But, she got Matt a t-shirt instead. Matt doesn't like t-shirts. What Matt really wants is a toy car. But, Matt has to hide how he feels, because if his aunt knows his real feelings, she'll never buy him anything again[1].

After a memory check asking about what Matt wanted, what he got, and what would happen if his aunt knew how he really felt, each child was asked to point to the face that showed how Matt really felt (1 point) and how his face looked (1 point). The posttest story was structurally the same but involved Joey's uncle giving him a ball instead of a bicycle.

Executive function. Executive function was assessed using the HTKS task (Cameron Ponitz et al., 2008; Ponitz, McClelland, Matthews, & Morrison, 2009). Children were instructed that when the experimenter says, "Touch your head," they should instead touch their toes, and when told "Touch your toes," they should instead touch their heads. Children completed four practice trials with feedback from the experimenter before moving on to 10 test trials without feedback. Each trial was scored from 0 to 2, with 0 indicating that children touched the indicated location, 1 indicating that children initially were incorrect but corrected themselves, and 2 indicating that children immediately touched the opposite of the indicated location as instructed. When children scored at least 10 points, they continued to the Knees–Shoulders part of the task. For this part, additional instructions were given regarding touching their knees and shoulders. They again completed four practice trials on just the knees–shoulders commands, followed by 10 more trials using all four instructions. Possible total scores ranged from 0 to 40.

Social problem solving. One object-acquisition story from the Social Problem Solving Test-Revised (SPST-R) was used (Rubin, 1988). Children featured in the story illustrations matched the

[1] This wording deviates from the wording in Wellman and Liu (2004) but comes precisely from a Theory of Mind Scale script that Wellman and Liu provided to the first author in October 2009, which directs people who use the script to cite Wellman and Liu (2004).

participating child's race and gender, as is customary. Children were told, "[A reading child] has been looking at this book for a long time and [an onlooker] really wants to look at the book. What could [onlooker] do or say so he/she could have a look at the book?" Children's responses were quickly recorded by hand, and then children were asked, "What else could he/she do or say?" and finally, "What if it was you? What could you do or say so you could have a look at the book?" Responses were scored on their number of references to sharing and fairness, as in Lillard (2012), with a possible range from 0 to 3.

Reading, vocabulary, and math. Three subscales (Letter–Word ID, Picture Vocabulary, and Applied Problems) from the Woodcock–Johnson Tests of Achievement III (WJ III), a standardized norm-referenced scale (Woodcock, McGrew, & Mather, 2001), were administered according to the WJ III manual; raw scores were used because gain scores were analyzed.

Results

The classroom snapshot coding revealed that removing supplementary materials greatly reduced the time spent with non-Montessori materials. Prior to removing the materials, children were engaged with supplementary non-Montessori materials about 25% of the time and engaged with Montessori materials 34% of the time—the remaining time was spent in specials, outside, and so on. After the materials were removed from two classrooms, children in those classrooms were engaged with Montessori materials 58% of the time, and use of supplementary materials was minimal. (Apparently, some supplementary materials had reappeared, as 5% of the time children were using them!) In the unchanged classroom, 42% of activities involved Montessori materials and 24%—about the same as earlier—involved supplementary ones.

Next, we consider child outcomes. In January, children in the two types of classrooms (non-Montessori materials retained versus removed) scored similarly on all the tests except HTKS, on which children in the retained non-Montessori materials classroom performed significantly better (retained M = 29.82, SD = 7.12; removed M = 22.29, SD = 14.23). Thus these children were more advanced at the outset in executive function, but on all five other measures, the scores of children in the two types of classrooms were the same.

Of interest was how much children changed in the remaining 4 months of the school year, after the removal of the non-Montessori materials from two classrooms. Therefore, gain scores were calculated by subtracting the pretest (January) from the posttest (May) scores for each child; t-tests—one-tailed, because we had a specific hypothesis, based on Lillard (2012)—were performed on these scores, comparing children in the classroom that retained its non-Montessori materials with children in the two classrooms that removed the non-Montessori materials. Two tests yielded medium effect-size changes that were significant at the $p < .05$ level, and one yielded a small effect-size change that was nonsignificant given the small sample size.

First, children in the classrooms from which non-Montessori materials were removed advanced significantly more than children from the unchanged classroom over the 4 months in early reading, as indicated by their Letter–Word scores, $t(50) = 1.88$, $p = .035$, Cohen's $d = .58$; see Figure 1. Their Letter–Word scores improved on average by 4.54 points (SD = 4.46), whereas those of children in the unchanged classroom improved by 2.24 points (SD = 3.46).

Second, children in the classrooms in which the non-Montessori materials were removed advanced significantly more on the HTKS test, $t(50) = 1.71$, $p = .047$, Cohen's $d = .51$ (see Figure 2). They improved by a mean of 5.11 points (SD = 9.39), versus a 0.41- (SD = 9.12) point gain for the children in the classroom that retained its non-Montessori materials. This result was not caused by restricted range: Children in both groups were still well below ceiling on HTKS at posttest.

The third result was nonsignificant and yielded a small effect size (Cohen's $d = .19$) so should be viewed more cautiously. Children in the classrooms from which the non-Montessori materials were removed advanced slightly more in their applied math performance (M = 1.34, SD = 1.97) than children in the classroom that retained non-Montessori materials (M = 1.00, SD = 1.54); see Figure 3.

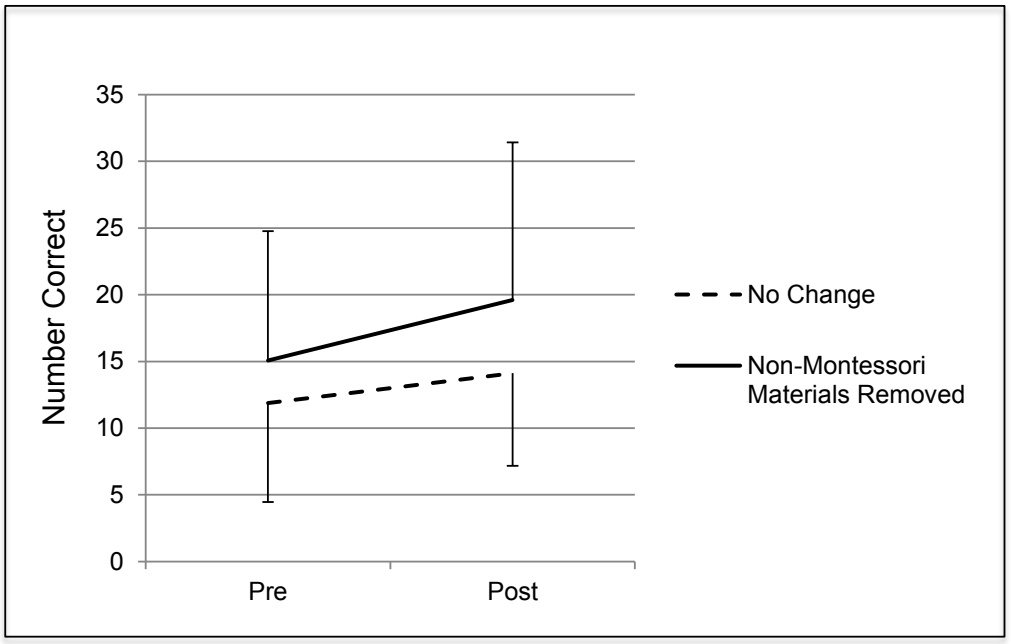

Figure 1. Change in Letter–Word scores from pretest to posttest. *Y*-axis represents number correct; error bars represent *SD*s.

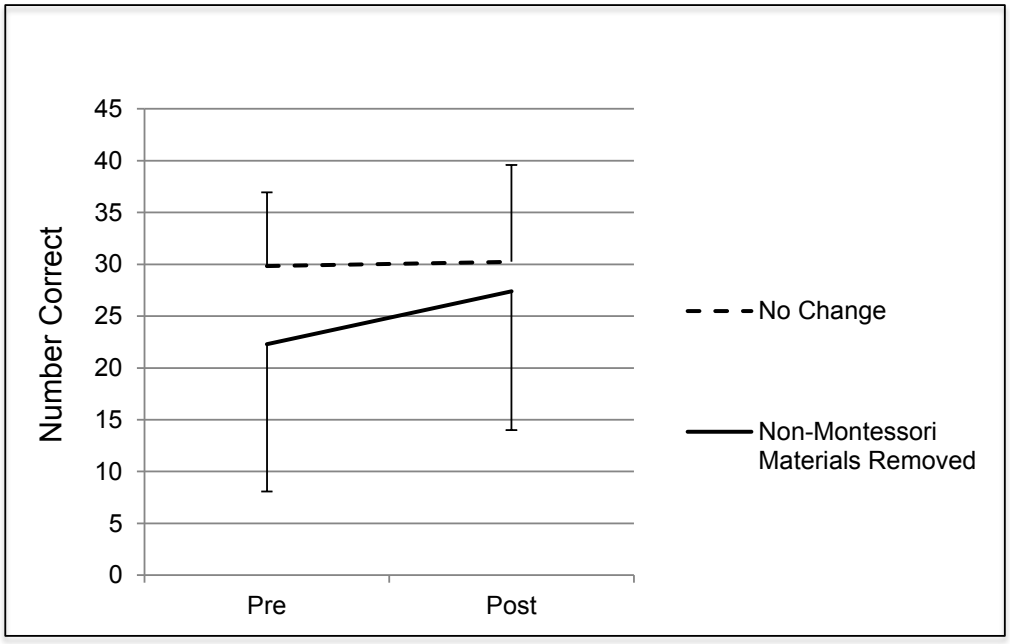

Figure 2. Change in HTKS scores from pretest to posttest. *Y*-axis represents number correct; error bars represent *SD*s.

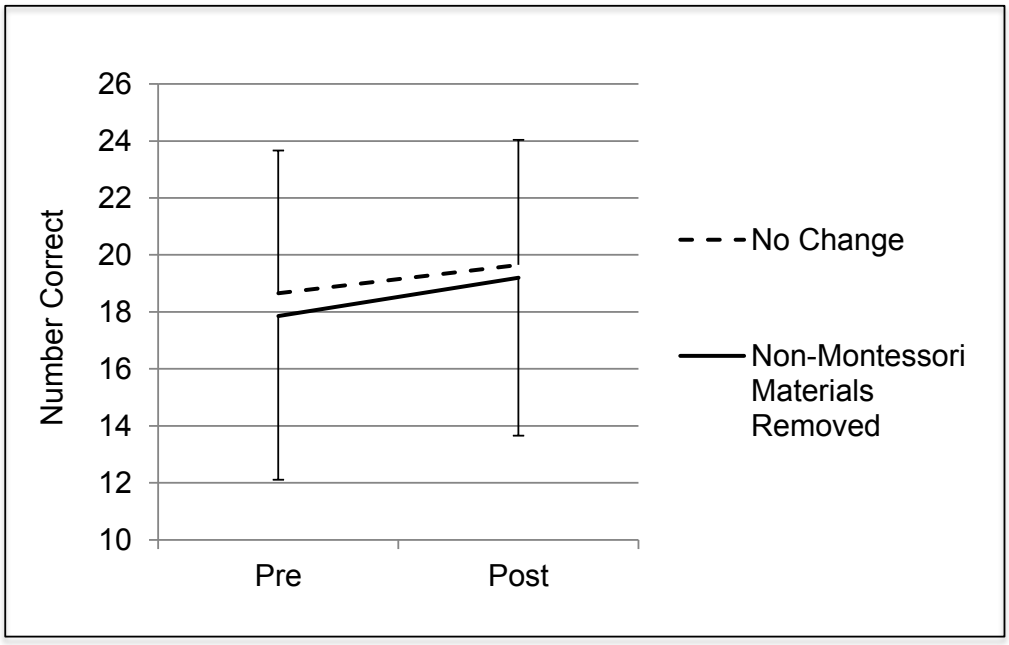

Figure 3. Change in Applied Problems scores from pretest to posttest. *Y*-axis represents number correct; error bars represent *SD*s.

On the Picture Vocabulary subtest, theory of mind scale, and the SPST-R, there was no difference in the degree of change between children in the two types of classrooms across the 4 months.

Discussion

In a prior study, children in classrooms in which children only had access to—and therefore virtually always used—Montessori materials advanced significantly more across the school year than did children in Montessori classrooms that supplemented their offerings with non-Montessori materials and in which children were using those non-Montessori materials roughly half the time (Lillard, 2012). The Montessori materials could have caused this difference, or they could have been a proxy for something else that actually caused the difference, ranging from teachers having different standards to parents' preferences for different classrooms (to the degree that parents specifically might have chosen a particular class). The best way to determine if the materials really matter is through an experiment.

A small-scale experiment was conducted to examine whether removing non-Montessori materials makes a difference on its own. Although the intervention lasted only 4 months, children in classrooms from which non-Montessori materials were removed advanced significantly more on two of six measures and slightly but nonsignificantly more on a third measure. These results are discussed in turn. As in prior research (Lillard, 2012), when supplementary materials were present, children used them a reasonable amount of the time; use of Montessori materials sharply increased (from 34% to 58% of the time) after the supplementary materials were removed.

First, although children in the unchanged classroom scored higher at both time points on the HTKS, a test of executive function, they did not improve at all on this measure across the 4 months, despite ample room for improvement. Children in the changed classrooms, by contrast, advanced a great deal over the 4 months. How might removing non-Montessori materials have influenced children's performance on this task? The task requires children to hold rules in mind and to inhibit the prepotent

response of touching the location that the experimenter's command, taken literally, told them to touch. Children, in addition, had to plan and execute the opposite response. One possible explanation for why children in the changed classrooms improved more on this task is that Montessori materials incorporate analogous demands to a greater degree than supplementary materials do. For instance, in several Sensorial exercises, children are asked to hold one sensory experience (the pitch of a Musical Bell or the length of a Red Rod) in mind as they cross the room to get its match (in the case of a Bell) or the Rod that is closest in length to the one they have. This task seems to challenge working memory in a way that, for example, putting together a commercial puzzle may not. In addition, all the Montessori materials are used according to specific steps, and children must keep these steps in mind as they plan and execute each action. In a Practical Life activity, for example, there is an order in which a child gathers the materials, lays them out for use, uses them, and finally puts them away. Perhaps Montessori teachers do not present supplementary materials with this same degree of precision. Hence, one possible reason for the rise in executive function when non-Montessori materials were removed concerns the Montessori materials themselves and how they are presented and used.

The second possible explanation we discuss is actually a by-product of the materials, and it concerns depth of concentration. Dr. Montessori repeatedly described seeing a child become transfixed by the wooden cylinders, such that even when others sang and danced around her, and even when her chair was lifted, her concentration was unbroken (Montessori, 1956, 1966, 1998). Dr. Montessori went on to observe this phenomenon in other children, with other materials. This deep concentration is something that Montessori teachers also observe today. Furthermore, Dr. Montessori claimed—and teachers today observe—that, after children had experienced this deep concentration, their personalities "normalized," and they became kinder and more compliant, made better choices, and had better self-control. Perhaps after the non-Montessori materials were removed and children had fewer options than to choose Montessori materials, they were more likely to have these concentration experiences and subsequent improvements in self-control, leading to higher scores on this task.

The advance in Letter–Word performance seems most likely to be related to the use of specific language materials. When commercial puzzles, games, crafts, projects, and other non-Montessori materials were no longer available, perhaps children went on to use the Language materials more, leading directly to this advance. Indeed, in an earlier, unpublished study involving nine Montessori classrooms, we found that the percentage of children engaged with Language materials in each Montessori classroom predicted the mean level of advance in Letter–Word performance in that classroom. Working with Language materials, like the Sandpaper Letters and the Moveable Alphabet, translates directly into doing well on the Letter–Word task, which requires children to read letters (*k* and *b*, for example) and then increasingly complex words.

This same factor may have led to the small improvement seen in the Applied Problems test, which begins with simple addition and subtraction and then moves to word problems, coins, and clock faces. With the Applied Problems test, the mapping from the materials to the test is less clear than for Letter–Word. The letter *p* on a Sandpaper Letter looks just like the *p* in the Woodcock–Johnson test, whereas the Applied Problems test has children count crayons and balloons rather than wooden spindles, red counters, and glass beads. Still, lack of non-Montessori materials may have led some children to engage more with Montessori Math materials than they otherwise might have, leading to this small increase.

Children did not advance more on the Picture Vocabulary subscale or on the two social tests (theory of mind and SPST-R). Although many Montessori materials teach nomenclature, the words taught are unlikely to align with the specific Woodcock–Johnson Picture Vocabulary test items. Vocabulary growth also accrues in conversation and reading, including books at circle time (Blachowicz, Fisher, Ogle, & Wattes-Taffe, 2006), but these activities are not likely to be influenced by the presence of materials. It is also possible that the Montessori materials would lead to better vocabulary over time, but not in 4 months.

The presence of non-Montessori materials might have little influence on social interaction in a Montessori classroom, especially if there was still only one of most materials, as was the case in these classrooms. Having only one copy of each material, regardless of its being a Montessori material or something else, might lead children to learn effective social problem-solving strategies to induce sharing behavior. Second, the degree of social interaction probably is not influenced by the amount of non-Montessori material, explaining the lack of difference on the theory of mind test. The fact that an earlier study did see classic–supplemented differences on tests of social cognition and behavior may suggest that a longer time period is needed to see differences or that the materials served as a proxy for some other classroom differences that led to different performance on the social tests. Another factor to consider is that in Lillard (2012) the children in classic Montessori classrooms used Montessori materials almost 100% of the time, whereas in the non-Montessori materials removed classrooms here, they used them only 58% of the time.

Limitations and Future Directions

Although removing non-Montessori materials did appear to influence how much children changed in the subsequent 4 months, which is consistent with Lillard (2012), the study has some clear weaknesses. First, the study was small: Only 52 children, from just three classrooms at one school, were studied, and the study was of only 4 months' duration. A larger sample would be especially useful. It is notable that the age ranges of children were similar across the different groups; the development of children of different ages is likely influenced differently by the presence or absence of different Montessori materials. Using only one school could be seen as a strength, as it means the children in the two samples were demographically similar. However, it is also possible that the individual teachers in the classrooms, rather than the change in materials, were responsible for the different levels of gain. Against this theory is the fact that children across the two types of classrooms scored the same at pretest on all but one measure. Finally, the short time frame of the study is a limitation; seeing children's trajectories over a whole school year, or several years, would be more revealing. Still, the results of this small study do suggest, using an experimental design, that children may be better served in Montessori classrooms that use only Montessori materials and that do not supplement that set of materials with commercially available toys.

Conclusions

Provision of materials is one important aspect of Montessori classrooms. Maria Montessori was very clear about this.

> The material should be limited in quantity. Properly understood, this principle is clear and logical. A normal child does not need stimuli to awaken him or put him in contact with the material world. He needs rather to bring order into the chaos created in his mind by the host of sensations coming to him from the outside world. [The child is] an ardent explorer of a world that is new to him. And what he needs, as an explorer, is a road (that is something which is straight and limited) which can lead him to his goal and keep him from wandering aimlessly about. He then passionately attaches himself to those things, limited and direct in scope, which bring order in to the chaos that has been created within him; and with this order, they provide light for his exploring mind and a guide for his researches. The explorer who was at first abandoned to himself then becomes an enlightened man who makes new discoveries at every step and advances with the strength which he receives from his inner satisfaction.
>
> Evidence of this kind should certainly modify the notion, still held by many, that a child is helped in proportion to the number of educational objects that are placed at his disposal. It is common, but false, belief that the child who has the most toys, the most help, should also be the most developed. Instead of that, the confused multitude of

> objects with which he is surrounded only aggravate the chaos of his mind (Montessori, 1967, pp. 104–105).

The materials Maria Montessori and her collaborators created were specifically designed to "bring order into the chaos" of the child's mind (p. 105), for example by abstracting the qualities of the sensory world, and engaging the child with specific routines that take care of and beautify the environment. The importance of the materials is an aspect of Montessori education that teachers appear too often to forget, as they often supplement the basic set of Montessori Primary materials with commercially available toys. The results of this small study, taken together with Lillard (2012), suggest their supplementation is a mistake and that children's development is helped when only the Montessori materials are made available.

AUTHOR INFORMATION

†Corresponding Author

Angeline S. Lillard† is a professor in the Department of Psychology at the University of Virginia in Charlottesville, Virginia. She can be reached at asl2h@Virginia.edu.

Megan J. Heise is a project coordinator in the Department of Psychology at the University of Virginia in Charlottesville, Virginia.

References

Ansari, A., & Winsler, A. (2014). Montessori public school pre-K programs and the school readiness of low-income Black and Latino children. *Journal of Educational Psychology, 106*(4), 1066–1079. doi: 10.1037/a0036799

Besançon, M., & Lubart, T. (2008). Differences in the development of creative competencies in children schooled in diverse learning environments. *Learning and Individual Differences, 18*(4), 381–389. doi: 10.1016/j.lindif.2007.11.009

Blachowicz, C. L. Z., Fisher, P. J. L., Ogle, D., & Wattes-Taffe, S. (2006). Vocabulary: Questions from the classroom. *Reading Research Quarterly, 41*(4), 524–539.

Brown, K. E., & Steele, A. S. L. (2015). Racial discipline disproportionality in Montessori and traditional public schools: A comparative study using the relative rate index. *Journal of Montessori Research, 1*(1), 14–27.

Cameron Ponitz, C. E., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., & Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly, 23*(2), 141–158. doi: 10.1016/j.ecresq.2007.01.004

Cox, M. V., & Rowlands, A. (2000). The effect of three different educational approaches on children's drawing ability: Steiner, Montessori and traditional. *British Journal of Educational Psychology, 70*, 485–503.

Dohrmann, K. R., Nishida, T. K., Gartner, A., Lipsky, D. K., & Grimm, K. J. (2007). High school outcomes for students in a public Montessori program. *Journal of Research in Childhood Education, 22*(2), 205–217.

Krafft, K. C., & Berk, L. E. (1998). Private speech in two preschools: Significance of open-ended activities and make-believe play for verbal self-regulation. *Early Childhood Research Quarterly, 13*(4), 637–658.

Lillard, A. S. (2011). What belongs in a Montessori primary classroom? Results from a survey of AMI and AMS teacher trainers. *Montessori Life, 23*(3), 18–32.

Lillard, A. S. (2012). Preschool children's development in classic Montessori, supplemented Montessori, and conventional programs. *Journal of School Psychology, 50*(3), 379–401. doi: 10.1016/j.jsp.2012.01.001

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

Lopata, C., Wallace, N. V., & Finn, K. V. (2005). Comparison of academic achievement between Montessori and traditional education programs. *Journal of Research in Childhood Education, 20*(1), 5–13.

Miller, L. B., & Bizzell, R. P. (1984). Long-term effects of four preschool programs: Ninth- and tenth-grade results. *Child Development, 55*(4), 1570–1587.

Montessori, M. (1956). *The child in the family* (N. R. Cirillo, Trans.). New York: Avon.

Montessori, M. (1966). *The secret of childhood* (M. J. Costello, Trans.). New York: Ballantine.

Montessori, M. (1967). *The discovery of the child.* (M. J. Costello, Trans.). New York: Ballantine.

Montessori, M. (1998). *Creative development in the child Vol 1* (R. Ramachandran, Trans.). Madras, India: Kalakshetra Press.

Ponitz, C. C., McClelland, M. M., Matthews, J. S., & Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology, 45*(3), 605–619. doi: 10.1037/a0015365

Rathunde, K. R., & Csikszentmihalyi, M. (2005a). Middle school students' motivation and quality of experience: A comparison of Montessori and traditional school environments. *American Journal of Education, 111*(3), 341–371.

Rathunde, K. R., & Csikszentmihalyi, M. (2005b). The social context of middle school: Teachers, friends, and activities in Montessori and traditional school environments. *Elementary School Journal, 106*(1), 59–79.

Rodriguez, L., Irby, B. J., Brown, G., Lara-Alecio, R., & Galloway, M. M. (2005). An analysis of second grade reading achievement related to pre-kindergarten Montessori and transitional bilingual education. In V. Gonzalez & J. Tinajero (Eds.), *Review of research and practice* (Vol. 3, pp. 47–68). Mahwah, NJ: Lawrence Erlbaum.

Rubin, K. H. (1988). *The Social Problem Solving Test–Revised* [Measurement instrument]. Waterloo, Canada: University of Waterloo.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*(2), 523–541.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III Tests Of Achievement* [Measurement instrument]. Rolling Meadows, IL: Riverside Publishing.