Year : 2020

When pedagogy matters : insights from Montessori education on the development of performance monitoring

Denervaud Solange

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Département de Radiologie et Radiodiagnostic
Centre Hospitalier Universitaire Vaudois

WHEN PEDAGOGY MATTERS: INSIGHTS FROM MONTESSORI EDUCATION ON THE DEVELOPMENT OF PERFORMANCE MONITORING

Thèse de doctorat en Neurosciences

Présentée à la

Faculté de Biologie et Médecine
de l'Université de Lausanne

par

SOLANGE DENERVAUD

Ingénieure diplômée de l'École Polytechnique Fédérale de Lausanne, Suisse

Jury

Prof. Andrea SERINO, Président
Prof. Patric HAGMANN, Directeur de thèse
Prof. Angeline S. LILLARD, Experte
Prof. Kerstin VON PLESSEN, Experte

Thèse n° 271

Lausanne 2020

Programme doctoral interuniversitaire en Neurosciences
des Universités de Lausanne et Genève

LEMANIC NEURO SCIENCE DOCTORAL SCHOOL

CHUV

UNIL | Université de Lausanne

UNIL | Université de Lausanne

UNIVERSITÉ DE GENÈVE

Programme doctoral interuniversitaire en Neurosciences des Universités de Lausanne et Genève

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président·e	Monsieur	Prof.	Andrea	**Serino**
Directeur·trice de thèse	Monsieur	Prof.	Patric	**Hagmann**
Expert·e·s	Madame	Prof.	Kerstin	**von Plessen**
	Madame	Prof.	Angeline	**S. Lillard**

le Conseil de Faculté autorise l'impression de la thèse de

Madame Solange Dénervaud

Ingénieure de l'École Polytechnique Fédérale de Lausanne - EPFL, Suisse

intitulée

When Pedagogy Matters: Insights from Montessori Education on the Development of Performance Monitoring

Lausanne, le 15.5.2020

pour Le Doyen
de la Faculté de Biologie et de Médecine

N. Geldner

Prof. Niko GELDNER
Directeur de l'Ecole Doctorale

The child seeks for independence by means of work; an independence of body and mind. Little he cares about the knowledge of others; he wants to acquire a knowledge of his own, to have experience of the world, and to perceive it by his own unaided efforts.

Maria Montessori, The Absorbent Mind

Abstract

The rapid pace of changes faced by todays young people calls for pedagogical practices that equip them not only with knowledge but also with the ability to think effectively, flexibly, and independently. This process rely on performance monitoring, a fundamental function of learning. When individuals notice something unexpected, such as an error, they tend to pause. In learning from this discrepant event, they adapt their behavior accordingly. Although performance monitoring is essential for academic learning and improves throughout childhood, its susceptibility to educational influences has not been studied.

Pedagogical traditions differ on how they teach children to learn from feedback and errors. Traditional education provides children from one age group with opportunities to engage in work, and then to learn about and correct their performance later based on a teachers feedback and evaluation. By contrast, Montessori education focuses on supporting children in self-correcting in real time. It utilizes specialized materials that encourage childrens self-discovery of relevant concepts, and multi-age classes in which children discuss answers as they work.

Here, we compared performance monitoring in children aged 4-15 years attending traditional or Montessori classes. Our multimodal approach (behavior, EEG, and MRI) revealed that 1) cortical regions related to performance monitoring undergo significant changes between the ages of 5 and 13 years; 2) children of that age do not process errors as adults do, and 3) pedagogical practices modulate both behavior and neural responses. More specifically, the behavioral, morphometric and EEG neural data reveal significant differences in how students notice and react to errors, and in how they self-correct. fMRI analyses reveal difference in brain network connectivity between students from the two groups, and suggest differences in error correction strategies. Finally, higher academic performances were not attributable to higher executive functions, but rather differences in creativity abilities.

Our work suggests that how students learn from errors reflects childhood schooling experience. Performance monitoring styles are also likely associated with youths cognitive flexibility more broadly, influencing how they react to novel or unexpected outcomes.

Résumé

Au vu du rythme effréné des changements auxquels sont confrontés les jeunes, il est essentiel que les pratiques pédagogiques ne se concentrent pas uniquement sur la transmission de connaissances, mais également sur leur capacité dapprendre de manière efficace, flexible et indépendante. L'élément central à cette entreprise est de favoriser une approche autodirigée et orientée sur les processus, dans laquelle les élèves développent la capacité d'apprendre de leurs erreurs. Ce processus est appelé *la gestion de la performance*. Bien que la gestion de la performance soit essentielle aux apprentissages scolaires et se développe durant l'enfance, sa susceptibilité aux influences pédagogiques n'a pas encore été étudiée.

Ici, nous avons comparé la gestion de la performance chez des enfants âgés de 4 à 15 ans, issus de classes traditionnelles ou Montessori. Alors que les pratiques pédagogiques traditionnelles mettent l'accent sur le fait que les élèves apprennent à partir des commentaires des enseignants, les pratiques pédagogiques Montessori encouragent les élèves à travailler de manière autonome avec du matériel spécialement conçu pour permettre de faire et dapprendre de leurs erreurs. Notre approche multimodale (comportement, EEG, IRM) nous a permis de dévoiler que 1) les régions corticales liées à la gestion de la performance subissent des changements importants entre 5 et 13 ans; 2) les enfants de cet âge ne traitent pas lerreur de la même manière que les adultes, et que 3) les pratiques pédagogiques modulent à la fois le comportement et les réponses cérébrales.

Ce travail constitue une première étape connectant la recherche sur la gestion de la performance avec l'émergence des habitudes mentales chez les enfants dans leurs environnements scolaires, avec des implications directes pour la recherche en développement, les professionnels de l'enfance, et les politiques.

Acknowledgments

Throughout this work, I was lucky to benefit from a multitude of scientific and human expertise that allowed me to grow scientifically and personally and, above all, carry out the different projects.

My first thanks go to **Prof. Patric Hagmann** for offering me a safe space to think and work in which I went my "unconventional" route and independently pursued my ideas. While such environments should abound for academics, they are sadly too rare and deserve to be loudly and gratefully acknowledged here. Thanks for your support. We both know how much this work was a battle, and you fought with me. You often said you were not the ideal mentor for these ideas; however, just trusting me and standing for them made you the perfect one for me. An everlasting and friendly thank you. Second, a huge thank you to **Prof. Mary Helen Immordino-Yang** for supervising some of this work, for your advice, for our night-long (swiss time!) discussion on science, education and life, and for genuinely feminizing science with a contagious passion. Thanks to **Prof. David Sander** for all the brainstorming and discussions we shared, for your creative ideas on the experimental design, for your generosity and for challenging the ideas and making them grow. A thank you to **Prof. Micah Murray** as well for help with setting experiments, ethics, recruitment and writing processes. A deep thank you to **Prof. Gilles Pourtois**: it was a real pleasure to work with you. I deeply appreciate your sharp scientific knowledge and efficiency, and thank you for taking the time to discuss all aspects and ideas. A special thanks to **Prof. Roger Beaty**, for hosting me while wrapping this work, and offering some bright insights for the next steps. **Prof. Edouard Gentaz**, despite our incompatibilities, you are the one I thank for getting this work started. Finally, a warm thank you to **Ulrike Toepel**, beyond being an extraordinary coordinator, you take such good care and time for each of us, making us feel considered, and guaranteeing that each of our Ph.D. stories carves its way in the jungle. I am also grateful to have met and worked with talented students and research assistants, who not only help with data collection, but also discussed the ideas, including **Adrien, Xavier, Morgane, Nora, Claire, Julien and Diane**. This work received precious extra funding from **the Eden Fondation**, **the Boninchi** and **the Schmidheiny Foundations** in Geneva, **the American Montessori Society**, in the United States, and **the Societe Academique Vaudoise,** in Lausanne.

My deepest gratitude goes to the experts for agreeing to be part of this journey. I admire both your work and your scientific path, and hope to collaborate with you more closely in the future. Beyond taking the time to discuss this work, I would like to thank you for your help so far. **Prof. Kerstin von Plessen**, you read our work in its early stages and provided precious feedback when it was crucially needed. **Prof. Angeline Lillard**, you were the first person I contacted when I decided to start this project. Since then, you have always replied to our requests, shared your tasks, given advice that helped me think about this work more in-depth.

Deep thanks to all my colleagues, for making tough work sweeter through the philosophical or less philosophical coffee or not-yogurt-BUT-chocolate breaks **Remy, Chiara, Joan, Hamza, Tommaso, German, Mehdi, Paola, Helene, Sebastien**. More specially, thank you to **Meri** for taking good care of the working climate at RC07. My gratitude to my colleagues, who became close friends, **Eleonora, Jef, Yasser, Fei, Emeline, Michael** and **Priscille**, with who I had the great chance to work with and be trained by. Thank you for taking the time to transfer some of your knowledge with generosity and kindness. **Emeline**, I don't know if the tarot would have predicted it, but we shared truly intense momentum, with tears of joy, laughter, tiredness, or madness. Thank you for your generosity, for always being there with deliciously baked cakes, for laughing at my not funny jokes, for our long reflections on personal and scientific topics, for the "just 5 minutes" ending in hours of coding and debugging. It was a sincere pleasure to sit back to back with you, and I genuinely wait our next pizza labmeetings! **Priscille-24-years-18-months**, what a great fun to have you around, I truly appreciate your sense of humor, your joyful Celine Dion songs and karaoke time, your frankness and honesty, your ability to share who you are that fosters highly qualitative time, and obviously, your still-not-mature prefrontal cortex that confers you this unique and so much appreciated cognitive flexibility ;-) that shakes our steady days. Thank you, **Jef**, for entering my Ph.D. journey just at the right time, for sharing your expertise and your supervision. Your generosity, wisely swiss advice and long discussions made this journey a memorable one. Finally, my gratitude to my (never-been-to-)sport-mate, **Yasser**, for being an inspiring friend with a great capacity to humbly share your time and knowledge.

I would like to acknowledge the support received from my long-standing friends, who still remain despite my absence and my obsessive talking about my work. Thanks for not judging, but sharing these interests, and for helping at the scientific and human levels; special thanks to **Melinda, Basile, Maïa, Jutzi, Helene, Alexis, Sophie, Agnes, Francois, Vanessa, Sylvain, Maïra, Dora, Christou, Adin, Myriam, Vincent and Jaqueline** and your **lovely children**. Thanks also to **Magali Whal** for our discussion on pedagogy, for your insights and for your support at every step. Thank you to **Muma**; by sharing your creative processes, you made merged science and art into a whole, translating some scientific explanations into realty. My deepest gratitude goes to **Mercedes, Isabelli, Maria, Valentine** and **Marlyse** for joyfully helping around with Jonas and Leonie, so I could dive headlong into my work without restriction. A special thanks to my dear friend **Sophie** for the last minute help with the kids, but also for all these long years of faithful support, for your clarity on life processes and for always being there, thank you.

I am lucky to have a heart family who played a key role in this process. Not only you made the Montessori pedagogy enter my life, shared your deep knowledge about it, opened the Montessori School Vevey's doors for all experiments, but also acted as a lighthouse in the storm; you showed up whenever I felt overwhelmed, brought me back to earth when ideas got blurry, made small victories sound like solid pillars for the next step and guided me wisely on this long journey toward my outer and inner research; a deep thank you to **Nadine, Pascale and Jean-Marc**. A deep thank you to **Gregoire**, for texting jokes and asking for news along the years, as well as for sharing your insights on life, thank you. My gratitude to **Sylvie**, for your delicious cookies that fed my stomach and soul during the whole thesis writing time, and for your support.

Finally, my gratitude goes to **Nicolas**, despite the roller-coaster nature of this journey, we learnt not to survive, but to live and ride the flow as it goes. I am not sure you expected such an adventure when you first pushed me to follow my dream of bridging Montessori and neuroscience, but thanks for doing it, and keeping supporting my choices despite often being out of our comfort zones! **Jonas**, you have been heavily involved in this work; since the day I told you that adults no longer remember most of their way of thinking as a child, you put so much effort into describing what was in your head and heart to help. **Leonie**, you came along and spontaneously started doing the same. You offered me so many insights into your world, with patience, simplicity and generosity. Thank you. I hope these years of hard work will leave you with memories not of my absence, but of the passionate woman I am, ready to be fully committed to my ideas, and that this will foster your journey toward yourself and your own passion(s) in life.

List of Abbreviations

ACC	Anterior Cingulate Cortex
CEN	Central Executive Network
dACC	dorsal Anterior Cingulate Cortex
DMN	Default Mode Network
EEG	Electroencephalographic
EF(s)	Executive Function(s)
ERN	error-related negativity
FC	Functional Connectivity
fMRI	functional Magnetic Resonance Imaging
Pe	error-related positivity
PES	post-error slowing
PIA	post-error improvement in accuracy
pMFC	posterior Medial Frontal Cortex
RL	reinforcement learning
rs-fcMRI	resting-state functional connectivity MRI
RT	Reaction Time
SN	Salience Network

Contents

1 Introduction

1.1 Skills for the 21st Century

We are facing rapid and substantial changes at the climatic, professional and social levels. Global warming will redraw the geographical contours, forcing large migratory movements. The advent of artificial intelligence will define a new professional landscape where purely executive jobs will no longer exist. At the social level, the individualist model is called upon to evolve toward a model in which cooperation more than competition dominates. These are significant challenges to be met, and skills such as flexibility, creativity and self-regulation have become even more important.

Professional organizations are already aware of these needs, they are trying to adapt their workplaces and train workers to improve their skills in this direction [Morieux, 2018]. Despite their efforts, however, changing adults' mental habits is not trivial. Indeed, the early years period presents the optimal plasticity for learning and developing such talents [Nelson and Bloom, 1997] [Gogtay et al., 2004]. This raises an important question: how do today's pedagogical practices in schools prepare tomorrow's adults for these critical skills and needs?

Current efforts in education focus on the curriculum's content on the one hand, and quantifying students' acquisitions on the other hand (e.g., formal assessment, grades, PISA [Grisay et al., 2007]). However, the above mentioned in-demand skills do not seem to result from supervised learning, nor can they be assessed using current metrics under the dominant pedagogical model in Western countries (the traditional system) [New, 2015]. Furthermore, the drawback of focusing solely on quantitative features is relegating to the background global and integrated child development, which may also be important to address the challenges ahead. Conversely, some alternative pedagogical approaches [Vygotsky, 1978] [Piaget, 1952] [Condliffe et al., 2017] do not target quantitative performance, but tend to address children's scholastic development in a more holistic way, such as the

Montessori pedagogy [Lillard, 2019]. These approaches aim to address the development of performance monitoring processes differently, with direct implications for a child's creative and cooperative abilities.

1.2 Montessori pedagogical practices

The Montessori pedagogy was born from years of empirical observations of developing children and their spontaneous self-directed activities. Students evolve within multi-grade classrooms (e.g., three to six years old children are working within the same class) equipped with a set of learning materials from which they are free to choose. These learning activities are self-corrective, which means that students discover the relevant concept on their own without requiring external feedback such as grades or other quantitative evaluations [Montessori, 1936]. They are allowed three uninterrupted hours of work at their own pace and are encouraged to cooperate with others [Lillard, 2011] [Ervin et al., 2010].

1.3 Quantitative studies of the Montessori pedagogy

The precited aspects, in isolation, are predictive of optimal development in children [Marshall, 2017]. However, quantitative studies on the effects of Montessori pedagogy are scarce. In general, Montessori students are reported to have increased scholastic outcomes, social understanding and development, mastery orientation, and well-being at school [Lillard and Else-Quest, 2006] [Lillard et al., 2017] [Rathunde and Csikszentmihalyi, 2005] [Ervin et al., 2010] [Ahmad and Reba, 2018] [Courtier, 2019]. A French study also reported advantages for Montessori pupils aged 6–10 years regarding both their divergent creativity (deriving new elements from a single element) and their convergent creativity (integrating diverse elements into a new, single element) [Besançon and Lubart, 2008]. Likewise, another study using the same task reported higher creativity skills in Montessori students [Fleming et al., 2019]. Moreover, children with lower socioeconomic status have been shown to catch up with their peers on scholastic outcomes within two years [Lillard et al., 2017].

While the majority of quantitative studies have confirmed the positive impact on scholastic, creative and social outcomes for Montessori-schooled students, these findings have been questioned [Lindenfors, 2007] [Mackinnon, 2007], found to be inconsistent across studies (e.g., higher math achievement [Dohrmann et al., 2007] versus higher language achievement [Peng and Md-Yunus, 2014]) or not reproduced (e.g., [Ruijs, 2017] [Lopota et al., 2005]). One possible explanation is the fidelity of implementation of the pedagogy. It appears to be a prerequisite to its success, as positive outcomes are reproducible as long as the pedagogy is strictly applied [Lillard, 2012]. Among the necessary criteria for high-fidelity implementation are the quality of teacher training, comprehensive didactic materials within the class environment, absence of grades or rewards, low teacher/pupil ratio, three hours of uninterrupted work and multi-grade classrooms [Montessori, 1936].

While positive outcomes are thought to reflect children's experiences with the pedagogical practices, the cognitive origins of these effects have not been studied yet. One dominant hypothesis is that Montessori-schooled students develop stronger executive functions (EFs) [Diamond and Lee, 2011], a set of higher cognitive abilities necessary for the direct monitoring of goals and planning of actions [Miyake et al., 2000]. In fact, many pedagogical features of Montessori specifically require EFs, such as the sequences of gestures to be memorized and repeated in new contexts, mindfulness activities, goal-directed movement and self-directed learning [Diamond, 2012] [Diamond, 2013] [Diamond, 2014]. Some research has specifically investigated EFs in Montessori students, reporting that five-year-old Montessori pupils achieve higher scores in a card-sorting task than their traditionally schooled peers [Lillard and Else-Quest, 2006]. A second longitudinal study of kindergarten children reported some effect of EFs over three years, but not as strong as one might have expected [Lillard et al., 2017]. An exploratory pre-test/post-test assessment in a small sample of Montessori preschoolers revealed an improvement that was correlated with the time spent within the Montessori environment rather than with age [Phillips-Silver and Daza, 2018]. However, it cannot be inferred that this advantage is specific to the Montessori setting, as these schoolchildren were from a single class, suggesting that a teacher effect could be at work.

In summary, there is clear evidence that Montessori schoolchildren score higher on scholastic tasks than traditionally schooled children, but without displaying a definite gain in EFs. Either the EF measures are not yet sensitive enough (e.g., measuring combined instead of separate core EFs) or these scholastic performance differences do not rely on EFs alone. One reasonable hypothesis is that the effects of Montessori school practices could partially stem from how children learn from erroneous responses, especially 4–13-year-olds. While traditional pedagogy teaches children by providing them information about when they have made a mistake so that they can avoid such mistakes in future, Montessori students are not given direct information about the correctness of their answers. Instead, Montessori teachers encourage children to notice their own incorrect thinking or help peers identify incorrect thinking in a pro-social manner [Montessori, 1936]. This could implicitly teach Montessori-schooled children to engage with errors and self-correction in a more autonomous, process-oriented and constructive way, while also helping raise their social skills. This may lead Montessori students to more readily guess and initially produce more incorrect responses; however, over time, they engage more effectively with errors to learn, a process called "performance monitoring."

1.4 Performance monitoring

1.4.1 Definition

The ability to flexibly adjust our behavior in response to a constantly changing environment is necessary for optimal goal achievement. This process is called performance monitoring and is depicted as a feedforward loop (see Fig. 1.1); unexpected events signal the necessity and form of adaptation required. It allows us to (i) quickly detect outcomes that deviate from expectation, (ii) integrate feedback, and (iii) fine-tune behavior and/or internal model for future decisions [Ullsperger et al., 2014a]. As a consequence, human beings can avoid repeating the same mistakes and evolve towards more appropriate behavior. It is not possible to directly measure performance monitoring itself. However, indirect informative markers are known and reported in the context of response error (Fig. 1.2).

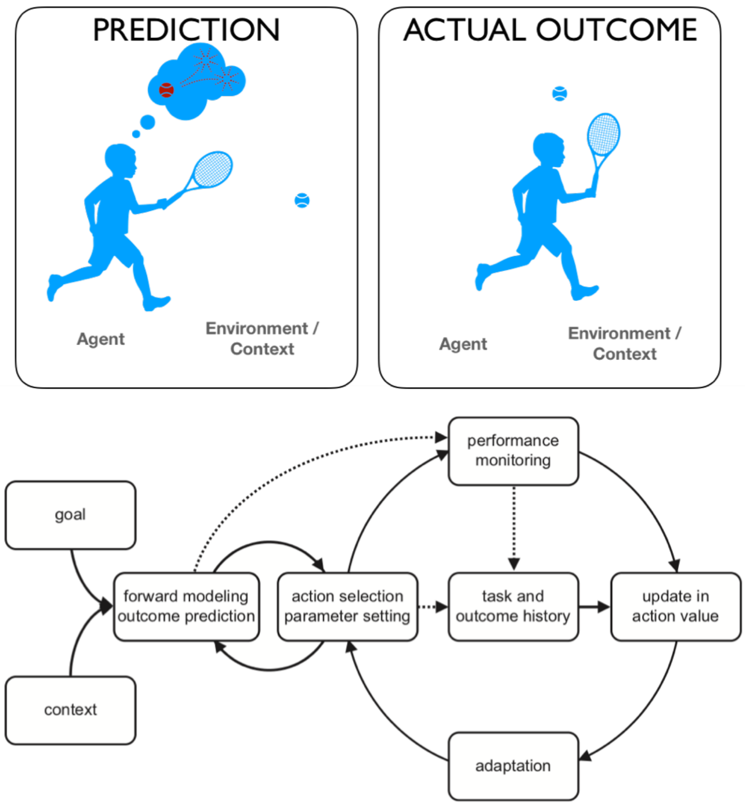


Figure 1.1 – Performance monitoring occurs when there is a mismatch between expected and actual outcomes (top panel). It is modeled as a feedforward loop; information resulting from unexpected event is integrated and used for later predictions (bottom panel, from [Ullsperger et al., 2014a]).

1.4.2 Behavioral measures

Post-error slowing

A large number of studies have investigated the cognitive mechanisms following errors by comparing the reaction time (RT) between a routine and an unexpected condition, such as response time in correct versus incorrect trials (Fig. 1.2 A). Any speeded task can be used for this neurophysiological measure. The Flanker task, for example, where participants track, respecting both speed and accuracy, the relevant stimuli while ignoring distractors [Eriksen and Eriksen, 1974]. Consistently, studies in adults report a higher differential RT for rare events, a phenomenon called Post-Error Slowing (PES) [Rabbitt, 1966].

PES amplitude correlates with different cognitive and emotional outcomes and relates to self-regulation capacity [Compton et al., 2011]. However, the root causes of PES are not clear yet and remain actively debated: is PES adaptive (enabling increase in post-error accuracy), or somewhat disruptive regarding subsequent performance? Indeed, slowed monitoring could stand for the higher cautious state that recruits selective attention and executive skills to optimize speed-accuracy tradeoff [Ridderinkhof et al., 2004]. However, very few studies found a positive correlation between post-error improvement in performance and PES [Danielmeier and Ullsperger, 2011]. Another view stands more for an orienting reflex [Notebaert et al., 2009], directly interfering with cognitive resources due to subsequently enlarged attention, which would not be task-relevant, thus referred to as maladaptive.

These opposite views could be more intertwined than previously thought; a recent study in humans and monkeys was able to explain PES as a reduced perceptual sensitivity counteracting increased decision boundaries and lowering the urgency signal [Purcell and Kiani, 2016]. There are known impact on visual perception, such as post-error blinding effect [Houtman and Notebaert, 2013], or pupillary responses reflecting arousal [Murphy et al., 2016]. [Ullsperger and Danielmeier, 2016] hypothesized that PES might not directly serve the prevention of future errors, but is rather a general phenomenon: an orienting reflex elicited by unexpected salient events reflecting a motivational component. They argued that a direct effect of PES might be not adaptive to the task at hand, but adaptive to resetting functional connectivity for neuronal adaptation. From a physiological perspective, noradrenaline release in forebrain structures impacts sensory processes by lowering their threshold, increasing at the same time cognitive flexibility and executive functioning, and offline memory consolidation [Sara and Bouret, 2012]. Altogether, these dual central-peripheral mechanisms (executive-arousal) allow cortical plasticity by resetting cortical networks. An increase in selective attention appears at a later stage to re-orient focus according to relevant stimuli dependent on cholinergic modulation [Ullsperger and Danielmeier, 2016].

Post-error improvement in accuracy

Another important marker of performance monitoring stands in post-error improvement in accuracy (PIA). It measures the number of correct actions that directly follow an incorrect response, over the total number of errors. PIA thus reflects the ability to bounce back from incorrect actions and come up with a correct one. PIA is an essential marker of behavioral adjustment that does not necessarily correlates with PES [Ullsperger and Danielmeier, 2016]. Reasons for poor PIA performance can be multiple, such as the emotional interferences (e.g., stress, anxiety), the affective bias toward error avoidance, or the lack of cognitive resources to overcome the adversity [Schroder and Moser, 2014].

1.4.3 Neural activity measures and underlying brain structures

Electroencephalographic (EEG) studies evaluated the time course of events following error responses (Fig. 1.2 B). A broad range of tasks emphasizing both speed and accuracy elicit the same brain response sequence in adults. Response-locked potentials of two components: an early frontocentral negativity (50 to 100 [ms] post-response), called error-related negativity (ERN), and a later and slower response (250-400 [ms]) more posterior, called error-related positivity (Pe). They respectively reflect an early task-unspecific detection of a need for adjustments (ERN), and a late task-specific selective attention orientation and learning (Pe) or conscious evaluation [Ullsperger et al., 2014a] [Ullsperger et al., 2014b] [Ullsperger and Danielmeier, 2016] [Danielmeier and Ullsperger, 2011]. Some psychological states alter the ERN amplitude suggesting dysfunction in underlying brain activity, such as in obsessive-compulsive disorder [Endrass and Ullsperger, 2014], social anxiety disorder [Endrass et al., 2014], anxiety or worry [Aarts and Pourtois, 2010] [Vocat et al., 2008] [Weinberg et al., 2010] [Hajcak et al., 2003], or mindsets [Moser et al., 2011]. Pe amplitude can be modulated by a happy mood (positivity), confirming motivational salience and/or conscious appraisal of errors [Paul et al., 2017].

Both source localization analyses and intracranial EEG recording, conclude on the role of distinct neural sub-regions in performance monitoring; ERN resulting mainly from the anterior cingulate cortex (ACC), and Pe from a more posterior region of the cingulate gyrus [Vocat et al., 2008] [Pourtois et al., 2010] [Vocat et al., 2008] (Fig. 1.2 C). Of note, the ACC contains a class of spindle-shaped neurons that is only seen in humans and great apes, suggesting a recent evolutionary specialization. Moreover, such cells appear postnatally, and their development is highly modulated by environmental conditions, such as stress [Allman et al., 2001]. EEG studies and functional magnetic resonance (fMRI) studies report activity within the cingulate gyrus, as well as the posterior medial frontal cortex (pMFC), the pre-supplementary motor area and dorsomedial prefrontal cortex (PFC) in the context of error compared to correct trials [Ullsperger et al., 2014b] [Danielmeier and Ullsperger, 2011]

[Danielmeier et al., 2011] [Carter et al., 2000] [Carter et al., 1998]. While early activation of the ACC may signal an unexpected event, such as surprising events [Wessel et al., 2012], the pMFC activity induces motor, sensory and social adaptation through a broad repertoire of processes [Danielmeier et al., 2011] [Laubach et al., 2015] [Izuma et al., 2015] [Klucharev et al., 2011].

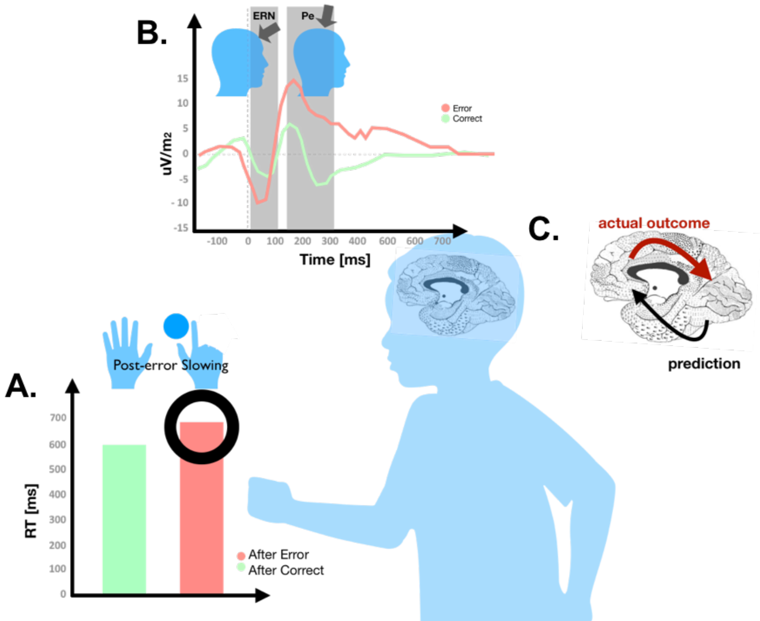


Figure 1.2 – Performance monitoring informative markers can be measured through (A) reaction time (RT) behavior (e.g., post-error slowing), and neural activation from (B) an electroencephalogram (e.g., error-related negativity -ERN- and positivity error -Pe-), or (C) from magnetic resonance imaging brain acquisitions (e.g., morphometric measures, functional activation, functional and/or structural connectivity) [Ullsperger et al., 2014a].

1.4.4 Affective aspects of performance monitoring: the self and the others

The affective processing of self-generated actions

Performance monitoring triggers cognitive reactions, but also social and emotional processes, as coactivation of ACC and amygdala in incorrect compared to correct response showed [Pourtois et al., 2010]. Previous research performed on adults showed that errors can even induce defensive motivational behavior [Hajcak and Foti, 2008]. In fact, response errors are processed as negative and aversive events in adults [Pourtois et al., 2010] [Koban et al., 2010] [Dignath et al., 2019]. A priming method (a speeded Go/noGo task coupled to an affective word categorization task) allowed to unveil this rapid and automatic reaction [Aarts et al., 2012]. Adult participants evaluated negative words faster and better than positive words after incorrect responses (see also [Aarts et al., 2013] [De Saedeleer and Pourtois, 2016] for replications), while barely assigning a positive value to correct responses. These findings suggest an asymmetry in the affective processing of self-generated actions in adult participants. Of note, the affective processing does not correlate with PES, implying that multiple intertwined components subserve performance monitoring at the adult age.

The social processing of performance monitoring

Witnessed errors also elicit the previously described performance monitoring markers [de Bruijn et al., 2009] [Kang et al., 2010] [Koban and Pourtois, 2014]. We learn from others incorrect responses by monitoring their actions [Metcalfe and Xu, 2018]. EEG and fMRI studies in humans robustly report activity in the pMFC when watching others performances. Human beings are thus able to learn from their social group (in the absence of direct genetic transmission), thanks to observational learning, no matter if socially or affectively driven [Bandura et al., 1963] [Call et al., 2005] [Clement and Dukes, 2013] [Meltzoff, 1995] [Nicol, 1995]. This process ingeniously optimizes learning efficiency and survival rate. However, the quality of the inter-personal context might play a role too. For instance, competitive or cooperative settings do not influence error-perception similarly. In both cases, self-generated errors elicited the previously described ERN. However, errors generated by others induced the same reaction only for cooperative tasks (observed ERN; oERN slightly delayed), whereas competition produced a later distinct one (smaller amplitude), therefore denoting different degrees of congruency with personal goals [Koban et al., 2010].

Altogether, these findings support coherently built-in competencies for learning from mistakes [Oudeyer et al., 2016], especially within a cooperative environment [Laubach et al., 2015].

1.4.5 The development of performance monitoring

We know from infant looking time and EEG studies that unexpected events are detected already by 2-3 months old babies [Dehaene-Lambertz, 2000] [Dehaene-Lambertz and Gliga, 2004]. However, measures of PES and ERN/Pe do not exist in children younger than three and four years old respectively [Jones et al., 2003] [Torpey et al., 2012], and tightly relates to the development of executive abilities [Downes et al., 2017]. Developmental studies are scarce and inconsistent in regard to the developmental trajectories of PES, ERN and Pe [Tamnes et al., 2013]. They sometimes show invariant [Davies et al., 2004] [Ladouceur et al., 2007] [Santesso et al., 2006] [Grammer et al., 2014], increasing [Santesso and Segalowitz, 2008] [Torpey et al., 2012] [Davies et al., 2004] or decreasing [Carrasco et al., 2013] [Hajcak et al., 2003] [Smulders et al., 2016] PES across age. The contradictory results on the developmental path may reflect task design features such as inter-stimulus timing or difficulty-level [Hogan et al., 2005] [Dutilh et al., 2012] or psychological factors such as anxiety [Torpey et al., 2012] [Meyer et al., 2013] [Meyer et al., 2012], or the social context of evaluation [Kim et al., 2005]. Still, instability of childrens performance monitoring system may also reflect the maturational change within the ACC [Smulders et al., 2016].

Performance monitoring highly depends on the ACC, which shows a protracted maturational path following an age-related caudal-ventral gradient of developmental change. Subregions of the ACC are associated with specific self-regulatory functions [Posner et al., 2007] [Margulies et al., 2007]: caudal (motor control), dorsal (attention/cognitive control), rostral (conflict monitoring), perigenual (mentalizing), and subgenual (emotional regulations). These cognitive competences are known to gradually become functional from early childhood to late adolescence, from motor to cognitive and socio-emotional control. A resting-state functional connectivity MRI (rs-fcMRI) study confirmed age-related differences in functional connectivity (FC) between the five precited subregions of the ACC, from diffuse in children, intermediate in adolescent, to higher and longer FC in adults. However, FC development was not uniform across functional networks: only the one associated with complex socio-emotional processing was significantly different in adults revealing plasticity that is still in process [Kelly et al., 2009]. The development of functional anatomy related to error processing was examined with fMRI in a population of 8-27 y.o., primarily focusing on ACC activity. First, rostral ACC activity was lower for correct versus incorrect responses, independently of age. Second, dorsal ACC (dACC) showed increased activity for incorrect response in adults compared to children and adolescents. Finally, adults recruited more PFC regions than posterior attentional areas [Velanova et al., 2008]. These studies corroborate others, suggesting that neural circuits for attentional control mature before those sub-serving social-emotional competencies [Tau and Peterson, 2010] (Fig. 1.3). Furthermore, it emphasizes that performance monitoring undergoes drastic and plastic change across childhood. It is possible that performance monitoring and its neural correlates are shaped by developmental experience, and in particular by schooling.

1.5 Thesis Outline

Pedagogical practices differ in the ways they support children in learning from their mistakes. Traditional approaches to schooling rely on teachers providing problems to students and providing feedback on their work. Learning outcomes are assessed using high-stakes tests graded for correct/incorrect answers. As students accommodate teachers feedback, they are discouraged from providing incorrect answers and encouraged to remember and reproduce correct answers. At the social level, students are grouped into classes with peers of a similar age and

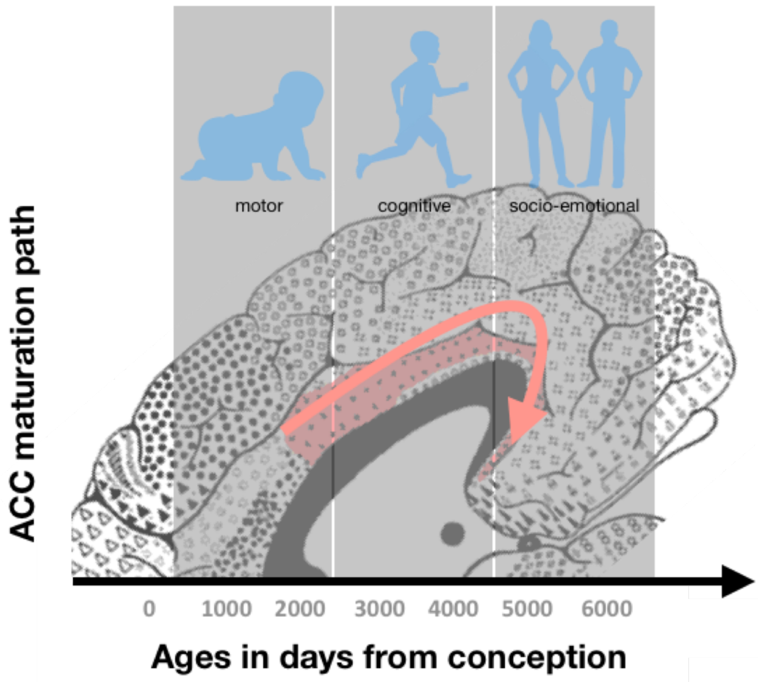


Figure 1.3 – The anterior cingulate cortex (ACC) is related to self-monitoring in general, different subregions related to specific self-regulatory functions; motor, cognitive and socio-emotional, that matures across age.

.

mainly interact during breaks for play. By contrast, in Montessori environments, teachers provide materials specifically designed to help children discover for themselves the underlying principles that lead to correct/incorrect outcomes in their work. Teachers avoid providing direct feedback about the correctness of childrens answers and instead focus on helping them engage properly with the materials to learn. The overarching aim is to help children actively learn and organize their own understanding. At the social level, students act in a socially cooperative manner in multi-grade classrooms, where they observe and sometimes teach each other, as the teacher/pupil ratio is deliberately low. These features may have a direct impact on performance monitoring development, with consequences on creativity and socio-emotional abilities.

1.5.1 The developmental cohort

To address this main developmental hypothesis, we recruited and organized a cohort of Montessori- and traditionally schooled subjects aged four to 15 years as well as adults. We adopted a multi-modal approach with behavioral as well as neuroimaging assessments to gain better insights into the underlying processes of performance monitoring. In total, 238 children and 55 adults were enrolled in the study. We controlled for between-group (Montessori and traditional) homogeneity at the age, fluid intelligence and socioeconomic levels, and created a questionnaire about parental pedagogical approach and home environment to control for selection bias as far as possible (see further discussion within chapter 3). The specific methodologies are detailed in each article.

All students underwent different behavioral measures (EFs, creativity, academic outcomes, well-being at school, post-error slowing). Some participants additionally undertook specific behavioral tasks to measure the affective bias post-error (speeded go/no go task with an affective categorization of words), understand their emotion recognition abilities (a social appraisal task and the offset emotional task) and characterize their multi-sensory capacities (a detection task and a multi-sensory continuous object recognition task). A subset of 65 respondents participated in an MRI scan session (with anatomical, diffusion, task, stimulus-directed, and resting-state fMRI acquisitions), and 36 also received an EEG (multi-sensory continuous recognition task). We only used adult participants for the affective bias post-error task. This dataset allowed us to answer the questions described in the subsequent section.

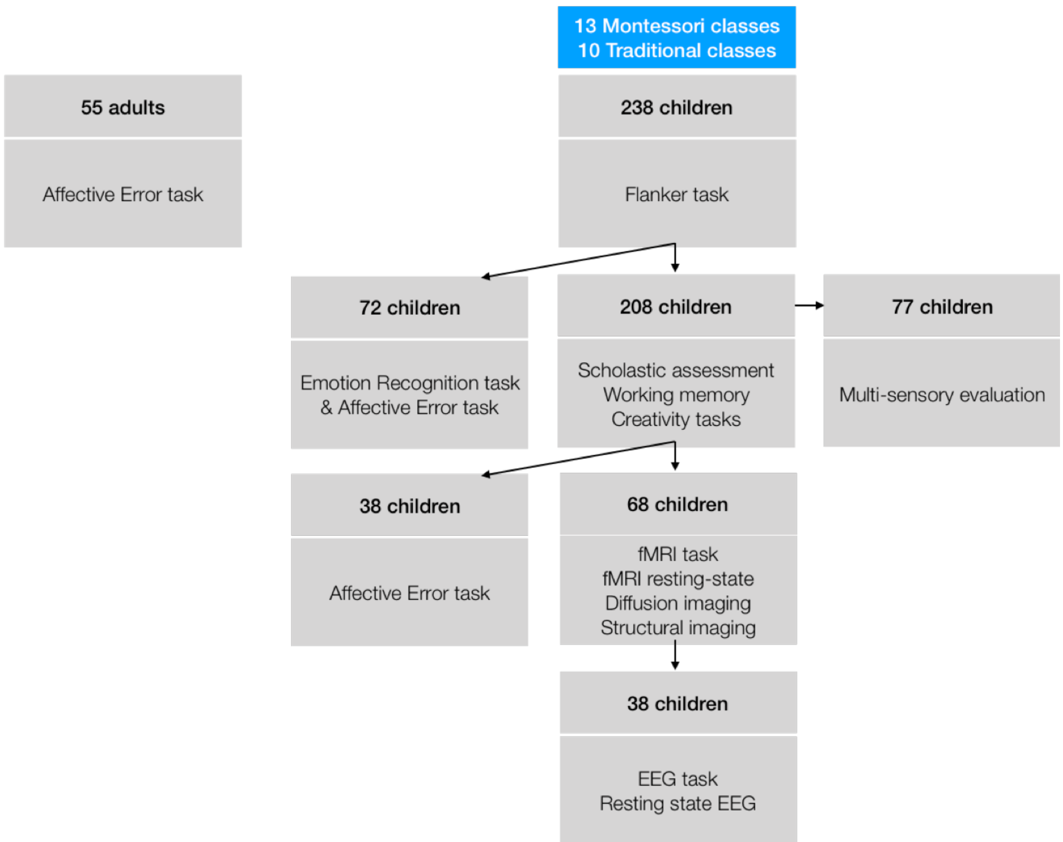


Figure 1.4 – Enrollment plan.

.

1.5.2 Research questions

During my Ph.D., I focused on specific research questions:

- **Replication study and EF investigation** The goal of this study was twofold: to replicate existing studies in another cultural context, namely Switzerland, and explore whether higher academic performance in Montessori-schooled students was attributable to higher EFs. As performance monitoring improvement parallels the development of executive abilities [Jones et al., 2003], we systematically investigated EF measures across ages and groups to understand their contribution to scholastic outcomes, such as other factor (creativity and well-being at school).
- **Multisensory integration profile** Some studies report a cascading effect from sensory functions to higher order cognitive abilities. Knowing the Montessori-schooled children make use of multiple senses through their learning materials, we investigated the extent to which perceptive abilities in the form of multisensory integration was more frequent than traditionally-schooled children. We further explored how these multisensory capacities were related to higher global cognitive outcomes (as measured through standardized tasks).
- **Performance monitoring characterization** The aim of these behavioral and neuroimaging studies was to explore the effects of age and pedagogical practices on the development of performance monitoring.

  PES and PIA Performance monitoring markers of PES and PIA were extracted from a standard and child-friendly flanker task. We further explored the PES-PIA relationship across age and schooling background.

  Cortical thickness and EEG components (ERN and Pe) relationship Morphological data were used to evaluate the development of cortical thickness in regions of interest within the ACC, and their relationship to ERN and Pe amplitudes. We further explored the pedagogical influences on this development.

  Error-monitoring fMRI task Neural activation and connectivity post-error responses were investigated through a tailored fMRI math task.

  Affective aspect of performance monitoring We explored whether response actions induced affective bias in children compared with adults and the extent to which pedagogical practices influence it. We customized an existing task to be child-friendly.
- **Socio-emotional competencies** Finally, we explored how pedagogical practices affect emotion recognition abilities as a more indirect social aspect of performance monitoring. We adapted a social appraisal task for this purpose and used another existing offset task.

2

Results

2.1 Summary of the findings

We conducted behavioral and neural studies of 4–15-year-old children in Switzerland exposed to high quality traditional and Montessori pedagogical practices. We replicated previous quantitative findings on scholastic development [Lillard et al., 2017] [Lillard and Else-Quest, 2006] [Besançon and Lubart, 2008], but in a different cultural environment. The effect sizes had a similar range to those of randomized studies [Lillard and Else-Quest, 2006] [Lillard et al., 2017], suggesting that some of the measured effects are attributable to Montessori pedagogical practices. We also found that higher academic outcomes did not relate to differences in EFs, but rather to higher creative thinking abilities (**2.2 Beyond executive functions, creativity skills benefit academic outcomes: Insights from Montessori education.**). We also showed that Montessori students benefit more from multi-sensory integration. We further report how multisensory integration has a direct relationship to children's cognitive abilities (**2.3 Multisensory gains in simple detection predict global cognition in schoolchildren**).

Performance monitoring was further explored. Behavioral measures revealed significant differences in how students notice, react and self-correct errors: Montessori students noticed errors younger and were better at self-correcting by adolescence than traditionally schooled students. We further observed a developmental shift across ages, with longer pauses associated with higher self-correction in younger children and shorter pauses predictive of better self-correction by adolescence (**2.4 Effects of traditional versus Montessori schooling on 4 to 15-year old childrens performance monitoring**).The morphometric data confirmed that the brain structures implicated in performance monitoring, namely the dorsal and rostral part of the ACC, undergo significant developmental changes during school years, correlated with ERN amplitude. The ERN component reflected the pedagogical practices experienced. These differences were also observed at the topographic level

(**2.5 Error-Monitoring is Modulated by School Pedagogy**). Brain activation was higher for correct than incorrect responses in an error-monitoring math task for both groups. However, the groups differed for other brain activity and the pattern of functional connectivity. On the one hand, traditionally schooled students had stronger FC between each seed region (cuneus, ACC, superior medial frontal) and the hippocampus on correct trials only. On the other hand, Montessori students had overall higher brain activity in brain regions implied in math, early visual and executive processes, as well as stronger connectivity between the ACC and superior frontal areas in error trials (**2.6 Error-monitoring is influenced by school pedagogy: fMRI evidence from Montessori and traditional schoolchildren**). In addition, we found children were biased toward correct responses, as opposed to incorrect ones, in contrast to adults. However, response actions elicited no affective bias for Montessori students, while traditionally schooled students were biased toward positive affect when behaving correctly (**2.7 Childrens automatic evaluation of self-generated actions is different from adults**). Finally, we also found differences in emotion perception; Montessori students more accurately integrated social contextual cues than traditionally-schooled students, and looked at positive emotional stimuli for longer (**2.8 Emotion recognition development: Preliminary evidence for an effect of school pedagogical practices**).

Together, these findings suggest that children's experience of performance monitoring differ to some extent from that of adults; correct responses elicit more brain activity and bias their responses. Further, performance monitoring seems to follow a non-linear developmental trajectory. Finally, there is evidence that error detection, evaluation and motivation differ in Montessori- and traditionally schooled children and that how 6–12-year-old children learn from errors is influenced by pedagogical practices.

2.2 Beyond executive functions, creativity skills benefit academic outcomes: Insights from Montessori education.

Denervaud, S., Knebel, J. F., Hagmann, P., & Gentaz, E. (2019). Beyond executive functions, creativity skills benefit academic outcomes: Insights from Montessori education. *PLoS One, 14*(11), e0225319. doi:10.1371/journal.pone.0225319

2.2.1 Abstract

Studies have shown scholastic, creative, and social benefits of Montessori education, benefits that were hypothesized to result from better executive functioning on the part of those so educated. As these previous studies have not reported consistent outcomes supporting this idea, we therefore evaluated scholastic development in a cross-sectional study of kindergarten and elementary school-age students, with an emphasis on the three core executive measures of cognitive flexibility, working memory update, and selective attention (inhibition). Two hundred and one (201) children underwent a complete assessment: half of the participants were from Montessori settings, while the other half were controls from traditional schools. The results confirmed that Montessori participants outperformed peers from traditional schools both in academic outcomes and in creativity skills across age groups and in self-reported well-being at school at kindergarten age. No differences were found in global executive functions, except working memory. Moreover, a multiple mediations model revealed a significant impact of creative skills on academic outcomes influenced by the school experience. These results shed light on the possibly overestimated contribution of executive functions as the main contributor to scholastic success of Montessori students and call for further investigation. Here, we propose that Montessori school-age children benefit instead from a more balanced development stemming from self-directed creative execution.

2.2.2 Personal Contribution

Conceptualization, Recruitment, Data curation, Formal analysis, Investigation, Project administration, Visualization, Writing original draft.

2.3 Multisensory gains in simple detection predict global cognition in schoolchildren.

Denervaud, S., Gentaz, E., Matusz, P.J.*, & Murray, M.M.* (*in revision*). Multisensory gains in simple detection predict global cognition in schoolchildren. *Scientific Report*

**both authors contributed equally*

2.3.1 Abstract

The capacity to integrate information from different senses is central for coherent perception across the lifespan from infancy onwards. Later in life, multisensory processes are related to cognitive functions, such as speech or social communication. During learning, multisensory processes can in fact enhance subsequent recognition memory for unisensory objects. These benefits can even be predicted; adults recognition memory performance is shaped by earlier responses in the same task to multisensory but not unisensory information. Everyday environments where learning occurs, such as classrooms, are inherently multisensory in nature. Multisensory processes may therefore scaffold healthy cognitive development. Here, we provide the first evidence of a predictive relationship between multisensory benefits in simple detection and higher- level cognition that is present already in schoolchildren. Multiple regression analyses indicated that the extent to which a child (N=68; aged 4.5 15years) exhibited multisensory benefits on a simple detection task not only predicted benefits on a continuous recognition task involving naturalistic objects (p=0.009), even when controlling for age. The same relative multisensory benefit also predicted working memory scores (p=0.023) and fluid intelligence scores (p=0.033) as measured using age- standardised test batteries. By contrast, gains in unisensory detection did not show significant prediction of any of the above global cognition measures. Our findings show that low-level multisensory processes predict higher-order memory and cognition already during childhood, even if still subject to ongoing maturation. These results call for revision of traditional models of cognitive development (and likely also education) to account for the role of multisensory processing, while also opening exciting opportunities to facilitate early learning through multisensory programs. More generally, these data suggest that a simple detection task could provide direct insights into the integrity of global cognition in schoolchildren and could be further developed as a readily-implemented and cost-effective screening tool for neurodevelopmental disorders,particularly in cases when standard neuropsychological tests are infeasible or unavailable.

2.3.2 Personal Contribution

Conceptualization, Recruitment, Data curation, Formal analysis, Investigation, Project administration, Visualization, Writing original draft.

2.4 Effects of traditional versus Montessori schooling on 4 to 15-year old childrens performance monitoring.

Denervaud, S., Knebel, J. F., Immordino-Yang, M.H.*, & Hagmann, P.* (2020). Effects of traditional versus Montessori schooling on 4 to 15-year old childrens performance monitoring. *Mind, Brain and Education*, doi:10.1111/mbe.12233

**both authors contributed equally*

2.4.1 Abstract

Through performance monitoring individuals detect and learn from unexpected outcomes, indexed by post-error slowing and post-error improvement in accuracy. Though performance monitoring is essential for academic learning and improves across childhood, its susceptibility to educational influences has not been studied. Here we compared performance monitoring on a flanker task in 234 children aged 4 through 15, from traditional or Montessori classrooms. While traditional classrooms emphasize that students learn from teachers feedback, Montessori classrooms encourage students to work independently with materials specially designed to support learners discovering errors for themselves. We found that Montessori students paused longer post-error in early childhood and, by adolescence, were more likely to self-correct. We also found that a developmental shift from longer to shorter pauses post-error being associated with self-correction happened younger in the Montessori group. Our findings provide preliminary evidence that educational experience influences performance monitoring, with implications for neural development, learning and pedagogy.

2.4.2 Personal Contribution

Conceptualization, Recruitment, Data curation, Formal analysis, Investigation, Project administration, Visualization, Writing original draft.

2.5 Error-Monitoring is Modulated by School Pedagogy.

Denervaud, S.*, Knebel, J.F.*, Mullier, E., Hagmann, P., & Murray, M.M. (*internal revisions*). Error-Monitoring is Modulated by School Pedagogy.

**both authors contributed equally*

2.5.1 Abstract

Within an inherently dynamic environment, unexpected outcomes are part of daily life. Performance monitoring allows us to detect these events and adjust behavior accordingly. The necessity of such an optimal functioning has made error-monitoring a prominent topic of research over the last decades. Event-related potentials (ERPs) have differentiated between two brain components involved in error-monitoring: the error-related negativity (ERN) and error-related positivity (Pe) that are thought to reflect detection vs. emotional/motivational processing of errors, respectively. Both ERN and Pe depend on the protracted maturation of the frontal cortices and anterior cingulate through adolescence. To our knowledge, the impact of schooling pedagogy on error-monitoring and its brain mechanisms remains unknown and was the focus of the present study. Swiss schoolchildren completed a continuous recognition task while 64-channel EEG was recorded and later analyzed within an electrical neuroimaging framework. They were enrolled either in a Montessori curriculum (N=13), consisting of self-directed learning through trial-and-error activities with sensory materials, or a traditional curriculum (N=14), focused on externally driven activities mainly based onreward feedback. The two groups were controlled for age, gender, socio-economic status, parental educational style, and scores of fluid intelligence. The ERN was significantly enhanced in Montessori schoolchildren (driven by a larger response to errors), with source estimation differences localized to the cuneus and precuneus. In contrast, the Pe was enhanced in traditional schoolchildren (driven by a larger response to correct trials), with source estimation differences localized to the ventral anterior cingulate. Receiver operating characteristic (ROC) analysis demonstrated that the ERN and Pe could reliably classify if a child was following a Montessori or traditional curriculum. Brain activity subserving error-monitoring is modulated differently according to school pedagogy.

2.5.2 Personal Contribution

Conceptualization, Task adaptation, Recruitment, Data curation, Formal analysis, Investigation, Project administration, Visualization, Writing original draft.

2.5.3 Follow-up analyses

We further investigated cortical thickness in specific subregions of the ACC (posterior, dorsal and rostral) accross age, and their relationship to both ERN and Pe amplitude.

MRI acquisition.
Structural imaging was collected at the Lemanic Biomedical Imaging Center (CIBM) of the University Hospital Lausanne (CHUV), on a Siemens 3T Prisma-Fit MR scanner, with a 64-channel head-coil. For each participant, a 3-dimensional high-resolution isotropic T1-weighted sequence (MPRAGE) was acquired (TR = 2000ms, TE = 2.47 ms, 208 slices; voxel size= 1x1x1, flip angle=8°) as anatomical individual reference.

Image processing.
Individual T1-weighted sequence were processed using the FreeSurfer 5.1.0 software (surfer.nmr.mgh.harvard.edu), a commonly-used open-source software package. The pipeline includes a first step of brain extraction, automated Talaraich transformation, brain tissue segmentation into white and grey matter and their boundaries, automated topological correction and deformation, and parcellation of cerebral cortex into anatomical structures automatically labelled into 34 regions per hemisphere. The second step computes voxels within each regions to derive morphometric measures. For the current study, we focused on the cortical thickness measures of predefined sub-regions of the ACC (posterior, caudal, rostral).

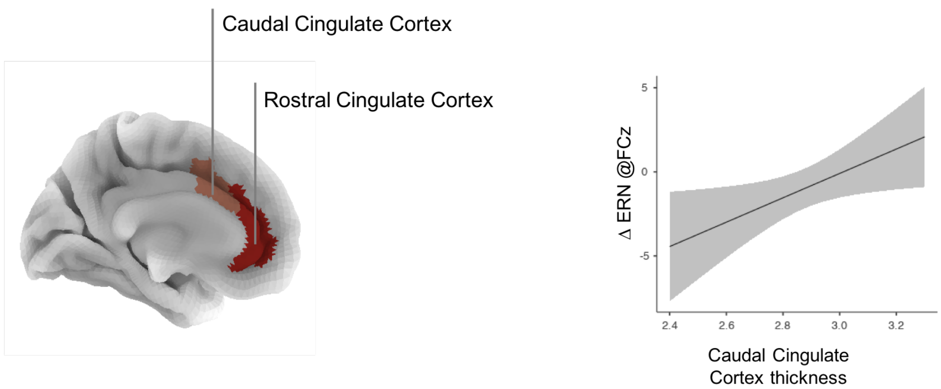


Figure 2.1 – There is significant cortical thinning (maturation) between 6-12 yo. within both the caudal (p =0.046) and rostral (p =0.052) subregions of the ACC (left panel), and thickness of the dorsal part of the ACC was predictive of the children's error-related negativity amplitude (t= 2.37, p =0.029; right panel).

Statistical analyses.
All neuroanatomical measures were examined for normality using the Shapiro-Wilk test. To statistically evaluate the effect of age on cortical thickness, as well as potential group differences (Montessori vs. traditional), we ran a multivariate analysis of covariance (MANCOVA) on the posterior, caudal and rostral subregions of the ACC (left and right). Finally, we performed regression analyses to investigate which factors between ACC subregions' cortical thickness, age, and group (Montessori, traditional) predicted error-related components (ERN and Pe). For these analyses, we considered significant results with a $p < 0.05$.

Results.
There was a significant effect of age on cortical thickness of the caudal part of the ACC ($F(23, 1)= 4.47$, $p = 0.046$), and near significant in the rostral part of the ACC ($F(23, 1)= 4.21$, $p = 0.052$). There also was a group difference in cortical thickness in the posterior subregion of the ACC ($F(23, 1)= 7.52$, $p = 0.012$). Finally, the ERN amplitude was reliably predicted by age ($t= 2.55$, $p = 0.020$), and dorsal ACC cortical thickness ($t= 2.37$, $p = 0.029$).

These preliminary findings further suggest a direct relationship between maturational change within the dorsal region of the ACC and error-monitoring abilities development.

2.6 Error-monitoring is influenced by school pedagogy: fMRI evidence from Montessori and traditional schoolchildren.

Denervaud, S., Fornari, E., Yang, X.F., Hagmann, P., Immordino-Yang, M.H.*, & Sander, D.* (*submitted*). Error-monitoring is influenced by school pedagogy: fMRI evidence from Montessori and traditional schoolchildren. **both authors contributed equally*

2.6.1 Abstract

The development of error-monitoring is central to learning and academic achievement. However, few studies exist on the neural correlates of childrens error- monitoring, and no studies have examined its susceptibility to educational influences. Pedagogical methods differ on how they teach children to learn from errors. Here, 32 students (aged 8-12 years) from high quality Swiss traditional or Montessori schools performed a math task with feedback during fMRI. Although the groups accuracies were similar, Montessori students skipped fewer trials, responded faster and showed more neural activity in right parietal and frontal regions involved in math processing. While traditionally-schooled students showed greater functional connectivity between the ACC, involved in error-monitoring, and hippocampus following correct trials, Montessori students showed greater functional connectivity between the ACC and frontal regions following incorrect trials. The findings suggest that pedagogical experience influences the development of error- monitoring and its neural correlates, with implications for neurodevelopment and education.

2.6.2 Personal Contribution

Conceptualization, Task development, Recruitment, Data curation, Formal analysis, Investigation, Project administration, Visualization, Writing original draft.

2.7 Childrens automatic evaluation of self-generated actions is different from adults.

Denervaud, S., Hess, A., Sander, D., & Pourtois, G. (*in review*). Childrens automatic evaluation of self-generated actions is different from adults. *Developmental Science*

2.7.1 Abstract

Performance monitoring (PM) is an important cognitive process enabling to swiftly detect mismatches between goal and action (such as response errors), as well as trigger subsequently behavioral adaptation (such as post error slowing - PES). Previous research showed that response errors are automatically processed as negative events in adults. Furthermore, this evaluative process could be dissociated from the subsequent behavioral adjustment captured by the PES, suggesting that PM is subserved by multiple components at the adult age. However, it remains unclear whether (i) the evaluation of errors as negative events is present at an earlier stage in life, namely childhood, and (ii) whether PM processes are fixed, or shaped by the environment. To address these questions, we tested the affective processing of self-generated actions and the PES in 8-12 years old schoolchildren enrolled either in traditional (N=56) or Montessori (N=45) schools, and adults (N=46). Results showed that children processed correct actions as positive events whereas adults processed errors as negative events. Moreover, this former effect was observed in traditional schoolchildren, but not in Montessori schoolchildren. In comparison, the PES was not modulated by age, nor pedagogy. These findings suggest that unlike the PES that likely reflects an age-invariant automatic attention orienting towards response errors, the affective processing of actions depends on both age and school environment.

2.7.2 Personal Contribution

Conceptualization, Task adaptation, Recruitment, Data curation, Formal analysis, Investigation, Project administration, Visualization, Writing original draft.

2.8 Emotion recognition development: Preliminary evidence for an effect of school pedagogical practices.

Denervaud, S., Mumenthaler, C., Gentaz, E., & Sander, D. (*in revision*). Emotion recognition development: Preliminary evidence for an effect of school pedagogical practices. *Journal of Learning and Instruction*

2.8.1 Abstract

Emotion recognition is shaped through social interactions from a childs early years through at least late adolescence. Experiences of culture or early life adversity have been reported to impact emotion recognition competencies. However, no emphasis thus far has been given to the effects of daily experience at school. Enriched and more diverse social interactions, such as fostered by some pedagogical practices, may contribute to emotion recognition competencies. Here, we investigated differences among schoolchildren experiencing Montessori versus traditional way of learning. Children were asked to categorize an ambiguous 50%fear-50%surprise face while the context was manipulated, and, in another task, to track positive-negative dynamic emotional changes. Results suggest that children experiencing traditional practices show higher fear recognition sensitivity. Conversely, schoolchildren experiencing Montessori practices integrated more social cues and perceived longer positive emotion. Such preliminary findings call for further research to determine both the extent to which pedagogy causes these effects, and their underlying mechanisms.

2.8.2 Personal Contribution

Conceptualization, Task adaptation, Recruitment, Data curation, Formal analysis, Investigation, Project administration, Visualization, Writing original draft.

3
Discussion

3.1 Interpretation and implications

The cognitive skills needed to face the rapidly changing 21st-century environment successfully include creativity, flexibility, and self-regulation. Central to these competencies stands the self-monitoring system, for which the ACC plays a central role and undergoes protracted development from early childhood to late adolescence (both at the structural and functional levels) [Posner et al., 2007]. An unanswered question is the extent to which pedagogical practices sustain its development. Specifically, the dorsal part of the ACC matures from around 6–12 years, along with EF maturation, improving the capacity to detect, evaluate and self-correct errors [Kelly et al., 2009] [Velanova et al., 2008]. The susceptibility of performance monitoring to pedagogical influence was the topic of this work, inspected through seven different studies. A detailed discussion for each study is provided within its related publication (see Annex). Here, I aim to broaden the reflection and provide a more global interpretation of the findings and their potential implications as well as possible future research directions.

3.1.1 What the switches and cascading effects suggest

Our findings corroborate previous studies of switches in learning signals [Peters et al., 2016] [van den Bos et al., 2009] [Zanolie et al., 2008] [Crone and van der Molen, 2007] [Crone et al., 2004] as well as those of cascading effects from sensory to cognitive processes [Rose et al., 2008]. Together, they reconcile some of the inconsistencies reported in the developmental literature on performance monitoring [Smulders et al., 2016]. In particular, depending on the age range studied and pedagogical training experienced, markers of performance monitoring may vary or not and they may look more or less similar toan adult pattern. Hence, the implications of a dynamic and flexible view of performance monitoring development are twofold: (i) non-linear development and (ii) the possibility of an optimal window of opportunity to acquire specific information (e.g., associate action with feedback).

Non-linear development

Post-mortem and neuroimaging studies have consistently reported the sequential and asynchronous organization of brain development. Primary sensory-motor regions mature first, followed by the adjacent unimodal associative cortices and, finally, the higher-order associative areas [Huttenlocher and Dabholkar, 1997] [Lebenberg et al., 2018] [Friedrichs-Maeder et al., 2017]. The developmental pattern observed in our data suggests that performance monitoring occurs in similarly sequential steps. First, the child seems to learn to detect discrepancy (higher PIA for greater PES), a process that multi-sensory processes may scaffold (higher multisensory gain related to higher cognitive abilities). Then, s/he seems to associate positive feedback with adequate behavior (higher brain activity for correct responses). Once these associations are integrated, the child may predict the causality of his/her actions and thus learn from his/her errors (higher PIA for slower PES). We further show that each step could benefit from pedagogical practices, as, for example, Montessori students detect errors earlier than traditionally schooled students.

This non-linear development could be tightly related to consciousness development, or the gaining of awareness about the physical, mental and social boundaries over time [Lewis, 2003]. This development may lead to successive, relevant information being processed as salient by the child [Uddin et al., 2011] [Menon and Uddin, 2010], before being automatized (unconsciously processed). In the performance monitoring framework, the child first needs to learn to detect incorrect actions, to raise awareness on discrepancy. Whenever this step is acquired, brain activity may shift from diffuse prefrontal areas to more segregated and integrated networks [Fair et al., 2007]; the child will then automatically react to unexpected events, while consciously integrating feedback thereafter [Metcalfe, 2017]. This idea aligns with the cascading effect measured in our study as well as in previous work: the quality of early awareness could predict later cognitive competencies. This dynamic may be recursive for learning to self-monitor behavior at the motor, cognitive and social level following the maturation path of the ACC [Kelly et al., 2009] [Posner et al., 2007] [Velanova et al., 2008].

This idea of consciousness also aligns with studies showing the effectiveness of error management training in adults through either mindfulness sessions (ref; interoceptive awareness) or reframing the error as a great opportunity for learning (ref; be aware that error avoidance is not an adaptive strategy) [Keith and Frese, 2005] [Keith and Frese, 2008] [Frese and Keith, 2015] [Smart and Segalowitz, 2017].

A possible window of opportunity

A window of opportunity is a neurodevelopmental stage in which neuronal activity is particularly sensitive to environmental cues, characterized by switches (the opening and closing of the sensitive period) and experience-dependent plasticity [Carcea and Froemke, 2013]. While the windows of opportunity for senses and language are well known and defined [Huttenlocher and Dabholkar, 1997] [Hensch, 2005a] (see Fig. 3.1, left panel), more recent studies report the same process to be true for higher-order executive functioning aspects [Tottenham and Gabard-Durnam, 2017]. For example, the fear response is highly modulated by maternal care in childhood (up to 12 years) but not adolescence, unveiling a sensitive period for the amygdala–prefrontal cortex connectivity and its related self-regulatory component [Callaghan and Tottenham, 2016]. Our findings provide compelling evidence for the presence of another window of opportunity for performance monitoring, that is self-regulation in the face of an unexpected event (e.g., error). If further confirmed, this would explain not only the switches and the influences of pedagogical practices, but also children's orientation to trial-and-error behavior.

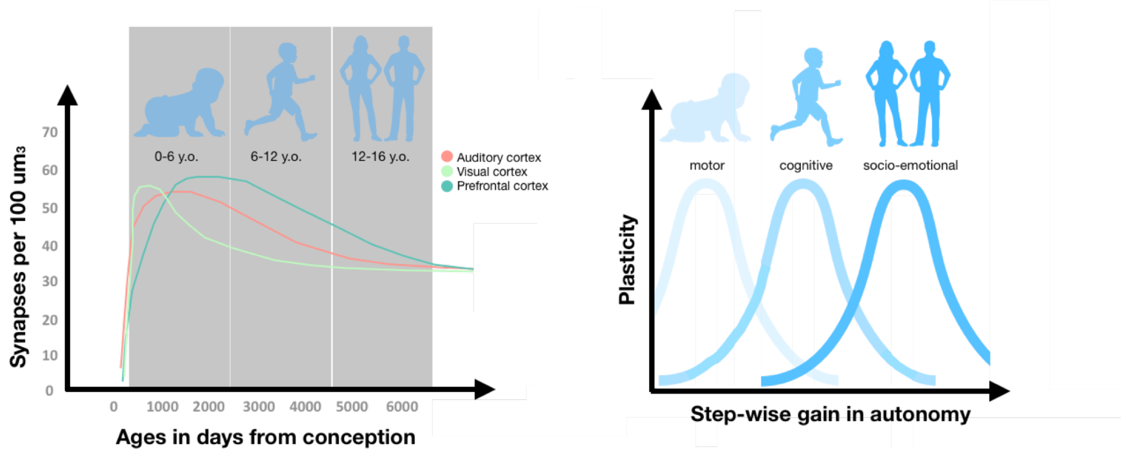


Figure 3.1 – Synaptogenesis across age was reported to be sequential; from sensory to prefrontal cortices (left panel, adapted from [Huttenlocher and Dabholkar, 1997]). Sensory modalities, as well as association between sensory abilities related to language, are known to be related to specific and temporally-defined windows of opportunity [Hensch and Fagiolini, 2005] [Hensch, 2005a]. There are preliminary evidence for windows of opportunities related to higher-order socio-emotional competencies [Tottenham and Gabard-Durnam, 2017]. We hypothesize that these successive windows of opportunity are meant to help the child acquire specific abilities toward autonomy and self-directed behavior (right panel). If true, performance monitoring may rely on such a specific time-window in development, related to the anterior cingulate cortex maturation [Velanova et al., 2008]

How and why switches exist. The cellular mechanisms underlying the windows of opportunity were shown for sensory development. They partially result from (i) released target signals that (ii) initiate a molecular cascade allowing inhibitory cells (interneurons) to enter maturation, where they (iii) show synapse proliferation similarly to inhibiting cells (non-specific) and finally (iv) pruning redundant or aberrant connections [Hensch and Fagiolini, 2005] [Hensch, 2005b] [Hensch, 2004]. Hence, a newly shaped receptive map results from the most frequently activated pattern. The maturation of interneurons balances the inhibitory-excitatory signal [Carcea and Froemke, 2013]. Plasticity switches off when the newly carved whole network is shaped. Both structural (peri-neuronal nets or myelin) and functional brakes ensure long-term stability [Bavelier et al., 2010] [Hensch and Bilimoria, 2012] [Thomas and Johnson, 2008]. Myelinization takes place to convey neural signals efficiently from one cortical region to another [Dubois et al., 2012]. Spontaneous brain activity at rest strengthens the connections between recently co-activated cortical areas, ensuring the integrity of the existing network [Vogel et al., 2010]. In the context of performance monitoring, we also report switches that could relate to how learning signals are first processed for positive feedback, but no longer when older, as ACC-related FC across ages partially denotes an experience-dependent organization.

Maturational change in the ACC and behavior. The relationship between neural dynamics and psychological/behavioral outcomes is poorly understood. However, some evidence shows that cortical [Tau and Peterson, 2010] and white matter maturation [Dubois et al., 2009] coincides with related cognitive functions. In addition, genotype–environment correlational studies report a relationship between neurodevelopment and behavior. Spontaneously chosen activities tend to reflect and/or match genetic developmental programs [Plomin, 2014] [Plomin et al., 2014]. While we observe a maturational change in the dorsal and rostral parts of the ACC, children may have a particular orientation for trial-and-error learning at both the cognitive and the social levels [Margulies et al., 2007] [Posner et al., 2007] (see Fig. 3.1, right panel). If pedagogical practices follow this tendency, as Montessori pedagogy does, students make more errors and inhibit less, leading them to develop strategies to face adversity and self-regulate over time [Metcalfe, 2017].

3.1.2 The pedagogical approach to feedback

The different pedagogical traditions may actually shape performance monitoring by influencing specific reinforcement learning (RL) parameters. From this perspective, it is relevant to recall that in the performance monitoring loop, incorrect actions are perceived as a mismatch between the expected and the actual outcome. Based on RL theories, the mismatch can be either model-free or model-based [Dolan and Dayan, 2013] [Glascher et al., 2010] [Neftci and Averbeck, 2019]. Whereas the former refers to how the value for a given outcome is computed and expected, the latter relates to a state transition for it, thereby enabling the optimal processing of sensory cues (unvalued) (see Fig. 3.2).

Performance monitoring is thus tightly related to the type of feedback received post-error. The pedagogical practices studied in this work differ largely in that

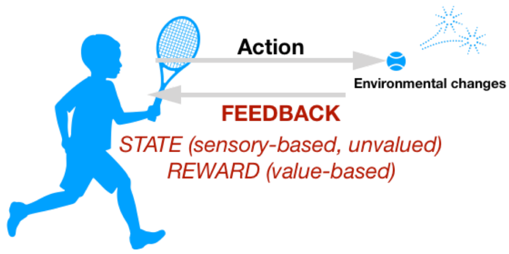


Figure 3.2 – The mismatch between expected and actual outcomes result in an error prediction signal that can be either valued (an expected quantity is not received, or higher than expected), or unvalued (change in state). These feedback are integrated within our internal model and used for forward modeling in later outcome predictions [Ullsperger et al., 2014a] [Dolan and Dayan, 2013]
.

aspect. Specifically, Montessori-schooled students rarely face evaluative feedback (from the teacher), grades or punishment for incorrect actions or behaviors. On the contrary, they have to self-discover relevant concepts using didactic material. In addition to the cognitive aspect of learning, there is an underestimated social component. Montessori pedagogical practices encourage the child to teach and learn from peers in multi-grade classrooms. Montessori students may thus rely more on unvalued feedback than traditionally-schooled students do [Lillard, 2012] [Rathunde and Csikszentmihalyi, 2005]. Indeed, we observed Montessori students detecting errors as early as four years old and they were also better at integrating multisensory cues. These findings may be intertwined. The Montessori curriculum offers children a sensory education from three years [Montessori, 1936] [Marshall, 2017] [Lillard, 2012]. In parallel, children start learning by doing; each incorrect action has a direct consequence that the child needs to fix. The child is invited to figure it out by him/herself. Therefore, as early as three years, the child trains to perceive his/her environment accurately and fix his/her incorrect actions autonomously through sensory feedback. By the age of 8–12 years, Montessori students showed no affective bias after their action responses. Together, these findings suggest that they may rely more on model-based RL [Shadmehr et al., 2010]. As the child certainly makes sense of his/her experience (e.g., incorrect actions), according to the feedback received and its weighting, both the academic materials and the social cues conveyed from the environment could have long-term implications for learning and curiosity [Oudeyer et al., 2016] [Oudeyer and Smith, 2016] [Gottlieb and Oudeyer, 2018] [Gottlieb et al., 2014].

In adults, this framework is extremely valuable, as it can account for a wide range of phenomena during RL, including the modulatory effects of feedback type and reward for it [Mattar et al., 2018]. The higher frequency of value feedback in 4–12-year-olds could relate to some extent to error avoidance in adulthood [Aarts et al., 2012]. Accordingly, it would be extremely informative for future studies to link more directly changes in performance monitoring to possible alterations of specific RL parameters.

Preliminary results on feedback processing

We analyzed the fMRI math task by focusing on different reward conditions. The participant solved the math problems; feedback were either in the form of points ("Point"), toy reward ("Toy"), or nothing ("No reward"). At the behavioral level, students performed better when the math problem was presented to "train," without reward, independently of the schooling background. We extracted the BOLD signals of the 180 seconds where the participant was playing for one of the three conditions, and computed the related FC matrices (correlation of brain activity between 128 cortical regions). We found significant differences between Montessori and traditionally-schooled students within each condition (Fig. 3.3). These findings provide preliminary evidence of a modulatory effect of feedback on neural connectivity in 8–12-year-old students.

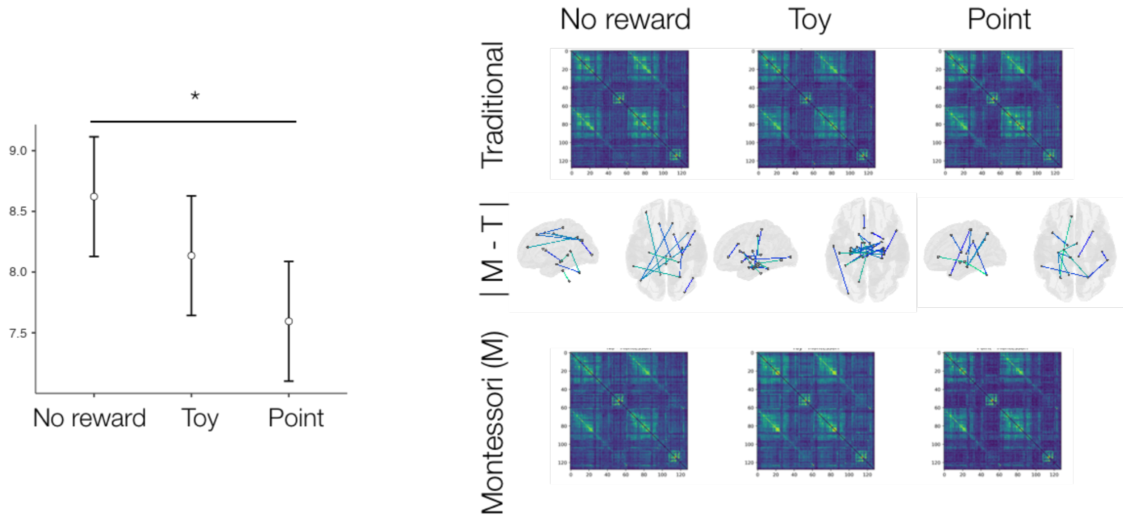


Figure 3.3 – At the behavioral level (left panel), students had higher accuracy when playing for "no reward" than for "points." At the functional connectivity level, connectivity matrices per condition and schooling system are displayed (right panel, up and down). Differences in functional connectivity were computed (absolute value of differences in FC between M and T) and plotted on a glass brain (right panel, middle).

3.1.3 Implications for cognitive flexibility

Central to the performance monitoring system is the dorsal part of the ACC. This subregion of the ACC, together with the anterior insula, form the salience network (SN). The SN acts as a switch turning off the default mode network (DMN) in the profit of the central executive network (CEN) [Menon and Uddin, 2010] (see Fig. 3.4, left panel). In the context of performance monitoring, when incorrect actions are more infrequent than correct responses [Ullsperger et al., 2014b] [Wessel et al., 2012], they undoubtedly trigger the SN, explaining the robust activation of the dACC across studies. The switching function then segregates intra- from extra-personal stimuli in a twofold mechanism: (i) fast and automatic perceptive bottom-up signaling and (ii) a higher-order system related to context-specific cues enhancing access to goal-directed actions [Menon, 2015]. Pediatric studies report the SN from the age of two. Still, studies show that its protracted development undergoes FC changes between seven and 20 years old [Uddin et al., 2011], suggesting that maturation responds to the environment for the optimal calibration of flexible control [Ryali et al., 2016] [Menon and Uddin, 2010].

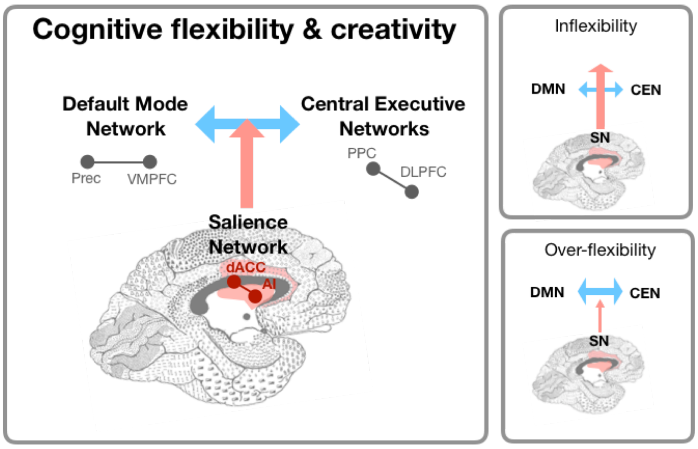


Figure 3.4 – Functional connectivity fluctuates over time between the different neural networks. Optimal coupling between the default mode network (DMN) and the central executive network (CEN) is orchestrated by the salience network (SN) (left panel). Dysregulation of the SN result in cognitive inflexibility or over-flexibility (right panel). Nodes for these networks are typically: Precuneus (Prec) and ventro-medial prefrontal cortex (VMPFC) for the DMN; dorsal anterior cingulate cortex (dACC) and anterior insula (AI) for the SN; posterior parietal cingulate (PCC) and dorso-lateral prefrontal cortex (DLPFC) for the CEN. Adapted from [Uddin, 2015] [Menon and Uddin, 2010].

The calibration of the SN has drastic consequences on how flexibly an adult acts and thinks. A growing body of research shows that over-activation of the SN leads to excessive inhibition, while weaky SN control impacts attentional and executive abilities [Uddin et al., 2010] [Taghia et al., 2018] [Uddin, 2015] (see Fig. 3.4, right panel). Different clinical populations, such as children with autism spectrum disorder and adults with schizophrenia, show impaired SNs [Odriozola et al., 2016] [Supekar et al., 2018]. On the contrary, the optimal calibration of the SN allows cognitive flexibility.

The neural dynamics underlying cognitive flexibility report CEN, SN and DMN interactions. In adults aged 22–35 years, transient connectivity and rapid switching is observed between the SN, CEN and DMN, with each presenting multiple short-lived states [Ryali et al., 2016]. The higher cognitive flexibility, the better creative thinking capacities are as a result of the tight coupling of the CEN and DMN [Beaty et al., 2016] [Kenett et al., 2018] [Beaty et al., 2018] [Adnan et al., 2019] [Shi et al., 2019].

As performance monitoring reflects the quality of ones voluntary control over oneself to self-regulate unexpected events, the quality of an individual's performance monitoring may predict his/her level of cognitive flexibility and creative abilities. Corroborating previous studies [Besançon and Lubart, 2008] [Lillard and Else-Quest, 2006] [Fleming et al., 2019], we showed that Montessori-schooled students score higher in both divergent and convergent creativity tasks. This may result from higher metastability in their neural networks and related to increased self-regulatory skills.

Preliminary results on cognitive flexibility

In this work, we lack a direct measure linking performance monitoring to cognitive flexibility. However, when running an independent component analysis to isolate DMN, CEN and SN from resting-state fMRI data (42 children aged 6-12, 22 from Montessori classes), we do find preliminary evidence that Montessori students have a stronger co-activation of the DMN and CEN than traditionally-schooled students. While this may denote increased cognitive flexibility and creative abilities, we show that correlating the DMN and CEN is highly dependent on the emotion regulation strategy. The higher the capacity for cognitive reappraisal [Ochsner et al., 2002] [Gross and John, 2003], the more the SN is related to the CEN, suggesting a voluntary control to self-regulate. This effect was significantly stronger in Montessori-schooled students (Fig. 3.5). These preliminary findings corroborate those of another study showing how SN FC changes across ages are related to the global development of emotion networks [Zhang et al., 2019]. The more intrinsic self-regulation the student grows, the higher his/her self-monitoring may be, with a direct impact on both flexible and creative competencies.

We further have preliminary evidence of differences in cognitive flexibility through the verbal fluency measure. In this task, the child produced words belonging to a defined semantic category (e.g., animals) during one minute. The analysis of the vocabulary field is done through semantic network pipelines using graph measures (here the Fruchterman Reingold algorithm, see [Kenett et al., 2017]). The extent and quality of semantic distance between words have been shown to reflect adults level of openness and creativity [Kenett and Faust, 2019] [Kenett et al., 2014] [Christensen et al., 2018]. Here, we found significant differences between five to twelve years of age students (N=65, half from each type of schooling). Montessori-schooled students had a higher capacity to associate distant concepts and showed a broader semantic network. We further have preliminary evidence of differences in cognitive flexibility through the verbal fluency measure. In this task, the child needs to produce words that belong to a defined semantic category (here; animals) during one minute. The analysis of the resulting vocabulary field is done through some semantic network pipelines using graph measures (here). The extent and quality of semantic distance between words have been shown, in adults, to reflect ones level

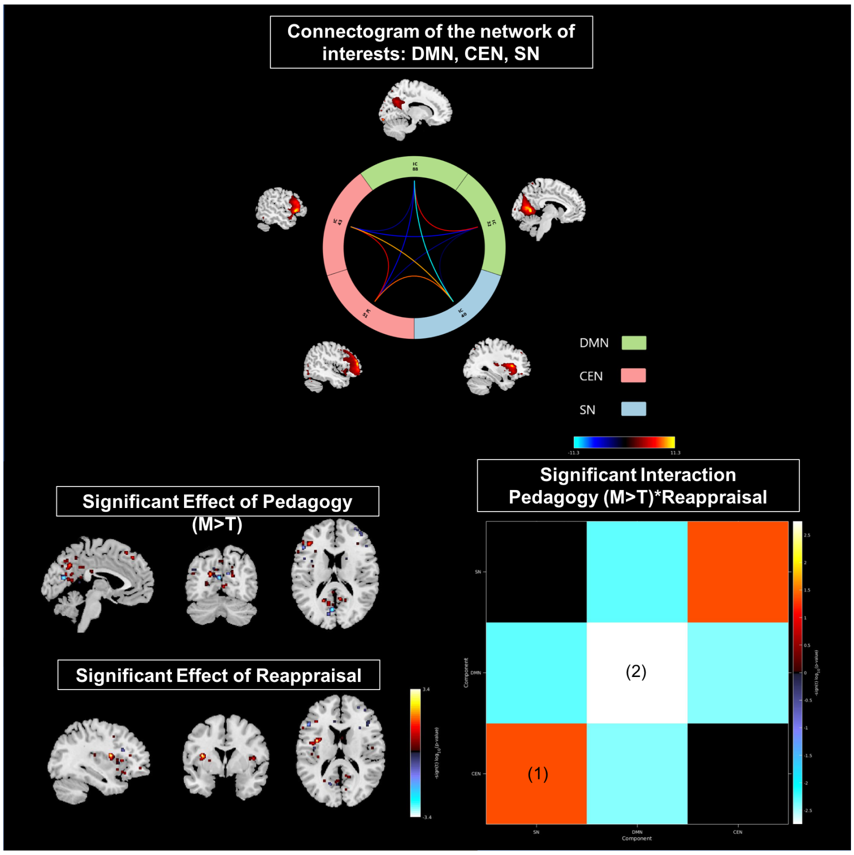


Figure 3.5 – Inter- and intra-networks of interest (DMN, CEN, SN) connectivity (top panel). Main effects of pedagogy (DMN more connected to the CEN in Montessori versus traditionally-schooled students) and emotion-regulation strategy (reappraisal was related to functional connectivity between the SN and the CEN) (left panel). Interactions between the pedagogical practices experienced by the child and its capacity to self-regulate his/her emotion (right panel); Montessori-schooled students reappraisal capacity was significantly correlated to (1) increased intra-DMN functional connectivity, and to (2) stronger inter-SN and CEN functional connectivity.

of openness and creativity. Here, we found significant differences with Montessori-schooled students having a higher capacity to associate distant concepts and showing a broader semantic network (Fig. 3.6).

3.2 Limitations

While these findings offer new insights into the development of children's performance monitoring system and its susceptibility to pedagogical practices, it comes along with limitations that future research should address to consolidate and broaden this work.

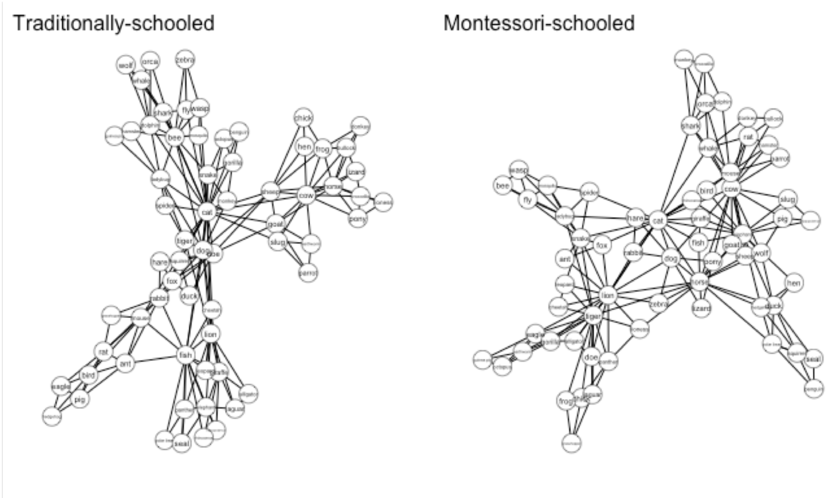


Figure 3.6 – Semantic network analysis comparing between-group capacity to cite less-related words, as well as between-group size and complexity of semantic network.

3.2.1 Selection bias and possible confounds

As a result of local policies in Switzerland, the Montessori schools partaking in these studies were all private, while all the traditional schools were public. We thus controlled for participants' fluid intelligence level as well as their parents' socioeconomic level (we further controlled for other factors in the neural studies; e.g., home environment, how much parents perceived life as stressful). Although the measured effects are similar to those of randomized studies, we cannot rule out selection bias. Indeed, by choosing Montessori education for their child, parents may have developed a more in-depth educational reflection and/or present a higher level of mental flexibility, with direct implications on their behavior and pedagogical practices at home. A feasible option to address this limitation would be to include children whose parents wished for Montessori education, but did not pursue it for practical reasons (e.g., the school was too far from their place of residence), but not because of financial limitations. Another would be to train the parents of prospective participants on pedagogical features and education before they enroll their child in one of the two schooling systems. Later, we could then start to collect longitudinal data on children to measure the extent to which pedagogical practices at school play a role in the development of performance monitoring compared with parental practices at home.

3.2.2 Small sample size

An explicit limitation of these studies is their relatively small sample size. We conducted power or sensitivity analyses to ensure that the effects measured were reliable. However, the findings result from a limited number of participants. When investigating the impact of pedagogical practices, where many other variables can interact, this may weaken the findings. Inter-individual variability may blur some of the effects or bias the outcomes. As a consequence, we may miss essential aspects of performance monitoring or overstate others because of the low number of participants. By other research groups replicating the studies and pooling data from different research groups in comparable areas (e.g., Swiss and French research groups), we could confirm the validity of the findings.

3.2.3 Cross-sectional design

We report developmental findings using only cross-sectional designs, whereas a state-of-the-art approach would be able to investigate the same measures under a longitudinal design. While the multi-modal approach could provide the redundant findings, future work should aim to assess some of the performance monitoring measures (PES, PIA, multisensory integration) over at least a year or two.

3.2.4 Quantitative approach

Finally, quantitative research, as presented in this work, can establish developmental benchmarks with a high degree of precision. However, this only represents a restricted view of the more complex reality. A view defined through a priori hypotheses and specific measurement tools can eliminate a large number of the variables that play a significant role in real life. Quantitative developmental studies have two direct consequences: (i) they lack a global and contextual perspective of the mechanisms studied and (ii) they raise questions of interest that derive from an adult "mature brain" that refrains from a more bottom-up developmental approach (an adult vision of what children's life experience is while children certainly do not sense reality as we do). A multi-modal approach makes it possible to overcome the first point partially; (i.e., we can approach the same question using complementary tools to broaden our understanding of a process); however, it does not address the second point. It would thus be useful to supplement present studies with qualitative observational approaches such as unsupervised interviews with children, record self-talk and gaze as well as investigate the content of spontaneous interactions with adults and peers when they face adversity in their endogenous environment. Less "controlled" approaches would offer knowledge and understanding about the child and the meaning of his/her developmental processes, such as those related to performance monitoring.

3.3 Future directions

Future work is needed to confirm the reported findings and address the above-mentioned limitations. However, these new insights in performance monitoring already raise further questions that would benefit from new research.

3.3.1 Isolate pedagogical features

In this work, we compared Montessori with traditionally schooled students. The dissimilarities found offer new insights into the development of the underlying brain mechanisms. However, this work cannot decipher the features of the Montessori pedagogy that make it possible for students to grow different performance monitoring abilities. As stated in the Introduction, traditional and Montessori pedagogical practices differ in at least two critical aspects: (i) the social environment (multi-grade classrooms with a low teacher to students ratio) and (ii) the learning approach (sensory education in kindergarten, self-corrective material, decision-making in daily curricula, uninterrupted work for three hours, no grade or reward). Studying these specific features in isolation could shed new light on the developmental processes of performance monitoring. For example, researchers could compare traditionally schooled groups of students that differ only in their grading system (in Switzerland, some alternative private schools do not use a grading system while keeping a traditional teaching approach). Alternatively, they might study specific social aspects such as multi-grade classrooms (public schools in rural areas pool children of different levels). Another research avenue could be to include alternative pedagogical approaches such as Waldorf or Freinet and look for any similarities between these approaches to help isolate crucial features. However, we suspect that even if some pedagogical features in isolation play an essential role in the performance monitoring process, the complexity of neurodevelopment could not be addressed other than by relying on a combination of features rather than simply one or two.

3.3.2 Peer-to-peer interactions

Performance monitoring studies in adults have paid much attention to the direct relation to supervised learning. However, an as-yet understudied side of performance monitoring deserves more investigation: unsupervised learning. How do children learn when not supervised? In particular, how do children learn from observing and interacting with peers of a similar age? And how do these peer-to-peer interactions impact learning strategies such as error monitoring, self-correction and action value attribution, all markers of performance monitoring. It might also be important to investigate the age at which children start using cues from peers as a source of information (Fig. 3.7). These aspects of learning are not yet well understood despite their fundamental importance for effective pedagogical practices. The investigation of such ideas should include eye-tracking studies, studies that characterize children's displacements and interactions with peers (frequency, duration) and portative EEG experiments to record the extent to which brainwaves synchronize with peers compared with an adult. While these studies would be less directly related to pedagogical practices, they would benefit from endogenous learning contexts such as classrooms in which children are free to move and interact, as found in Montessori environments.

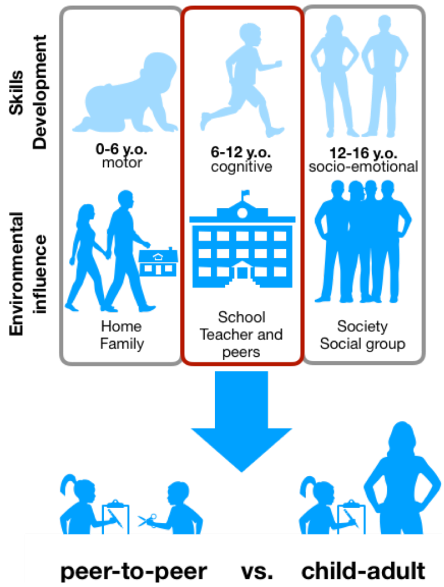


Figure 3.7 – It would be of interest to investigate how children learn with peers compared to with adults.

3.3.3 Longitudinal study

Our findings suggest that school pedagogical practices impact performance monitoring in 5–15-year-old children, with preliminary evidence for a window of plasticity for learning from errors. Confirming and characterizing the age range for the scaffolding of performance monitoring would rely on longitudinal studies. At least three research questions should be addressed. The first is the extent to which sensory integration capacity at kindergarten age predicts a child's ERN brain component. If any relationship were to be found, this would shed new light on the brainwaves that could reflect a time-wise record of experience and explain the cascading effect. It would also mean that the early sensory window of opportunity serves as a first step to learning how to learn autonomously, that is *performance monitoring*. Second, we could test the possibility of a cascading effect of error-monitoring abilities in childhood on an adolescent's coping with conflict in a social environment. One's relationship with error might predict the quality of ones interpersonal inter-

actions, especially the capacity for cooperative behaviors. The third is to track and characterize the behavioral shift and study its relation with brain maturation and how experience fastens or optimizes this process (Fig. 3.8).
These studies would benefit from multi-modal approaches and the use of more endogenous settings that better mat reality.

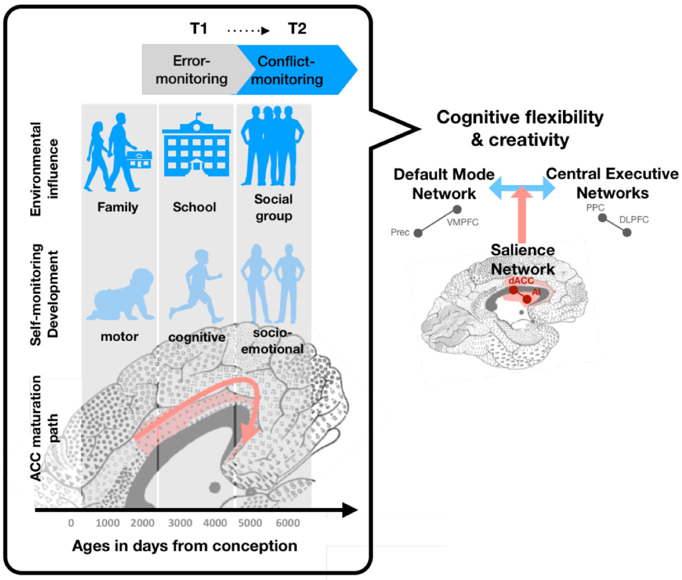


Figure 3.8 – We hypothesize that the quality of error-monitoring in school-aged children is predictive or conflict-monitoring in adolescents, based on the maturational path of the anterior cingulate cortex (ACC) and related self-moniotring competences [Posner et al., 2007], with direct effects on individual's cognitive flexibility in social contexts. Nodes of the networks related to cognitive flexibility and creativity are reported in the figures: Precuneus (Prec) and ventro-medial prefrontal cortex (VMPFC) for the default mode network; dorsal anterior cingulate cortex (dACC) and anterior insula (AI) for the salience network; posterior parietal cingulate (PCC) and dorso-lateral prefrontal cortex (DLPFC) for the central executive network.

3.4 Conclusion

This work sheds light on the emergence of mental habits when facing unexpected events such as response errors. While learning signals for performance monitoring shift across ages, pedagogical practices play a crucial role in shaping cognitive and social reactions. Indeed, when comparing Montessori students with traditionally-schooled students, differences in performance monitoring are unveiled not only at the behavioral level, but also at the neural one. More work is needed to consolidate these findings and investigate the long-term implications on mindsets and interpersonal behaviors. Nonetheless, this work opens two essential research avenues. The first is that performance monitoring in children seems to be oriented more toward correct responses than incorrect ones, contrary to adults, with more sigmoidal than linear developmental dynamics. The second is the influence of school experience in the development of performance monitoring, which strongly suggests the existence of a specific window of opportunity for the acquisition of that competency. Together, these two aspects may have a drastic impact on education and thus they urgently call for further research. In a social context in which rapid changes and uncertainty predominate and where AI development and implementation occur at a sustained pace, students need to foster their flexibility, creativity and self-regulatory skills to ensure their future. It therefore seems wise to think about school pedagogical practices from an enlightened neurodevelopmental perspective.

Every person is the author of his own skills.

Maria Montessori, The Absorbent Mind

Bibliography

[New, 2015] (2015). New vision for education; unlocking the potential of technology. Report, World Economic Forum.

[Aarts et al., 2012] Aarts, K., De Houwer, J., and Pourtois, G. (2012). Evidence for the automatic evaluation of self-generated actions. *Cognition*, 124(2):117–27.

[Aarts et al., 2013] Aarts, K., De Houwer, J., and Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: evidence from erps. *Emotion*, 13(5):960–73.

[Aarts and Pourtois, 2010] Aarts, K. and Pourtois, G. (2010). Anxiety not only increases, but also alters early error-monitoring functions. *Cogn Affect Behav Neurosci*, 10(4):479–92.

[Adnan et al., 2019] Adnan, A., Beaty, R., Lam, J., Spreng, R. N., and Turner, G. R. (2019). Intrinsic default-executive coupling of the creative aging brain. *Soc Cogn Affect Neurosci*, 14(3):291–303.

[Ahmad and Reba, 2018] Ahmad, S. and Reba, A. (2018). Social and moral development of students: A comparative study on montessori and non-montessori students. *Journal of Elementary Education*, 28(1):53–64.

[Allman et al., 2001] Allman, J. M., Hakeem, A., Erwin, J. M., Nimchinsky, E., and Hof, P. (2001). The anterior cingulate cortex. the evolution of an interface between emotion and cognition. *Ann N Y Acad Sci*, 935:107–17.

[Bandura et al., 1963] Bandura, A., Ross, D., and Ross, S. A. (1963). Vicarious reinforcement and imitative learning. *J Abnorm Psychol*, 67:601–7.

[Bavelier et al., 2010] Bavelier, D., Levi, D. M., Li, R. W., Dan, Y., and Hensch, T. K. (2010). Removing brakes on adult brain plasticity: from molecular to behavioral interventions. *J Neurosci*, 30(45):14964–71.

[Beaty et al., 2016] Beaty, R. E., Benedek, M., Silvia, P. J., and Schacter, D. L. (2016). Creative cognition and brain network dynamics. *Trends Cogn Sci*, 20(2):87–95.

[Beaty et al., 2018] Beaty, R. E., Kenett, Y. N., Christensen, A. P., Rosenberg, M. D., Benedek, M., Chen, Q., Fink, A., Qiu, J., Kwapil, T. R., Kane, M. J., and Silvia, P. J. (2018). Robust prediction of individual creative ability from brain functional connectivity. *Proc Natl Acad Sci U S A*, 115(5):1087–1092.

[Besançon and Lubart, 2008] Besançon, M. and 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.

[Call et al., 2005] Call, J., Carpenter, M., and Tomasello, M. (2005). Copying results and copying actions in the process of social learning: chimpanzees (pan troglodytes) and human children (homo sapiens). *Anim Cogn*, 8(3):151–63.

[Callaghan and Tottenham, 2016] Callaghan, B. L. and Tottenham, N. (2016). The neuro-environmental loop of plasticity: A cross-species analysis of parental effects on emotion circuitry development following typical and adverse caregiving. *Neuropsychopharmacology*, 41(1):163–76.

[Carcea and Froemke, 2013] Carcea, I. and Froemke, R. C. (2013). Cortical plasticity, excitatory-inhibitory balance, and sensory perception. *Changing Brains Applying Brain Plasticity to Advance and Recover Human Ability*, 207:65–90.

[Carrasco et al., 2013] Carrasco, M., Harbin, S. M., Nienhuis, J. K., Fitzgerald, K. D., Gehring, W. J., and Hanna, G. L. (2013). Increased error-related brain activity in youth with obsessive-compulsive disorder and unaffected siblings. *Depression and Anxiety*, 30(1):39–46.

[Carter et al., 1998] Carter, C. S., Braver, T., Barch, D. M., Botvinick, M., Noll, D., and Cohen, J. D. (1998). The role of the anterior cingulate cortex in error detection and the on-line monitoring of performance: An event related fmri study. *Biological Psychiatry*, 43:13s–13s.

[Carter et al., 2000] Carter, C. S., Macdonald, A. M., Botvinick, M., Ross, L. L., Stenger, V. A., Noll, D., and Cohen, J. D. (2000). Parsing executive processes: Strategic vs. evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 97(4):1944–1948.

[Christensen et al., 2018] Christensen, A. P., Kenett, Y. N., Cotter, K. N., Beaty, R. E., and Silvia, P. J. (2018). Remotely close associations: Openness to experience and semantic memory structure. *European Journal of Personality*, 32(4):480–492.

[Clement and Dukes, 2013] Clement, F. and Dukes, D. (2013). The role of interest in the transmission of social values. *Front Psychol*, 4:349.

[Compton et al., 2011] Compton, R. J., Arnstein, D., Freedman, G., Dainer-Best, J., Liss, A., and Robinson, M. D. (2011). Neural and behavioral measures of error-related cognitive control predict daily coping with stress. *Emotion*, 11(2):379–390.

[Condliffe et al., 2017] Condliffe, B., Quint, J., Visher, M. G., Bangser, M. R., D. S., Saco, L., and Nelson, E. (2017). Project-based learning; a literature review.

[Courtier, 2019] Courtier, P. (2019). *Limpact de la pédagogie Montessori sur le développement cognitif, social et académique des enfants en maternelle.* Thesis.

[Crone et al., 2004] Crone, E. A., Ridderinkhof, K. R., Worm, M., Somsen, R. J., and van der Molen, M. W. (2004). Switching between spatial stimulus-response mappings: a developmental study of cognitive flexibility. *Dev Sci*, 7(4):443–55.

[Crone and van der Molen, 2007] Crone, E. A. and van der Molen, M. W. (2007). Development of decision making in school-aged children and adolescents: evidence from heart rate and skin conductance analysis. *Child Dev*, 78(4):1288–301.

[Danielmeier et al., 2011] Danielmeier, C., Eichele, T., Forstmann, B. U., Tittgemeyer, M., and Ullsperger, M. (2011). Posterior medial frontal cortex activity predicts post-error adaptations in task-related visual and motor areas. *J Neurosci*, 31(5):1780–9.

[Danielmeier and Ullsperger, 2011] Danielmeier, C. and Ullsperger, M. (2011). Post-error adjustments. *Front Psychol*, 2:233.

[Davies et al., 2004] Davies, P. L., Segalowitz, S. J., and Gavin, W. J. (2004). Development of response-monitoring erps in 7- to 25-year-olds. *Dev Neuropsychol*, 25(3):355–76.

[de Bruijn et al., 2009] de Bruijn, E. R., de Lange, F. P., von Cramon, D. Y., and Ullsperger, M. (2009). When errors are rewarding. *J Neurosci*, 29(39):12183–6.

[De Saedeleer and Pourtois, 2016] De Saedeleer, L. and Pourtois, G. (2016). Evaluative priming reveals dissociable effects of cognitive versus physiological anxiety on action monitoring. *Emotion*, 16(4):498–514.

[Dehaene-Lambertz, 2000] Dehaene-Lambertz, G. (2000). Development of phonological perception in children: electrophysiological studies. *Revue De Neuropsychologie*, 10(4):519–533.

[Dehaene-Lambertz and Gliga, 2004] Dehaene-Lambertz, G. and Gliga, T. (2004). Common neural basis for phoneme processing in infants and adults. *Journal of Cognitive Neuroscience*, 16(8):1375–1387.

[Diamond, 2012] Diamond, A. (2012). Activities and programs that improve children's executive functions. *Curr Dir Psychol Sci*, 21(5):335–341.

[Diamond, 2013] Diamond, A. (2013). Executive functions. *Annual Review of Psychology, Vol 64*, 64:135–168.

[Diamond, 2014] Diamond, A. (2014). Want to optimize executive functions and academic outcomes?: Simple, just nourish the human spirit. *Minn Symp Child Psychol Ser*, 37:205–232.

[Diamond and Lee, 2011] Diamond, A. and Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, 333(6045):959–64.

[Dignath et al., 2019] Dignath, D., Eder, A. B., Steinhauser, M., and Kiesel, A. (2019). Conflict monitoring and the affective signaling hypothesis an integrative review. *Psychonomic Bulletin Review*.

[Dohrmann et al., 2007] Dohrmann, K., Nishida, T., Gartner, A., Lipsky, D., and Grimm, K. J. (2007). High school outcomes for students in a public montessori program. *Journal of Research in Childhood Education*, 22(2):103–114.

[Dolan and Dayan, 2013] Dolan, R. J. and Dayan, P. (2013). Goals and habits in the brain. *Neuron*, 80(2):312–25.

[Downes et al., 2017] Downes, M., Bathelt, J., and De Haan, M. (2017). Event-related potential measures of executive functioning from preschool to adolescence. *Dev Med Child Neurol*, 59(6):581–590.

[Dubois et al., 2012] Dubois, J., Dehaene-Lambertz, G., Mangin, J. F., Le Bihan, D., Huppi, P. S., and Hertz-Pannier, L. (2012). Brain development of infant and mri by diffusion tensor imaging. *Neurophysiologie Clinique-Clinical Neurophysiology*, 42(1-2):1–9.

[Dubois et al., 2009] Dubois, J., Hertz-Pannier, L., Cachia, A., Mangin, J. F., Le Bihan, D., and Dehaene-Lambertz, G. (2009). Structural asymmetries in the infant language and sensori-motor networks. *Cerebral Cortex*, 19(2):414–423.

[Dutilh et al., 2012] Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., and Wagenmakers, E. J. (2012). Testing theories of post-error slowing. *Atten Percept Psychophys*, 74(2):454–65.

[Endrass et al., 2014] Endrass, T., Riesel, A., Kathmann, N., and Buhlmann, U. (2014). Performance monitoring in obsessive-compulsive disorder and social anxiety disorder. *J Abnorm Psychol*, 123(4):705–14.

[Endrass and Ullsperger, 2014] Endrass, T. and Ullsperger, M. (2014). Specificity of performance monitoring changes in obsessive-compulsive disorder. *Neurosci Biobehav Rev*, 46 Pt 1:124–38.

[Eriksen and Eriksen, 1974] Eriksen, B. A. and Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a non- search task. *Perception and Psychophysics*, 16:143149.

[Ervin et al., 2010] Ervin, B., Wash, P. D., and Mecca, M. E. (2010). A 3-year study of self-regulation in montessori and non-montessori classrooms. *Montessori Life*, (2).

[Fair et al., 2007] Fair, D. A., Dosenbach, N. U., Church, J. A., Cohen, A. L., Brahmbhatt, S., Miezin, F. M., Barch, D. M., Raichle, M. E., Petersen, S. E., and Schlaggar, B. L. (2007). Development of distinct control networks through segregation and integration. *Proc Natl Acad Sci U S A*, 104(33):13507–12.

[Fleming et al., 2019] Fleming, D. J., Culclasure, B., and Zhang, D. (2019). The montessori model and creativity. *Journal of Montessori Research*, 5(2).

[Frese and Keith, 2015] Frese, M. and Keith, N. (2015). Action errors, error management, and learning in organizations. *Annual Review of Psychology, Vol 66*, 66:661–687.

[Friedrichs-Maeder et al., 2017] Friedrichs-Maeder, C. L., Griffa, A., Schneider, J., Huppi, P. S., Truttmann, A., and Hagmann, P. (2017). Exploring the role of white matter connectivity in cortex maturation. *PLoS One*, 12(5):e0177466.

[Glascher et al., 2010] Glascher, J., Daw, N., Dayan, P., and O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66(4):585–95.

[Gogtay et al., 2004] Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., r., Herman, D. H., Clasen, L. S., Toga, A. W., Rapoport, J. L., and Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc Natl Acad Sci U S A*, 101(21):8174–9.

[Gottlieb et al., 2014] Gottlieb, J., Lopes, M., and Oudeyer, P. Y. (2014). Active learning and decision making: an introduction to the collection. *F1000Res*, 3:276.

[Gottlieb and Oudeyer, 2018] Gottlieb, J. and Oudeyer, P. Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nat Rev Neurosci*, 19(12):758–770.

[Grammer et al., 2014] Grammer, J. K., Carrasco, M., Gehring, W. J., and Morrison, F. J. (2014). Age-related changes in error processing in young children: a school-based investigation. *Dev Cogn Neurosci*, 9:93–105.

[Grisay et al., 2007] Grisay, A., de Jong, J. H., Gebhardt, E., Berezner, A., and Halleux-Monseur, B. (2007). Translation equivalence across pisa countries. *J Appl Meas*, 8(3):249–66.

[Gross and John, 2003] Gross, J. J. and John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *J Pers Soc Psychol*, 85(2):348–62.

[Hajcak and Foti, 2008] Hajcak, G. and Foti, D. (2008). Errors are aversive: defensive motivation and the error-related negativity. *Psychol Sci*, 19(2):103–8.

[Hajcak et al., 2003] Hajcak, G., McDonald, N., and Simons, R. F. (2003). Anxiety and error-related brain activity. *Biological Psychology*, 64(1-2):77–90.

[Hensch, 2004] Hensch, T. K. (2004). Critical period regulation. *Annu Rev Neurosci*, 27:549–79.

[Hensch, 2005a] Hensch, T. K. (2005a). Critical period mechanisms in developing visual cortex. *Curr Top Dev Biol*, 69:215–37.

[Hensch, 2005b] Hensch, T. K. (2005b). Critical period plasticity in local cortical circuits. *Nat Rev Neurosci*, 6(11):877–88.

[Hensch and Bilimoria, 2012] Hensch, T. K. and Bilimoria, P. M. (2012). Re-opening windows: Manipulating critical periods for brain development. *Cerebrum*, 2012:11.

[Hensch and Fagiolini, 2005] Hensch, T. K. and Fagiolini, M. (2005). Excitatory-inhibitory balance and critical period plasticity in developing visual cortex. *Prog Brain Res*, 147:115–24.

[Hogan et al., 2005] Hogan, A. M., Vargha-Khadem, F., Kirkham, F. J., and Baldeweg, T. (2005). Maturation of action monitoring from adolescence to adulthood: an erp study. *Dev Sci*, 8(6):525–34.

[Houtman and Notebaert, 2013] Houtman, F. and Notebaert, W. (2013). Blinded by an error. *Cognition*, 128(2):228–36.

[Huttenlocher and Dabholkar, 1997] Huttenlocher, P. R. and Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *J Comp Neurol*, 387(2):167–78.

[Izuma et al., 2015] Izuma, K., Akula, S., Murayama, K., Wu, D. A., Iacoboni, M., and Adolphs, R. (2015). A causal role for posterior medial frontal cortex in choice-induced preference change. *J Neurosci*, 35(8):3598–606.

[Jones et al., 2003] Jones, L. B., Rothbart, M. K., and Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science*, 6(5):498–504.

[Kang et al., 2010] Kang, S. K., Hirsh, J. B., and Chasteen, A. L. (2010). Your mistakes are mine: Self-other overlap predicts neural response to observed errors. *Journal of Experimental Social Psychology*, 46(1):229–232.

[Keith and Frese, 2005] Keith, N. and Frese, M. (2005). Self-regulation in error management training: emotion control and metacognition as mediators of performance effects. *J Appl Psychol*, 90(4):677–91.

[Keith and Frese, 2008] Keith, N. and Frese, M. (2008). Effectiveness of error management training: a meta-analysis. *J Appl Psychol*, 93(1):59–69.

[Kelly et al., 2009] Kelly, A. M., Di Martino, A., Uddin, L. Q., Shehzad, Z., Gee, D. G., Reiss, P. T., Margulies, D. S., Castellanos, F. X., and Milham, M. P. (2009). Development of anterior cingulate functional connectivity from late childhood to early adulthood. *Cereb Cortex*, 19(3):640–57.

[Kenett et al., 2014] Kenett, Y. N., Anaki, D., and Faust, M. (2014). Investigating the structure of semantic networks in low and high creative persons. *Frontiers in Human Neuroscience*, 8.

[Kenett and Faust, 2019] Kenett, Y. N. and Faust, M. (2019). A semantic network cartography of the creative mind. *Trends Cogn Sci*, 23(4):271–274.

[Kenett et al., 2017] Kenett, Y. N., Levi, E., Anaki, D., and Faust, M. (2017). The semantic distance task: Quantifying semantic distance with semantic network path length. *J Exp Psychol Learn Mem Cogn*, 43(9):1470–1489.

[Kenett et al., 2018] Kenett, Y. N., Medaglia, J. D., Beaty, R. E., Chen, Q., Betzel, R. F., Thompson-Schill, S. L., and Qiu, J. (2018). Driving the brain towards creativity and intelligence: A network control theory analysis. *Neuropsychologia*, 118(Pt A):79–90.

[Kim et al., 2005] Kim, E. Y., Iwaki, N., Uno, H., and Fujita, T. (2005). Error-related negativity in children: effect of an observer. *Dev Neuropsychol*, 28(3):871–83.

[Klucharev et al., 2011] Klucharev, V., Munneke, M. A., Smidts, A., and Fernandez, G. (2011). Downregulation of the posterior medial frontal cortex prevents social conformity. *J Neurosci*, 31(33):11934–40.

[Koban and Pourtois, 2014] Koban, L. and Pourtois, G. (2014). Brain systems underlying the affective and social monitoring of actions: an integrative review. *Neurosci Biobehav Rev*, 46 Pt 1:71–84.

[Koban et al., 2010] Koban, L., Pourtois, G., Vocat, R., and Vuilleumier, P. (2010). When your errors make me lose or win: event-related potentials to observed errors of cooperators and competitors. *Soc Neurosci*, 5(4):360–74.

[Ladouceur et al., 2007] Ladouceur, C. D., Dahl, R. E., and Carter, C. S. (2007). Development of action monitoring through adolescence into adulthood: Erp and source localization. *Dev Sci*, 10(6):874–91.

[Laubach et al., 2015] Laubach, M., Caetano, M. S., and Narayanan, N. S. (2015). Mistakes were made: neural mechanisms for the adaptive control of action initiation by the medial prefrontal cortex. *J Physiol Paris*, 109(1-3):104–17.

[Lebenberg et al., 2018] Lebenberg, J., Mangin, J. F., Thirion, B., Poupon, C., Hertz-Pannier, L., Leroy, F., Adibpour, P., Dehaene-Lambertz, G., and Dubois, J. (2018). Mapping the asynchrony of cortical maturation in the infant brain: A mri multiparametric clustering approach. *Neuroimage*.

[Lewis, 2003] Lewis, M. (2003). The emergence of consciousness and its role in human development. *Ann N Y Acad Sci*, 1001:104–33.

[Lillard, 2019] Lillard, A. (2019). Shunned and admired: Montessori, self-determination, and a case for radical school reform. *Educational Psychology Review*.

[Lillard and Else-Quest, 2006] Lillard, A. and Else-Quest, N. (2006). The early years. evaluating montessori education. *Science*, 313(5795):1893–4.

[Lillard, 2011] Lillard, A. S. (2011). Mindfulness practices in education: Montessoris approach. *Mindfulness*, 2(2):78–85.

[Lillard, 2012] Lillard, A. S. (2012). Preschool children's development in classic montessori, supplemented montessori, and conventional programs. *J Sch Psychol*, 50(3):379–401.

[Lillard et al., 2017] Lillard, A. S., Heise, M. J., Richey, E. M., Tong, X., Hart, A., and Bray, P. M. (2017). Montessori preschool elevates and equalizes child outcomes: A longitudinal study. *Front Psychol*, 8:1783.

[Lindenfors, 2007] Lindenfors, P. (2007). Studying students in montessori schools. *Science*, 315(5812):596–7; author reply 596–7.

[Lopota et al., 2005] Lopota, C., Wallace, N. V., and Finn, K. V. (2005). Comparison of academic achievement between montessori and traditional education programs. *Journal of Research in Childhood Education*, 20(1):5–13.

[Mackinnon, 2007] Mackinnon, P. (2007). Studying students in montessori schools. *Science*, 315(5812):596–7; author reply 596–7.

[Margulies et al., 2007] Margulies, D. S., Kelly, A. M., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., and Milham, M. P. (2007). Mapping the functional connectivity of anterior cingulate cortex. *Neuroimage*, 37(2):579–88.

[Marshall, 2017] Marshall, C. (2017). Montessori education: a review of the evidence base. *npj Science of Learning*, 2(1).

[Mattar et al., 2018] Mattar, M. G., Thompson-Schill, S. L., and Bassett, D. S. (2018). The network architecture of value learning. *Netw Neurosci*, 2(2):128–149.

[Meltzoff, 1995] Meltzoff, A. N. (1995). What infant memory tells us about infantile amnesia: long-term recall and deferred imitation. *J Exp Child Psychol*, 59(3):497–515.

[Menon, 2015] Menon, V. (2015). *Salience Network*, pages 597–611.

[Menon and Uddin, 2010] Menon, V. and Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Struct Funct*, 214(5-6):655–67.

[Metcalfe, 2017] Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology, Vol 68*, 68:465–489.

[Metcalfe and Xu, 2018] Metcalfe, J. and Xu, J. (2018). Learning from one's own errors and those of others. *Psychonomic Bulletin Review*, 25(1):402–408.

[Meyer et al., 2013] Meyer, A., Hajcak, G., Torpey, D. C., Kujawa, A., Kim, J., Bufferd, S., Carlson, G., and Klein, D. N. (2013). Increased error-related brain activity in six-year-old children with clinical anxiety. *J Abnorm Child Psychol*, 41(8):1257–66.

[Meyer et al., 2012] Meyer, A., Weinberg, A., Klein, D. N., and Hajcak, G. (2012). The development of the error-related negativity (ern) and its relationship with anxiety: evidence from 8 to 13 year-olds. *Dev Cogn Neurosci*, 2(1):152–61.

[Miyake et al., 2000] Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1):49–100.

[Montessori, 1936] Montessori, M. (1936). *The Secret of Childhood*. Ballantine, New York, 1981 edition.

[Morieux, 2018] Morieux, Y. (2018). Bringing manager back to work. *BCG*, pages 1–32.

[Moser et al., 2011] Moser, J. S., Schroder, H. S., Heeter, C., Moran, T. P., and Lee, Y. H. (2011). Mind your errors: evidence for a neural mechanism linking growth mind-set to adaptive posterror adjustments. *Psychol Sci*, 22(12):1484–9.

[Murphy et al., 2016] Murphy, P. R., van Moort, M. L., and Nieuwenhuis, S. (2016). The pupillary orienting response predicts adaptive behavioral adjustment after errors. *PLoS One*, 11(3):e0151763.

[Neftci and Averbeck, 2019] Neftci, E. O. and Averbeck, B. B. (2019). Reinforcement learning in artificial and biological systems. *Nature Machine Intelligence*, 1(3):133–143.

[Nelson and Bloom, 1997] Nelson, C. A. and Bloom, F. E. (1997). Child development and neuroscience. *Child Dev*, 68(5):970–987.

[Nicol, 1995] Nicol, C. J. (1995). The social transmission of information and behavior. *Applied Animal Behaviour Science*, 44(2-4):79–98.

[Notebaert et al., 2009] Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., and Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition*, 111(2):275–9.

[Ochsner et al., 2002] Ochsner, K. N., Bunge, S. A., Gross, J. J., and Gabrieli, J. D. (2002). Rethinking feelings: an fmri study of the cognitive regulation of emotion. *J Cogn Neurosci*, 14(8):1215–29.

[Odriozola et al., 2016] Odriozola, P., Uddin, L. Q., Lynch, C. J., Kochalka, J., Chen, T., and Menon, V. (2016). Insula response and connectivity during social and non-social attention in children with autism. *Soc Cogn Affect Neurosci*, 11(3):433–44.

[Oudeyer et al., 2016] Oudeyer, P. Y., Gottlieb, J., and Lopes, M. (2016). Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. *Prog Brain Res*, 229:257–284.

[Oudeyer and Smith, 2016] Oudeyer, P. Y. and Smith, L. B. (2016). How evolution may work through curiosity-driven developmental process. *Top Cogn Sci*, 8(2):492–502.

[Paul et al., 2017] Paul, K., Walentowska, W., Bakic, J., Dondaine, T., and Pourtois, G. (2017). Modulatory effects of happy mood on performance monitoring: Insights from error-related brain potentials. *Cogn Affect Behav Neurosci*, 17(1):106–123.

[Peng and Md-Yunus, 2014] Peng, H. and Md-Yunus, S. (2014). Do students in montessori schools perform better on achievement tests? a taiwanese perspective. *International Journal of Early Childhood*, 46(2):299–311.

[Peters et al., 2016] Peters, S., Van Duijvenvoorde, A. C., Koolschijn, P. C., and Crone, E. A. (2016). Longitudinal development of frontoparietal activity during feedback learning: Contributions of age, performance, working memory and cortical thickness. *Dev Cogn Neurosci*, 19:211–22.

[Phillips-Silver and Daza, 2018] Phillips-Silver, J. and Daza, M. T. (2018). Cognitive control at age 3: Evaluating executive functions in an equitable montessori preschool. *Frontiers in Education*.

[Piaget, 1952] Piaget, J. (1952). *The origins of intelligence in children*. International Universities Press, New York.

[Plomin, 2014] Plomin, R. (2014). Genotype-environment correlation in the era of dna. *Behav Genet*, 44(6):629–38.

[Plomin et al., 2014] Plomin, R., Shakeshaft, N. G., McMillan, A., and Trzaskowski, M. (2014). Nature, nurture, and expertise. *Intelligence*, 45:46–59.

[Posner et al., 2007] Posner, M. I., Rothbart, M. K., Sheese, B. E., and Tang, Y. (2007). The anterior cingulate gyrus and the mechanism of self-regulation. *Cogn Affect Behav Neurosci*, 7(4):391–5.

[Pourtois et al., 2010] Pourtois, G., Vocat, R., N'Diaye, K., Spinelli, L., Seeck, M., and Vuilleumier, P. (2010). Errors recruit both cognitive and emotional monitoring systems: simultaneous intracranial recordings in the dorsal anterior cingulate gyrus and amygdala combined with fmri. *Neuropsychologia*, 48(4):1144–59.

[Purcell and Kiani, 2016] Purcell, B. A. and Kiani, R. (2016). Neural mechanisms of post-error adjustments of decision policy in parietal cortex. *Neuron*, 89(3):658–71.

[Rabbitt, 1966] Rabbitt, P. M. (1966). Errors and error correction in choice-response tasks. *J Exp Psychol*, 71(2):264–72.

[Rathunde and Csikszentmihalyi, 2005] Rathunde, K. and Csikszentmihalyi, M. (2005). Middle school students motivation and quality of experience: A comparison of montessori and traditional school environments. *American Journal of Education*, 111(3).

[Ridderinkhof et al., 2004] Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, 306(5695):443–7.

[Rose et al., 2008] Rose, S. A., Feldman, J. F., Jankowski, J. J., and Van Rossem, R. (2008). A cognitive cascade in infancy: Pathways from prematurity to later mental development. *Intelligence*, 36(4):367–378.

[Ruijs, 2017] Ruijs, N. (2017). The effects of montessori education: Evidence from admission lotteries. *Economics of Education Review*, 61:19–34.

[Ryali et al., 2016] Ryali, S., Supekar, K., Chen, T., Kochalka, J., Cai, W., Nicholas, J., Padmanabhan, A., and Menon, V. (2016). Temporal dynamics and developmental maturation of salience, default and central-executive network interactions revealed by variational bayes hidden markov modeling. *PLoS Comput Biol*, 12(12):e1005138.

[Santesso and Segalowitz, 2008] Santesso, D. L. and Segalowitz, S. J. (2008). Developmental differences in effor-related erps in middle- to late-adolescent males. *Developmental Psychology*, 44(1):205–217.

[Santesso et al., 2006] Santesso, D. L., Segalowitz, S. J., and Schmidt, L. A. (2006). Error-related electrocortical responses in 10-year-old children and young adults. *Dev Sci*, 9(5):473–81.

[Sara and Bouret, 2012] Sara, S. J. and Bouret, S. (2012). Orienting and reorienting: the locus coeruleus mediates cognition through arousal. *Neuron*, 76(1):130–41.

[Schroder and Moser, 2014] Schroder, H. S. and Moser, J. S. (2014). Improving the study of error monitoring with consideration of behavioral performance measures. *Front Hum Neurosci*, 8:178.

[Shadmehr et al., 2010] Shadmehr, R., Smith, M. A., and Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience, Vol 33*, 33:89–108.

[Shi et al., 2019] Shi, L., Beaty, R. E., Chen, Q., Sun, J., Wei, D., Yang, W., and Qiu, J. (2019). Brain entropy is associated with divergent thinking. *Cereb Cortex*.

[Smart and Segalowitz, 2017] Smart, C. M. and Segalowitz, S. J. (2017). Respond, don't react: The influence of mindfulness training on performance monitoring in older adults. *Cognitive Affective Behavioral Neuroscience*, 17(6):1151–1163.

[Smulders et al., 2016] Smulders, S. F., Soetens, E., and van der Molen, M. W. (2016). What happens when children encounter an error? *Brain Cogn*, 104:34–47.

[Supekar et al., 2018] Supekar, K., Cai, W., Krishnadas, R., Palaniyappan, L., and Menon, V. (2018). Dysregulated brain dynamics in a triple-network saliency model of schizophrenia and its relation to psychosis. *Biol Psychiatry*.

[Taghia et al., 2018] Taghia, J., Cai, W., Ryali, S., Kochalka, J., Nicholas, J., Chen, T., and Menon, V. (2018). Uncovering hidden brain state dynamics that regulate performance and decision-making during cognition. *Nat Commun*, 9(1):2505.

[Tamnes et al., 2013] Tamnes, C. K., Walhovd, K. B., Torstveit, M., Sells, V. T., and Fjell, A. M. (2013). Performance monitoring in children and adolescents: a review of developmental changes in the error-related negativity and brain maturation. *Dev Cogn Neurosci*, 6:1–13.

[Tau and Peterson, 2010] Tau, G. Z. and Peterson, B. S. (2010). Normal development of brain circuits. *Neuropsychopharmacology*, 35(1):147–68.

[Thomas and Johnson, 2008] Thomas, M. and Johnson, M. (2008). New advances in understanding sensitive periods in brain development. *Current Direction in Psychological Science*, 17.

[Torpey et al., 2012] Torpey, D. C., Hajcak, G., Kim, J., Kujawa, A., and Klein, D. N. (2012). Electrocortical and behavioral measures of response monitoring in young children during a go/no-go task. *Dev Psychobiol*, 54(2):139–50.

[Tottenham and Gabard-Durnam, 2017] Tottenham, N. and Gabard-Durnam, L. J. (2017). The developing amygdala: a student of the world and a teacher of the cortex. *Current Opinion in Psychology*, 17:55–60.

[Uddin, 2015] Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nat Rev Neurosci*, 16(1):55–61.

[Uddin et al., 2010] Uddin, L. Q., Supekar, K., and Menon, V. (2010). Typical and atypical development of functional human brain networks: insights from resting-state fmri. *Front Syst Neurosci*, 4:21.

[Uddin et al., 2011] Uddin, L. Q., Supekar, K. S., Ryali, S., and Menon, V. (2011). Dynamic reconfiguration of structural and functional connectivity across core neurocognitive brain networks with development. *J Neurosci*, 31(50):18578–89.

[Ullsperger and Danielmeier, 2016] Ullsperger, M. and Danielmeier, C. (2016). Reducing speed and sight: How adaptive is post-error slowing? *Neuron*, 89(3):430–2.

[Ullsperger et al., 2014a] Ullsperger, M., Danielmeier, C., and Jocham, G. (2014a). Neurophysiology of performance monitoring and adaptive behavior. *Physiol Rev*, 94(1):35–79.

[Ullsperger et al., 2014b] Ullsperger, M., Fischer, A. G., Nigbur, R., and Endrass, T. (2014b). Neural mechanisms and temporal dynamics of performance monitoring. *Trends Cogn Sci*, 18(5):259–67.

[van den Bos et al., 2009] van den Bos, W., Guroglu, B., van den Bulk, B. G., Rombouts, S. A. R. B., and Crone, E. A. (2009). Better than expected or as bad as you thought? the neurocognitive development of probabilistic feedback processing. *Frontiers in Human Neuroscience*, 3.

[Velanova et al., 2008] Velanova, K., Wheeler, M. E., and Luna, B. (2008). Maturational changes in anterior cingulate and frontoparietal recruitment support the development of error processing and inhibitory control. *Cerebral Cortex*, 18(11):2505–2522.

[Vocat et al., 2008] Vocat, R., Pourtois, G., and Vuilleumier, P. (2008). Unavoidable errors: a spatio-temporal analysis of time-course and neural sources of evoked potentials associated with error processing in a speeded task. *Neuropsychologia*, 46(10):2545–55.

[Vogel et al., 2010] Vogel, A. C., Power, J. D., Petersen, S. E., and Schlaggar, B. L. (2010). Development of the brain's functional network architecture. *Neuropsychology Review*, 20(4):362–375.

[Vygotsky, 1978] Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. MA: Harvard University Press, Cambridge.

[Weinberg et al., 2010] Weinberg, A., Olvet, D. M., and Hajcak, G. (2010). Increased error-related brain activity in generalized anxiety disorder. *Biol Psychol*, 85(3):472–80.

[Wessel et al., 2012] Wessel, J. R., Danielmeier, C., Morton, J. B., and Ullsperger, M. (2012). Surprise and error: common neuronal architecture for the processing of errors and novelty. *J Neurosci*, 32(22):7528–37.

[Zanolie et al., 2008] Zanolie, K., Teng, S., Donohue, S. E., van Duijvenvoorde, A. C., Band, G. P., Rombouts, S. A., and Crone, E. A. (2008). Switching between colors and shapes on the basis of positive and negative feedback: an fmri and eeg study on feedback-based learning. *Cortex*, 44(5):537–47.

[Zhang et al., 2019] Zhang, Y., Padmanabhan, A., Gross, J. J., and Menon, V. (2019). Development of human emotion circuits investigated using a big-data analytic approach: Stability, reliability, and robustness. *J Neurosci*.

4

Annexes

PLOS ONE

RESEARCH ARTICLE

Beyond executive functions, creativity skills benefit academic outcomes: Insights from Montessori education

Solange Denervaud[1,2,3,4]*, **Jean-François Knebel**[3], **Patric Hagmann**[4], **Edouard Gentaz**[1,2]

1 The Center for Affective Sciences (CISA), Campus Biotech, University of Geneva, Geneva, Switzerland, **2** Faculty of Psychology and Educational Sciences (FAPSE), University of Geneva, Geneva, Switzerland, **3** The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology and Department of Clinical Neurosciences, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), Lausanne, Switzerland, **4** Connectomics Lab, Department of Radiology, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), Lausanne, Switzerland

* solange.denervaud@chuv.ch

Citation: Denervaud S, Knebel J-F, Hagmann P, Gentaz E (2019) Beyond executive functions, creativity skills benefit academic outcomes: Insights from Montessori education. PLoS ONE 14 (11): e0225319. https://doi.org/10.1371/journal.pone.0225319

Editor: Ronald B. Gillam, Utah State University, UNITED STATES

Received: March 15, 2019

Accepted: November 1, 2019

Published: November 21, 2019

Data Availability Statement: All relevant data are uploaded to Zenodo: https://zenodo.org/record/3445416#.XcnAOdVKiUk0 DOI: http://doi.org/10.5281/zenodo.3445416.

Funding: SD was supported by the National Center of Competence in Research (NCCR) Affective sciences financed by the Swiss National Science Foundation (no51NF40-104897). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Studies have shown scholastic, creative, and social benefits of Montessori education, benefits that were hypothesized to result from better executive functioning on the part of those so educated. As these previous studies have not reported consistent outcomes supporting this idea, we therefore evaluated scholastic development in a cross-sectional study of kindergarten and elementary school-age students, with an emphasis on the three core executive measures of cognitive flexibility, working memory update, and selective attention (inhibition). Two hundred and one (201) children underwent a complete assessment: half of the participants were from Montessori settings, while the other half were controls from traditional schools. The results confirmed that Montessori participants outperformed peers from traditional schools both in academic outcomes and in creativity skills across age groups and in self-reported well-being at school at kindergarten age. No differences were found in global executive functions, except working memory. Moreover, a multiple mediations model revealed a significant impact of creative skills on academic outcomes influenced by the school experience. These results shed light on the possibly overestimated contribution of executive functions as the main contributor to scholastic success of Montessori students and call for further investigation. Here, we propose that Montessori school-age children benefit instead from a more balanced development stemming from self-directed *creative execution*.

Introduction

In a professional context where artificial intelligence is expected to surpass humans in the execution of routine tasks, we need to ensure that pedagogical approaches support a workforce capable of *creative executions* to retain its cooperative advantage and benefit from technological advances in autonomy and freedom. Some argue that a human advantage is fundamentally

Competing interests: The authors have declared that no competing interests exist.

his ability to create; to efficiently execute *individually-driven* thoughts [1]. In the traditional and dominant pedagogical system in Western countries, the metric of school success relies mostly on academic outcomes (e.g., PISA [2]), directly encouraging executive abilities. The drawback of such measures focusing solely on performance and execution is to relegate to the background a more global and integrated child development assessments, which may well also be very important to address the challenges ahead. Conversely, some alternative pedagogical approaches, such as Freinet, Waldorf, Montessori [3, 4, 5], do not target performance per se and tend to address school curricula in a more global and interdisciplinary fashion. They may address academic development differently.

The Montessori pedagogy was born of years of empirical observations of self-directed activities from developing children [6] and feature multi-age classes and a focus on peer-to-peer teaching. Children are free to choose their own learning activities from a specific set of sensory and self-corrective materials, without external feedback such as grades or evaluations [6, 7]. Many pedagogical aspects of the Montessori approach were individually shown to require and train executive functions (EFs), such as goal-directed movement, sequence of gestures to be memorized and repeated in new contexts, and so on [8–10]. Based on preliminary evidence showing that young Montessori schoolchildren (5 years old, on average) achieved higher scores at a card-sorting task [11] than children from traditional schools, it was hypothesized that a Montessori curriculum should more effectively promote EF development [12]. A second longitudinal study of kindergarten children reported some effect on EFs over three years, but not as strong as one could have expected [13]. An exploratory pretest/posttest assessment in a small sample of Montessori preschoolers revealed an improvement that was correlated not with age but with the time spent within the Montessori environment, and beyond the national normed data [14]. However, it cannot be inferred that this advantage is specific to the Montessori setting, as schoolchildren were issued from one single class, this could be a confound with a teacher-effect. While these studies do not report clear and robust effects on EFs, they do not discard this possibility, and investigating EF outcomes in older Montessori students could confirm this hypothesis.

On the other hand, despite no clear differences in EFs, Montessori students were reported to have increased scholastic outcomes, higher creativity skills as well as better well-being at school [11, 13, 15, 16]. Notably, Lillard and Else-Quest (2006) [11] have shown, through a lottery design in U.S. public schools, that children who received a Montessori-based education exhibited cognitive and socio-emotional advantages. These benefits are sometimes debated [17, 18] but seem reproducible as long as the quality and fidelity of pedagogical implementation is observed [19]. In addition, a French study reported advantages for Montessori pupils, regarding both divergent (deriving new elements from a single element) and convergent (integrating diverse elements into a new, single element) creativity over a period of 2 years in children ranging from 6 to 10 years of age [16]. Finally, a more recent and randomized study [13], followed children over the three years of public preschool. Children improved faster in academic achievement, social understanding, and mastery orientation.

In summary, there is core evidence that Montessori schoolchildren score higher on scholastic tasks than traditionally schooled children, but without displaying a definite gain in EFs. Either the EF measures were not yet sensitive enough (measuring combined instead of separated core EFs), or these scholastic performance differences do not rely on EFs alone. In this study, we tested whether reported findings held in another socio-cultural environment, namely Switzerland, while emphasizing the three core EF measures (selective attention, working memory, and cognitive flexibility). We further investigated how creativity, well-being at school, and executive functions mediate academic outcomes. Finally, we assessed the global development of both groups.

We addressed these questions in a large cohort of 201 schoolchildren (M_{age} = 9.01 years old, SD = 2.34, 96 girls and 105 boys) through a controlled observational study. As there are no public Montessori schools in Switzerland, and accordingly no option for a lottery design study, we matched pupils from Montessori private schools with peers from traditional public schools controlling for their SES, fluid intelligence, and age.

Materials and methods

Participants and procedures

The study's experimental design was based on existing literature [11, 16]. It was conducted in accordance with the Declaration of Helsinki and ethical committee from the Psychology and Education Faculty, University of Geneva (First approved on the 3rd of December 2015 under the name "Evaluation comportementale des compétences cognitives et émotionnelles chez les enfants de 5–6 ans, 9–10 ans et 12–13 ans scolarisés dans différents environnements pédagogiques -Montessori et Système Traditionnel-"). Teacher participation was voluntary.

Montessori private schools were selected according to the criteria set by the International Montessori Association (S1 Text). For the control group, traditional public schools were selected in specific areas, given the city's official statistical data on mean salary to include the upper class–salary population only, and were controlled to apply the official local study plan. In total, 21 different classes (13 Montessori classes and 8 traditional classes) from 10 different schools (5 Montessori schools and 5 traditional schools) were included in the study. The 30 teachers who participated were equally experienced (in each group, one teacher was in the early stage of her career, and all others were in the mid-to-late stages of their careers) and trained across systems (all teachers had graduated with an official pedagogical diploma).

Written consent was obtained for each child from his or her parent. Selection criteria included age group (from kindergarten age up to 7 years old, and from elementary age up to 13 years old) and school system (children had to have been enrolled in their school system since the year of their fourth birthday, or for at least 3 years). In total, 208 children were enrolled.

Data from children reported to benefit from psychological support because of learning difficulties (n = 2), with low fluid intelligence or low socio-economic status (lower than 2 standard deviations [SDs] from the mean; n = 2), outside the target age range (more than 13 years old; n = 2) as well as data from nonnative French speakers (as reported by parents or teachers; n = 1), were excluded from the study. In total, 201 children from 4.37 to 13.40 years of age (M_{age} = 9.01 years old, SD = 2.34, 96 girls and 105 boys) were retained for the study. Ninety-nine (99) participants were schooled in the Montessori educational system (54 girls; Table 1), while 102 were enrolled in the traditional group (42 girls). Descriptive check of age confirmed a bimodal distribution (S1 Fig); children were then assigned to either the kindergarten (M_{age} = 5.9, SD = 0.82, 4.4–7.8 years old) or elementary group (M_{age} = 10.3, SD = 1.4, 7.6–13.4 years old), according to their current school enrollment (Table 2).

Table 1. Study participants.

Control variable (N = 201)	Montessori (n = 99)	Traditional (n = 102)
Age (SD)	8.91 (2.40)	9.10 (2.28)
Age min, max	4.37, 13.37	4.62, 13.28
Gender, # of girls	42	54
Fluid intelligence	30.5 (7.18)	29.4 (6.63)
Socio-economic status	0.70 (0.11)	0.70 (0.12)

https://doi.org/10.1371/journal.pone.0225319.t001

Table 2. Study participant subgroups.

Kindergarten	Montessori (n = 30)	Traditional (n = 28)
Age (SD) Min, max	5.93 (0.89) 4.37–7.83	5.87 (0.75) 4.62–7.83
# of girls	16	16
Fluid intelligence	22.8 (8.79)	21.7 (7.26)
Socio-economic status (SD)	0.64 (0.12)	0.70 (0.13)
Elementary	**Montessori (n = 69)**	**Traditional (n = 74)**
Age (SD) Min, max	10.22 (1.53) 7.69–13.4	10.30 (1.21) 7.58–13.3
# of girls	26	38
Fluid intelligence	33.8 (1.98)	32.4 (3.09)
Socio-economic status (SD)	0.73 (0.09)	0.70 (0.11)

https://doi.org/10.1371/journal.pone.0225319.t002

Children were tested in schools in a dedicated, separated room. Tasks were either paper or computer based. The total duration of the experiments was of 2 hours, interrupted by brief breaks depending on the participant's fatigue.

Group comparison

To ensure the homogeneity of the two groups, we controlled for age, gender, socio-economic status, and fluid intelligence.

(i) Socio-economic status (SES). SES was assessed through a parental questionnaire [20] based on education level and both professional situation and category that 79% of parents filled out.

(ii) Fluid intelligence (FI). FI evaluation was made with the help of a black-and-white version of Raven's Progressive Matrices (PM-47) test [21] (S2 Fig). The task comprised 36 items. For each item, an incomplete matrix was presented, and the child was asked to identify the missing element that completes the matrix. Each correct item granted 1 point, with the maximum score being 36.

Scholastic assessment

Each child's global scholastic development was evaluated using well-established metrics based on four aspects (i–iv): executive functions, academic outcomes, well-being at school, and creativity.

(i) Exectuive Functions (EFs). EFs were evaluated with the help of two different types of tasks: (a) selective attention (inhibition) and (b) cognitive flexibility measures were derived from reaction time (RT) of the Flanker fish task (a child-friendly version of the flanker task, where arrows are replaced by fishes) [22]. In this particular experiment (performed using Presentation® software), the child was asked to indicate the orientation of fish (replacing the pointing arrow) by pressing keys during three different blocks. Rules were switched from the first block (focus on the fish at the center of a line of five blue fishes—17 trials) to the second block (focus on the four fish flanking the central one, all pink—17 trials). The final block randomly mixed both instructions (line of five blue fish or five pink fish for 45 trials). Response time limit was 2,000 ms for children up to 6 years old and 1,500 ms for older children. Trials with valid RT (within 2 SD) were computed as follows: for selective attention (inhibition), mean RT of congruent trials were subtracted from mean RT of incongruent trials within the first block. For cognitive flexibility, switching was computed as the mean of RT differences between successive blocks with a switch in the rules (i.e., from a line of blue fish, to a line of pink fish, last block only). (c) Working memory update was measured from the Ascending Digit (up to 6 years old) or Digit-Letter (more than 6 years old) span tasks (item from the

WISC-IV)[23]; the child was asked to listen to and memorize a string of mixed digits or mixed digit-letters, and to repeat them in an orderly ascending manner. The game started with a two-digit string; when the child successfully performed two trials in a row, an extra digit was added to the string. If the child missed a trial, a digit was removed from the string. If the child missed either three trials in a row or three trials at a single level, the game ended. The final score was age standardized.

(ii) Academic outcomes. Academic outcomes were assessed using both literacy and numeracy standardized tasks. Younger children (up to 6 years old) were evaluated through oral comprehension [24], early reading competence [25], and verbal problems [26]. (a) Oral comprehension: 27 items from Pierre Lecocq's "Epreuve de Compréhension Syntaxico-sémantique" (E.CO.S.SE) to evaluate oral comprehension were selected. Children were told a sentence and had to select among four pictures the one corresponding to that sentence. Correct responses were summed to obtain the final score (maximum 25; S3 Fig). (b) Early reading competence: First, phonemic and syllabic awareness was measured using items as cited in Gentaz et al. (2013). The child was told a pseudoword and had to repeat the same pseudoword without the first syllable (10 items) or the first phoneme (24 items). Length and difficulty increased throughout the task (maximum score, 34). Second, each child performed a decoding task (Word attack); reading 30 pseudowords within 1 minute (maximum score, 30). Accuracy across language tasks was summed and expressed as a percentage. (c) Verbal problem: children were told orally 10 different verbal problems and had to report their answer each time (S4 Fig). Accuracy (0 or 1) was computed, the final score being the sum with a maximum of 10 and expressed as a percentage of accuracy.

For older children, we evaluated language and mathematical skills through standardized competence scales [27]. (a) Language competence: Based on a story the child was asked to first read, several skills were successively tested: reading comprehension (questions on the story), grammar, and spelling tests. The maximum score was 100% of correct answers. (b) Mathematical competence: The child had to perform some arithmetical, logical, and geometric tasks. The maximum score was 100% of correct answers.

(iii) Self-reported well-being at school. Well-being at school was evaluated through questionnaires. Children up to 6 years old answered the "Feeling about School" questionnaire [11] using a graduated faces scale (from a very sad face to a very happy one) corresponding to a 5-level Likert scale. Older children filled out the Buss and Plomin questionnaire for the sociability measure [28]. Children answered statements about their feelings using a Likert scale ranging from 0 to 4. The final score was expressed as a percentage.

(iv) Creativity. Creativity was measured using both divergent and convergent abstract drawing items from a standardized test [29]. (a) Divergent creativity: The child was asked to draw as many different drawings as possible from one imposed abstract form, within a time frame of 10 minutes. The final score was the sum of all valid creations, where the initial imposed abstract form was correctly integrated within a new concept. (b) Convergent creativity: The child was asked to pick at least three different abstract forms out of eight and to create one new drawing that combined them, within a time frame of 15 minutes. Drawings were blindly scored by three different judges following the referenced scale (maximum of 7, from 1 = very poor creativity to 7 = highly creative). Criteria were originality and storytelling of the drawing. The final score was expressed as a percentage.

Statistical analysis

Statistical analyses were computed using R, and, in part, jamovi (Version 0.9) Computer Software.

Group comparison. Prior to group comparison, statistical t tests were run on the control variables (age, fluid intelligence, and socio-economic status) to ensure group homogeneity (S1 Table).

Scholastic outcomes.

T tests. Assuming a selection bias, scores per task were tested statistically using bootstrapping Yuen t test [30] with 20% trimming and 600 repetitions for bootstrapping. This test was used to determine significant differences between the two groups of schoolchildren (Montessori vs. traditional) at the two age levels (kindergarten and elementary), with a false discovery rate (FDR) p-value correction at $q = 0.05$. Additionally, we controlled for age by running ANCOVA on each measure with age as a covariate.

Multiple mediation model. Z-scored data from the same cognitive measure (executive functions, academic outcomes, well-being at school, or creativity skills) were averaged across subjects. A multiple mediator model was built and computed on the pooled data to evaluate the effects of multiple factors (executive functions, creativity skills, well-being at school) simultaneously on academic outcomes, when the predictor was school system (contrast Montessori-Traditional). The full model was *Academic outcomes ~ executive functions + well-being at school + creativity skills + system*, and the mediator model was *executive functions~system (M-T)*, *well-being at school~system (M-T)*, and *creativity skills~system (M-T)*. We used the large sample z-test of the mediated effect, known to be slightly more accurate than the Sobel test, with 1,000 bootstrap repetitions (percentile method) [31].

Radial plot. Finally, a radial plot was designed to qualitatively represent the scholastic development of each child and the mean for both groups (Montessori or traditional) with the pooled dataset. There were four axes in the radial plot; each edge standing for the maximal score possible for the core skills (academic outcomes, EFs, creativity, and well-being), and the center standing for the minimal score for all the skills. Each child's averaged z-score was reported as a distance along each axis and joined between axes.

Results

Children were proficient at all tasks, and no one was excluded due to missing data or outlier outcomes. The scores were individually computed and reported before the statistical comparison between the two groups at both school-level, controlling for age. We then built the multiple mediation model to investigate the relationships between EFs, creativity, well-being at school, and academic outcomes. Finally, we plotted a qualitative measure of global scholastic development through the radial representation.

Scholastic outcomes

At kindergarten age, between group comparison revealed that, even when controlling for age, Montessori schoolchildren score higher than same-age children from traditional schools on language, math, well-being, working memory, convergent and divergent creativity tasks (Table 3, top panel). At elementary age, results revealed that language, math, working memory, convergent and divergent creativity scores were higher in the Montessori schoolchildren than in same-age children from traditional schools, even when controlling for age (Table 3, bottom panel). Our findings are of medium to large effect sizes (Cohen's d), that are at least comparable to previous studies comparing Montessori and traditional schoolchildren (Table 3, right column).

Multiple mediation model

There was a significant indirect effect of creativity skills only on academic outcomes, $z > 2$, $p = 0.04$. As Fig 1 illustrates, for Montessori schoolchildren, the standardized regression

Table 3. Scores per age level and group (mean, SD), and statistics.

Kindergarten	Montessori mean (SD)	Traditional mean (SD)	Yuen's test bootstrapped (*p*-values FDR corrected)	Main effect of pedagogy when controlling for age (ANCOVA)	Effect size Cohen's *d*	Effect size Cohen's *d* from randomized studies [11, 13]
Language (%)	**66.1 (26.7)**	**51.8 (23.8)**	**2.05, p = 0.06**	**5.26, p = 0.026**	0.56	0.44 (Letter-Word) & 0.63 (Word Attack); 0.36 & 0.41 (Academic achievement at time 1 and time 2)
Math (%)	**45.1 (27.8)**	**23.9 (31.0)**	**3.52, *p* = 0.012**	**8.66, *p* = 0.005**	**0.72**	0.55 (Applied problem)
Well-being at school (%)	**87.2(12.0)**	**75.8 (13.9)**	**3.69, *p* = 0.008**	**11.13, *p* = 0.002**	**0.88**	
Convergent creativity (score)	**3.88 (1.49)**	**2.74 (1.27)**	**3.54, *p* = 0.013**	**11.8, *p* = 0.001**	**0.82**	
Divergent creativity (score)	**6.63 (4.32)**	**3.36 (2.72)**	**2.89, *p* = 0.016**	**13.6, *p* < 0.001**	**0.90**	
Working memory (score)	**5.30 (1.85)**	**4.16 (1.56)**	**3.13, *p* = 0.016**	**7.01, *p* = 0.010**	**0.66**	0.61 (Dimensional Card Sort); 0.35 (at time 3 for the Head-Toes-Knees-Shoulders and Copy-Design tasks)
Selective attention (ms)	74.5 (203)	144 (245)	−1.77, *p* = 0.100	1.29, *p* = 0.260	0.31	
Cognitive flexibility (ms)	46.7 (104)	31.0 (76.7)	0.22, *p* = 0.380	0.355, *p* = 0.554	0.17	
Elementary	**Montessori mean (SD)**	**Traditional mean (SD)**	**Yuen's test bootstrapped (*p*-values FDR corrected)**	**Main effect of pedagogy when controlling for age (ANCOVA)**	**Effect size Cohen's *d***	**Effect size Cohen's *d* from randomized studies**
Language (%)	**74.4 (14.8)**	**57.6 (26.7)**	**3.74, *p* = 0.004**	**29.0, *p* < 0.001**	**0.78**	
Math (%)	**66.1 (25.0)**	**45.1 (26.7)**	**4.28, *p* = 0.003**	**30.6, *p* < 0.001**	**0.81**	
Well-being at school (%)	65.7(20.7)	63.4 (19.7)	0.42, *p* = 0.77	0.60, *p* = 0.442	0.12	0.54 (positive school feeling)
Convergent creativity (score)	**5.13 (1.46)**	**3.53 (1.51)**	**6.07, *p* = 0.003**	**43.93, *p* < 0.001**	**1.07**	0.71 (Creativity of narrative)
Divergent creativity (score)	**10.8 (4.08)**	**7.42 (4.65)**	**4.61, *p* = 0.003**	**22.76, *p* < 0.001**	**0.76**	
Working memory (score)	**7.32 (2.12)**	**6.29 (2.35)**	**1.99, p = 0.053**	**7.79, p = 0.006**	0.46	
Selective attention (ms)	21.8 (85.4)	7.18 (87.6)	1.38, *p* = 0.24	0.94, *p* = 0.335	0.17	
Cognitive flexibility (ms)	51.3 (68.6)	43.7 (54.5)	0.12, *p* = 0.900	0.44, *p* = 0.508	0.12	

https://doi.org/10.1371/journal.pone.0225319.t003

coefficient between system and creativity as well as creativity and academic outcomes were statistically significant ($p < 0.001$ and $p = 0.036$, respectively). The standardized indirect effect was ß = 0.39 ($p < 0.001$), with partial mediation of Montessori system by creativity on the academic outcomes. Of note, the standardized regression coefficient between system and executive functions was not significant ($p = 0.18$) (S2 Table).

Radial representation

Fig 2A depicts the radial representation of the group means. Each core skill *z*-score center is tracked with the red line, allowing a visual assessment of mean global development. The pattern shows that only creativity skills and academic outcomes differ between groups in favor of Montessori schoolchildren. Fig 2B (Montessori schoolchildren) and Fig 2C (traditional schoolchildren) are radial plots, where individual outcomes are depicted; of note, there is a visible difference on the "creativity" corner.

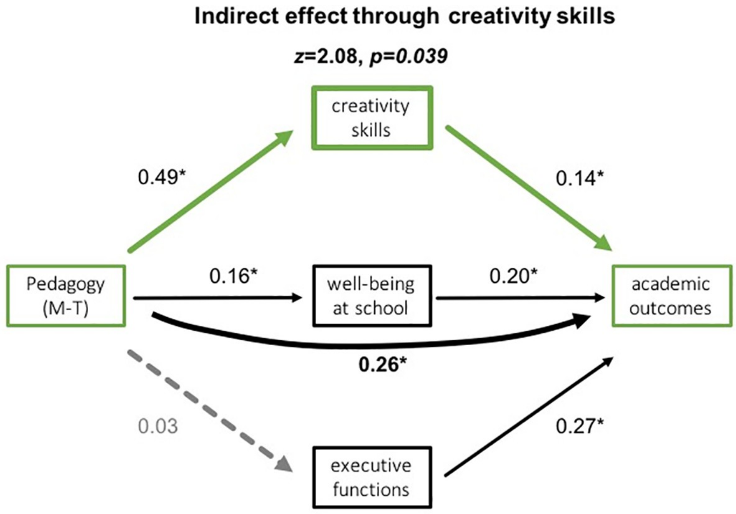


Fig 1. Multiple mediation model (according to [31]) for the indirect effect of children's school system (Montessori vs. traditional) via multiple mediators (executive functions, well-being at school, and creativity skills) on academic outcomes. The only significant ($z > 2$) indirect mediation effect on academic outcomes was creativity skills in Montessori schoolchildren (green path). The standardized solution coefficients (ß) and significant p-values < 0.05 (depicted with a star) are reported next to related path.

https://doi.org/10.1371/journal.pone.0225319.g001

Discussion

We evaluated cognitive measures that were studied separately in previous works done in the field, comparing Montessori with traditional schoolchildren on scholastic, creativity and well-being outcomes [11, 13, 15, 16].

Regarding scholastic and creativity scores, our findings corroborate previous studies [11, 16] but in a different cultural environment, suggesting that some of the measured effects could reflect the child pedagogical experience in a Montessori setting. Kindergarten Montessori schoolchildren also reported a better sense of well-being at school than schoolchildren from traditional pedagogy. This is in line with previous studies using the same tasks; however, we did not find a similar difference among elementary schoolchildren. Based on the existing literature reporting that the Montessori pedagogy promote students' sense of belonging to the school [11, 15], with higher autonomy usually leading to well-being [32], this result is contrary to our expectations. This may reflect more generally a developmental shift in how schoolchildren orient and evaluate their social interest at school; from the teacher at kindergarten-age to their peers from 6 years old onwards. The general attenuation in well-being with age may thus reflect the usual appearance of socio-cognitive conflicts in children and/or the social bias of self-reported questionnaires [33].

Concerning EFs, which include cognitive flexibility, working memory update, and selective attention (inhibition) according to Miyake's model [34], no difference was found between school settings. The exception was for working memory, which was found to be different in favor of Montessori students. As opposed to cognitive flexibility and inhibition, which were measured based on RT (speeded response task) through a computerized task, working memory was measured as a score (no time restriction). Time limit and/or the screen interface could artifact the outcomes, since Montessori schoolchildren are not accustomed to this type of activity within their school environment, nor to work under time pressure. Previous studies making use of screen-free tasks with no account for RT reported advantages for the Montessori schoolchildren. For example, the Head-Toes-Knees-Shoulders and Copy-Design tasks

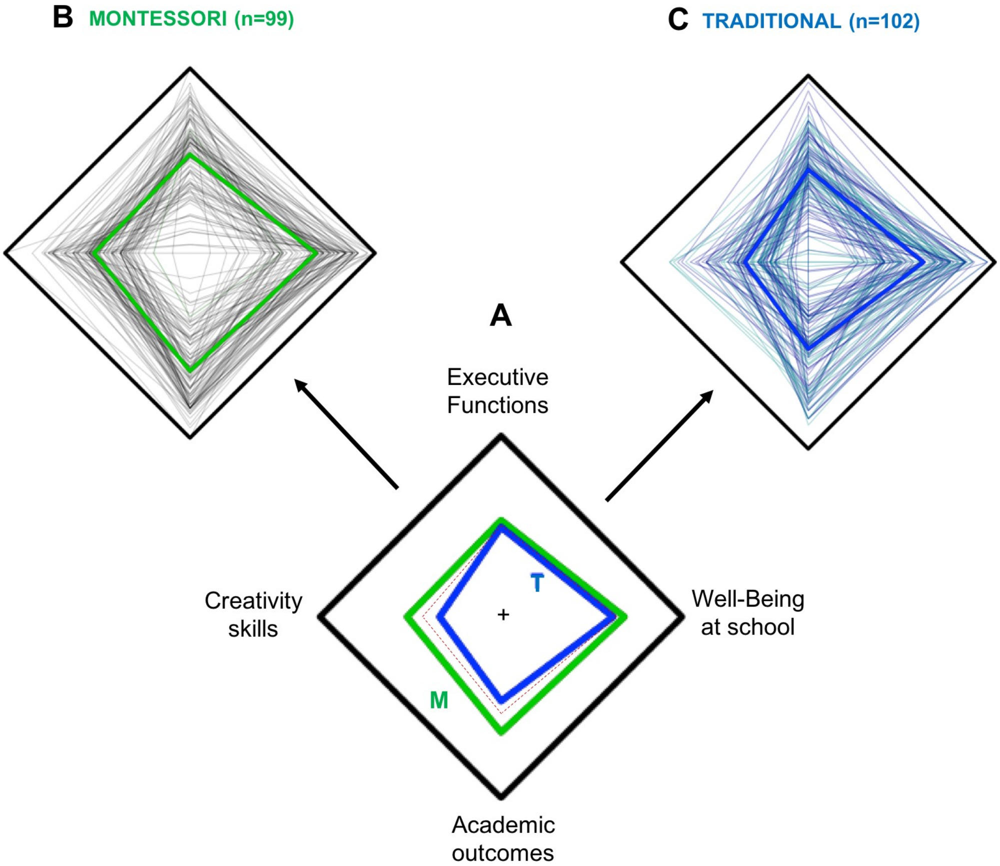


Fig 2. Radial qualitative representation of the four different cognitive measures: executive functions, creativity skills, well-being at school, and academic outcomes, each located at a summit. The scales depend on the measured cognitive skill; however, all run from the minimum at the center to the maximum at the border of the square. Individual results are represented with a thin line (Montessori schoolchildren on top left 2B, and control on the top right 2C), and mean for each group is reported with a bold line in the central square 2A, where the dotted red line marks the 0 of each cognitive measure's z-score scale. Montessori (M) depicted in green, traditional (T) in blue. Group differences (M vs. T) are observed for creativity skills and academic outcomes.

https://doi.org/10.1371/journal.pone.0225319.g002

showed an improvement over two years in Montessori schoolchildren compared to controls from traditional schools [13]. Either these tasks pull more on the child's working memory capacity, or time/computer constrains their actual competences. We further addressed the issue of timing by looking at the error-rate of the flanker task instead of RT, but none of these analyses revealed group differences ($p>0.5$). Another possible explanation for the absence of a clear global EF difference in our cohort could also stem from the known relationship between EFs and SES [35]. Indeed, in the context of our study, participants come from relatively high-income family environments, which likely influenced their EF capacity in a similar way. In addition, these children attend schools with high-quality settings, where great emphasis was placed on EF trainings in the last decade.

While there are no differences in EFs, their self-monitoring could still differ. Empirical studies describe Montessori schoolchildren with the capacity for a deep concentration state [9,18], which certainly rely on combined self-regulatory features rather than just selective attention capacity (high focus). In this context, it would be of interest to measure self-directed EF [36], as the Montessori children are trained for more autonomous thinking behaviors that could promote more "intrinsically" driven executive control, and also explain their higher creativity skills.

This was further explored through the multiple mediator analysis. Our data shows that beyond EFs, creative competencies specifically modulate academic success in Montessori schoolchildren, suggesting good execution of self-generated ideas. This aspect of EFs is currently understudied in the framework of academic outcomes in school years.

Finally, pupils attending a Montessori school were shown to have a more balanced global development. This may play a key role in promoting academic performance. This finding raises the question of the limit of emphasizing a unique aspect of scholastic development, such as EFs. In fact, cognition with less control (lower EFs) as during the childhood years presents many advantages, such as faster learning rate and higher creative abilities [37]. Seeking for cognitive performance may be at the cost of qualitative and long-term learning[38]. Accordingly, expecting schoolchildren to maintain a high level of selective attention, or placing too much emphasis on other core EFs, may well be counterproductive and impair the individual's innate capacity for learning and creative execution abilities.

More broadly, one can wonder whether educating and directing competencies in isolation does not prevent schoolchildren from making connections or unrelated links later on, and thus prevent from nurturing individual creative thinking. Indeed, creativity is frequently attributed to genius or pure talent—an innate spark found only in the Albert Einsteins, Pablo Picassos, or Steve Jobses of this world. However, creative thinking is a fundamental competency, present in all of us to different degrees, and something that can be nurtured. We need to address and educate *creative execution abilities*, not simply by allocating more hours for painting or crafting within curricula (there is little of these activities within the Montessori education, for instance), but rather by investigating which aspects of pedagogical approaches fostering creativity, such as Montessori, make it possible for schoolchildren to grow this way of thinking. We suspect that it results from a combination of features more than one; such as using more naturalistic activities that are inherently inter-disciplinary, interacting with peers from different ages, making a choice amongst different activities, taking the lead over projects, or seeking for answers and solutions on their own. These pedagogical aspects are not easy to capture scientifically and will highly benefit from extensive multimodal research in the future.

The main limitation of our study is the fact that, due to local policies in Switzerland, the Montessori classes included in the study are all in private schools, whereas the traditional schools are public. We chose public schools in areas of similar wealth to that of Montessori school candidates and controlled for their SES. This may constitute, despite all precautions, a selection bias. Nevertheless, our basic findings are in agreements with the two existing randomized studies made in public Montessori schools [11, 13], suggesting that this bias may be weak or negligible and that the observed effect is mainly attributable to schooling differences. However, in our effort to match the schoolchildren based on their SES, we did not account for the possible bias that parents enrolling their children within Montessori curricula could themselves present higher creative thinking. If so, interactions with their child could also influence the higher level of creativity measured in our study. Further studies should be conducted to deepen these findings and would benefit either from a longitudinal or a lottery design study instead of the use of matched controls to clarify some of the uncertainty raised in this discussion.

Supporting information

S1 Text. School selection criteria.
(PDF)

S1 Table. Control variables' t-Test. Age, Fluid Intelligence (FI) and socio-economic status (SES).
(PDF)

S2 Table. Mediation model.
(PDF)

S1 Fig. Age distribution.
(PDF)

S2 Fig. Fluid Intelligence was measured with the help of the Raven matrices task, examples displayed here.
(PDF)

S3 Fig. Oral comprehension, examples from the E.C.O.S.S.E task.
(PDF)

S4 Fig. Verbal problem examples.
(PDF)

Acknowledgments

We thank all the children, parents, teachers and school Directors who participated in this study; we are grateful to Prof. Diamond and David Abott, Ph.D., Prof. Angeline Lillard, and Prof. Todd Lubart and Maud Besançon, Ph.D., for sharing some of their tasks and advice, we also thank Morgane Budry, Diane Gervaise, Julien Fedrizzi, and Claire Battilotti for help with data collection and coding; Andres Posada, Ph.D., for technical support; Adin Ross-Gillespie, Ph.D., for his expertise in statistical analysis and editing; Jean-Marc Randin, Prof. David Sander, and Nicolas Denervaud, Ph.D., for their critical insights, Prof. Liliane Sprenger-Charolles and Prof. Kerstin von Plessen for comments on an earlier version of this manuscript; both reviewers for highly constructive remarks; Nora Yang, Tim Pfahlert, and Cambridge Proofreading for editing the final version of this manuscript.

Author Contributions

Conceptualization: Solange Denervaud, Edouard Gentaz.

Data curation: Solange Denervaud.

Formal analysis: Solange Denervaud, Jean-François Knebel, Patric Hagmann.

Funding acquisition: Edouard Gentaz.

Investigation: Solange Denervaud.

Project administration: Solange Denervaud, Edouard Gentaz.

Supervision: Patric Hagmann.

Visualization: Solange Denervaud, Patric Hagmann.

Writing – original draft: Solange Denervaud.

Writing – review & editing: Jean-François Knebel, Patric Hagmann, Edouard Gentaz.

References

1. Puccio GJ, State University College at Buffalo. Center for Studies in Creativity. Creativity rising: creative thinking and creative problem solving in the 21st century. Buffalo, N.Y.: ICSC Press, International Center for Studies in Creativity; 2012. 144 p. p.

2. Grisay A, de Jong JH, Gebhardt E, Berezner A, Halleux-Monseur B. Translation equivalence across PISA countries. J Appl Meas. 2007; 8(3):249–66. Epub 2007/09/07. PMID: 17804893.
3. Vygotsky L. Mind in society: The development of higher psychological processes. Cambridge: MA: Harvard University Press; 1978.
4. Piaget J. The origins of intelligence in children. New York: International Universities Press; 1952.
5. Condliffe B, Quint J, Visher MG, Bangser MR, Drohojowska S, Saco L, et al. Project-Based Learning; A Literature Review. New York: 2017.
6. Montessori M. The Secret of Childhood. 1981 ed. New York: Ballantine; 1936.
7. Marshall C. Montessori education: a review of the evidence base. npj Science of Learning. 2017; 2(1). https://doi.org/10.1038/s41539-017-0012-7 PMID: 30631457
8. Diamond A. Activities and Programs That Improve Children's Executive Functions. Curr Dir Psychol Sci. 2012; 21(5):335–41. Epub 2012/10/01. https://doi.org/10.1177/0963721412453722 PMID: 25328287; PubMed Central PMCID: PMC4200392.
9. Diamond A. Executive Functions. Annu Rev Psychol. 2013; 64:135–68. https://doi.org/10.1146/annurev-psych-113011-143750 WOS:000316383600007. PMID: 23020641
10. Diamond A. Want to Optimize Executive Functions and Academic Outcomes?: Simple, Just Nourish the Human Spirit. Minn Symp Child Psychol Ser. 2014; 37:205–32. Epub 2014/11/02. PMID: 25360055; PubMed Central PMCID: PMC4210770.
11. Lillard AS, Else-Quest N. The early years. Evaluating Montessori education. Science. 2006; 313(5795):1893–4. Epub 2006/09/30. https://doi.org/10.1126/science.1132362 PMID: 17008512.
12. Diamond A, Lee K. Interventions shown to aid executive function development in children 4 to 12 years old. Science. 2011; 333(6045):959–64. Epub 2011/08/20. https://doi.org/10.1126/science.1204529 PMID: 21852486; PubMed Central PMCID: PMC3159917.
13. Lillard AS, Heise MJ, Richey EM, Tong X, Hart A, Bray PM. Montessori Preschool Elevates and Equalizes Child Outcomes: A Longitudinal Study. Front Psychol. 2017; 8:1783. Epub 2017/11/23. https://doi.org/10.3389/fpsyg.2017.01783 PMID: 29163248; PubMed Central PMCID: PMC5670361.
14. Phillips-Silver J, Daza MT. Cognitive Control at Age 3: Evaluating Executive Functions in an Equitable Montessori Preschool. Frontiers in Education. 2018.
15. Rathunde K, Csikszentmihalyi M. Middle School Students' Motivation and Quality of Experience: A Comparison of Montessori and Traditional School Environments. American Journal of Education. 2005; 111(3).
16. Besançon M, Lubart T. Differences in the development of creative competencies in children schooled in diverse learning environments. Learning and Individual Differences. 2008; 18(4):381–9. https://doi.org/10.1016/j.lindif.2007.11.009
17. Mackinnon P. Studying students in Montessori schools. Science. 2007; 315(5812):596–7; author reply -7. Epub 2007/02/06. PMID: 17278256.
18. Lindenfors P. Studying students in Montessori schools. Science. 2007; 315(5812):596–7; author reply -7. Epub 2007/02/03. https://doi.org/10.1126/science.315.5812.596b PMID: 17272701.
19. Lillard AS. Preschool children's development in classic Montessori, supplemented Montessori, and conventional programs. J Sch Psychol. 2012; 50(3):379–401. Epub 2012/06/05. https://doi.org/10.1016/j.jsp.2012.01.001 PMID: 22656079.
20. Genoud PA. Indice de position socioéconomique (IPSE): un calcul simplifié: Fribourg University; 2011 [cited 2015]. Available from: http://www3.unifr.ch/cerf/fr/indice-de-position-socioéconomique.html.
21. Raven J, Raven JC, Court JH. Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 2: The Coloured Progressive Matrices. San Antonio, TX: Harcourt Assessment; 1998.
22. Eriksen BA, Eriksen CW. Effects of noise letters upon identification of a target letter in a non- search task. Perception and Psychophysics. 1974; 16:143–9.
23. Wechsler D. WISC-IV Echelle d'intelligence de Wechsler pour enfants et adolescents. 4th ed: ECPA; 2005.
24. Lecocq P. L'E.Co.S.Se: une épreuve de compréhension syntaxico-sémantique. Lille1996.
25. Gentaz E, Sprenger-Charolles L, Theurel A, Cole P. Reading comprehension in a large cohort of French first graders from low socio-economic status families: a 7-month longitudinal study. PLoS One. 2013; 8(11):e78608. Epub 2013/11/20. https://doi.org/10.1371/journal.pone.0078608 PMID: 24250802; PubMed Central PMCID: PMC3826761.
26. von Aster M. ZAREKI-R Batterie pour l'évaluation du traitement des nombres et du calcul chez l'enfant.: ECPA; 2005.
27. Simonart G. ECHAS; échelle d'apprentissages scolaires primaires2008.

28. Buss AH, Plomin R. Temperament: Early developing personality traits. London: Psychology Press; 1984. 196 p.
29. Lubart T, Besançon M, Barbot B. EPoC: Evaluation du potentiel créatif des enfants. France: Hogrefe 2011.
30. Yuen KK. The two-sample trimmed t for unequal population variances. Biometrika. 1974; 61:165–70.
31. Selker R. Simple Mediation and Moderation Analysis https://github.com/raviselker/medmod2017. Package].
32. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol. 2000; 55(1):68–78. Epub 2001/06/08. https://doi.org/10.1037//0003-066x.55.1.68 PMID: 11392867.
33. Rosenman R, Tennekoon V, Hill LG. Measuring bias in self-reported data. Int J Behav Healthc Res. 2011; 2(4):320–32. Epub 2011/10/01. https://doi.org/10.1504/IJBHR.2011.043414 PMID: 25383095; PubMed Central PMCID: PMC4224297.
34. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. Cognitive Psychol. 2000; 41(1):49–100. https://doi.org/10.1006/cogp.1999.0734 WOS:000088656100002. PMID: 10945922
35. Lawson GM, Hook CJ, Farah MJ. A meta-analysis of the relationship between socioeconomic status and executive function performance among children. Dev Sci. 2018; 21(2). Epub 2017/05/31. https://doi.org/10.1111/desc.12529 PMID: 28557154; PubMed Central PMCID: PMC5821589.
36. Barker JE, Semenov AD, Michaelson L, Provan LS, Snyder HR, Munakata Y. Less-structured time in children's daily lives predicts self-directed executive functioning. Front Psychol. 2014; 5:593. Epub 2014/07/30. https://doi.org/10.3389/fpsyg.2014.00593 PMID: 25071617; PubMed Central PMCID: PMC4060299.
37. Thompson-Schill SL, Ramscar M, Chrysikou EG. Cognition without control: When a little frontal lobe goes a long way. Curr Dir Psychol Sci. 2009; 18(5):259–63. Epub 2009/01/01. https://doi.org/10.1111/j.1467-8721.2009.01648.x PMID: 20401341; PubMed Central PMCID: PMC2855545.
38. Soderstrom NC, Bjork RA. Learning versus performance: an integrative review. Perspect Psychol Sci. 2015; 10(2):176–99. Epub 2015/04/25. https://doi.org/10.1177/1745691615569000 PMID: 25910388.

1. SUPPORTING INFORMATION

School selection criteria

The criteria set by the International Montessori Association (AMI) (https://montessori-ami.org):

(i) all teachers were AMI trained,

(ii) a complete set of Montessori material was available in each classroom,

(iii) children had a 3-hour continuous working time, and

(iv) there were at least 3 different age-levels per class.

The official local study plan for traditional public schools strictly implies, as from 6 years old:

(i) frontal teaching,

(ii) tests and formal evaluations,

(iii) breaks every hour, and

(iv) one age-level per class.

N.B. In Switzerland, it is quite common to have two part-time teachers that share the lead of one class, which was the case for many of the public traditional classes included in the study. In addition, Montessori classes had often two teachers per class, if the latter was large (according to the Swiss law).

Control variable	t	df	p	Cohen's d
Age	-0.59	199	0.56	-0.08
FI	1.09	199	0.28	0.15
SES	0.92	156	0.36	0.15

Table S1: Control variables' t-Test. Age, Fluid Intelligence (FI) and socio-economic status (SES)

Models Info

Mediators Models		
	m1	Creativity skills ~ Pedagogy
	m2	Well-being at school ~ Pedagogy
	m3	Executive functions ~ Pedagogy
Full Model		
	m4	Academic outcomes ~ Creativity skills + Well-being at school + Executive functions + Pedagogy
Indirect Effects		
	IE 1	Pedagogy ⇒ Creativity skills ⇒ Academic outcomes
	IE 2	Pedagogy ⇒ Well-being at school ⇒ Academic outcomes
	IE 3	Pedagogy ⇒ Executive functions ⇒ Academic outcomes

Indirect and Total Effects

				95% C.I. (a)				
Type	**Effect**	**Estimate**	**SE**	**Lower**	**Upper**	**β**	**z**	**p**
Indirect	**Pedagogy1 ⇒ Creativity skills ⇒ Academic outcomes**	**0.0304**	**0.01465**	**-0.00257**	**0.0562**	**0.0684**	**2.07**	**0.038**
	Pedagogy1 ⇒ Well-being at school ⇒ Academic outcomes	0.0141	0.00793	0.00126	0.0324	0.0317	1.78	0.076
	Pedagogy1 ⇒ Executive functions ⇒ Academic outcomes	0.0115	0.00970	-0.00644	0.0324	0.0258	1.18	0.238
Component	**Pedagogy1 ⇒ Creativity skills**	**0.1796**	**0.02208**	**0.13393**	**0.2228**	**0.4870**	**8.13**	**< .001**
	Creativity skills ⇒ Academic outcomes	**0.1691**	**0.08028**	**-0.01448**	**0.3086**	**0.1405**	**2.11**	**0.035**
	Pedagogy1 ⇒ Well-being at school	**0.0630**	**0.02812**	**0.00627**	**0.1187**	**0.1585**	**2.24**	**0.025**
	Well-being at school ⇒ Academic outcomes	**0.2235**	**0.06837**	**0.08783**	**0.3556**	**0.1999**	**3.27**	**0.001**
	Pedagogy1 ⇒ Executive functions	0.0385	0.02980	-0.02243	0.0934	0.0949	1.29	0.196
	Executive functions ⇒ Academic outcomes	**0.2974**	**0.06360**	**0.17958**	**0.4324**	**0.2719**	**4.68**	**< .001**

Indirect and Total Effects

Type	Effect	Estimate	SE	95% C.I. (a) Lower	Upper	β	z	p
Direct	**Pedagogy1 ⇒ Academic outcomes**	**0.1167**	**0.02993**	**0.06126**	**0.1801**	**0.2629**	**3.90**	**< .001**
Total	Pedagogy1 ⇒ Academic outcomes	0.1726	0.02905	0.11563	0.2295	0.3873	5.94	< .001

Note. (a) Confidence intervals computed with method: Bootstrap percentiles Pedagogy 1 stands for M-T.

Table S2: Mediation model.

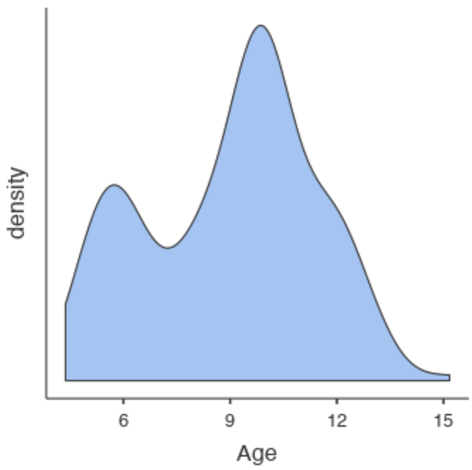


Figure S1: Age distribution.

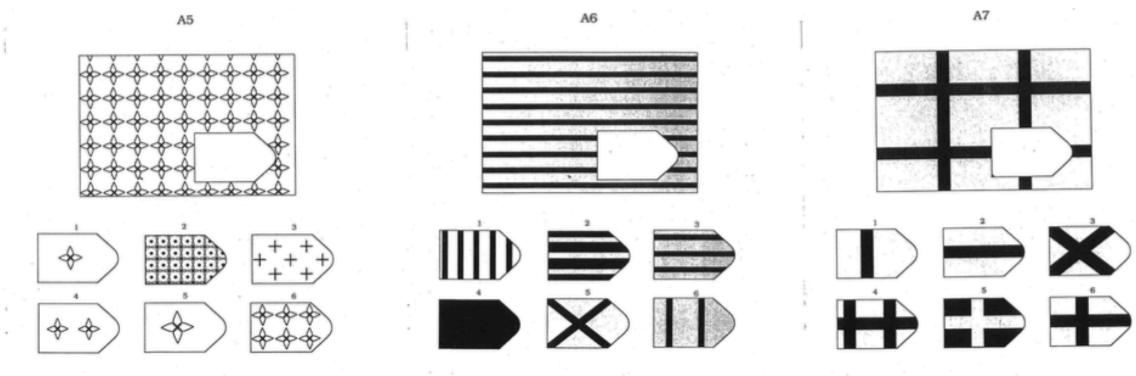


Figure S2: Fluid Intelligence was measured with the help of the Raven matrices task, examples displayed here.

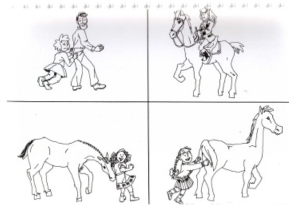
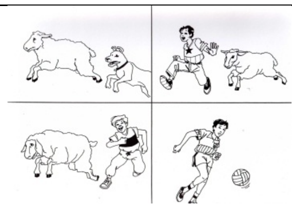
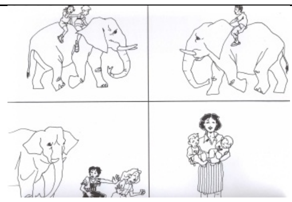
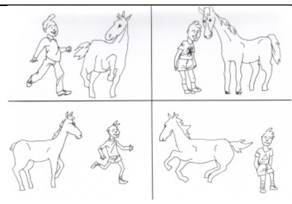
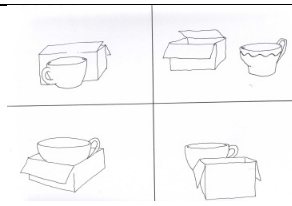
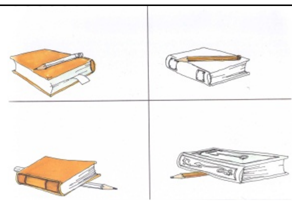
E.CO.S.SE	Examples of sentence for oral comprehension (CO)	Pictures
Active sentences	**OC**- La fille pousse le cheval. *The girl pushes the horse*	
Passive sentences	**OC** - La fille est poursuivie par le cheval. *The girls is pursued by the horse*	
Pronouns	**OC** - L'éléphant les porte. *The elephant carries them*	
Double negation	**CO** - Ni le garçon ni le cheval ne courent. *Neither the boy nor the horse run*	
Spatial relation	**CO** - Le crayon est derrière la boîte. *The pencil is behind the box*	
Embedded relative clause with spatial relation	**CO** - Le crayon qui est sur le livre est jaune. *The pencil on the book is yellow*	

Figure S3: Oral comprehension, examples from the E.C.O.S.S.E task.

Oral Sentence	**Expected Answer**
"Jean a 4 cerises. Il en mange 1. Combien de cerises lui reste-t-il ?" *Jean has 4 cherries. He eats one. How many cherries are left?*	4 – 1 = **3**
"Pierre a 12 billes. Il donne 5 billes à sa copine Anne. Combien de billes a Pierre maintenant ?" *Pierre has 12 marbles. He gives 5 marbles to his girlfriend Anne. How many marbles has Peter now?*	12 – 5 = **7**
"Il y a 4 poissons dans le bocal. David ajoute des poissons. Maintenant, il y a 6 poissons dans le bocal. Combien David a-t-il ajouté de poissons ?" *There are four fishes in the jar. David adds fishes. Now there are six fishes in the jar. How many fish did David add?*	4 + x = 6 x = **2**

Figure S4: Verbal Problem examples.

www.nature.com/scientificreports

SCIENTIFIC REPORTS

natureresearch

OPEN

Multisensory Gains in Simple Detection Predict Global Cognition in Schoolchildren

Solange Denervaud[1,2], Edouard Gentaz[2,3], Pawel J. Matusz[1,4,5,8] & Micah M. Murray[1,5,6,7,8*]

The capacity to integrate information from different senses is central for coherent perception across the lifespan from infancy onwards. Later in life, multisensory processes are related to cognitive functions, such as speech or social communication. During learning, multisensory processes can in fact enhance subsequent recognition memory for unisensory objects. These benefits can even be predicted; adults' recognition memory performance is shaped by earlier responses in the same task to multisensory – but not unisensory – information. Everyday environments where learning occurs, such as classrooms, are inherently multisensory in nature. Multisensory processes may therefore scaffold healthy cognitive development. Here, we provide the first evidence of a predictive relationship between multisensory benefits in simple detection and higher-level cognition that is present already in schoolchildren. Multiple regression analyses indicated that the extent to which a child (N = 68; aged 4.5–15years) exhibited multisensory benefits on a simple detection task not only predicted benefits on a continuous recognition task involving naturalistic objects ($p = 0.009$), even when controlling for age, but also the same relative multisensory benefit also predicted working memory scores ($p = 0.023$) and fluid intelligence scores ($p = 0.033$) as measured using age-standardised test batteries. By contrast, gains in unisensory detection did not show significant prediction of any of the above global cognition measures. Our findings show that low-level multisensory processes predict higher-order memory and cognition already during childhood, even if still subject to ongoing maturation. These results call for revision of traditional models of cognitive development (and likely also education) to account for the role of multisensory processing, while also opening exciting opportunities to facilitate early learning through multisensory programs. More generally, these data suggest that a simple detection task could provide direct insights into the integrity of global cognition in schoolchildren and could be further developed as a readily-implemented and cost-effective screening tool for neurodevelopmental disorders, particularly in cases when standard neuropsychological tests are infeasible or unavailable.

When a child wishes to cross the street, simply looking left and right for incoming cars is not always sufficient to make a safe choice. Sensitivity to additional cues, like the noise generated by an approaching car, will also guide their judgement, and may save their life. There are two aspects of this capacity to integrate information from different senses that are likely themselves synergistic. First, multisensory information may accelerate perceptual decision-making and result in faster and more accurate responses (reviewed in[1–4]). Second, multisensory information may provide a more efficient means for learning and memory than unisensory stimuli, which in turn can guide future behaviour (reviewed in[5–8]). Learning in multisensory contexts is thus of clear adaptive benefit during development and throughout the lifespan, particularly given the fact that multisensory contexts are reflective of naturalistic settings[9]. It thus logically follows that the gain afforded by multisensory processes may themselves

[1]The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology, Vaudois University Hospital Center and University of Lausanne, Lausanne, Switzerland. [2]The Center for Affective Sciences (CISA), University of Geneva, Geneva, Switzerland. [3]Faculty of Psychology and Educational Sciences (FAPSE), University of Geneva, Geneva, Switzerland. [4]Information Systems Institute at the University of Applied Sciences Western Switzerland (HES-SO Valais), 3960, Sierre, Switzerland. [5]Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA. [6]Department of Ophthalmology, Fondation Asile des aveugles and University of Lausanne, Lausanne, Switzerland. [7]Sensory, Cognitive and Perceptual Neuroscience Section, Center for Biomedical Imaging (CIBM) of Lausanne and Geneva, Lausanne, Switzerland. [8]These authors contributed equally: Pawel J. Matusz and Micah M. Murray. *email: micah.murray@chuv.ch

provide a scaffold for improved higher-level cognitive functions such as learning, recognition memory, working memory, and fluid intelligence (among others) (reviewed in[10]). The aim of the present study was to assess the presence of such relationships in school-aged children.

Considerable research points to a general link between processing speed and measures of intelligence[11–14], in adults as well as in school-aged children[15]. One potential consideration with that research is that the tasks used to evaluate processing speed were all visual in nature and thus did not assess the contribution of other sensory systems to cognitive abilities, including intelligence. At the same time, it stands to reason that individuals capable of capitalising on situations that improve processing speed (e.g. multisensory contexts) should also demonstrate stronger cognitive abilities; what Rose and colleagues refer to as a "cognitive cascade"[2,16].

Longitudinal studies have linked cross-modal pattern matching in infants with their later reading abilities, such as in the seminal work of Birch and Belmont[17–19] in 220 elementary (5–12 years old) scholars. This link was then extended to infants, including those born prematurely, by Rose and colleagues. An infant's ability to match information (typically temporal patterns) between the senses is predictive of later reading skills. In particular, matching abilities *between* the senses has been shown to be a better predictor of reading skills than matching skills for patterns *within* a given sensory modality[20]. The capacity to establish sensory-independent or multisensory representations may be a core underlying skill for cognitive functions to develop and thus are indicative of core intelligence.

While the literature in infants and young children appears to support links between multisensory processes and higher-level cognition, establishing these links in school-aged children has proven more elusive. There is evidence that school-aged children (8–12 years-old) do benefit from multisensory compared to unisensory learning contexts, with facilitated later (unisensory) recognition memory[21,22]. Similar conclusions are garnered from the works of Broadbent and colleagues. These authors found incidental learning to be improved by multisensory cues[23], and that retention of category learning over a 24-hour delay to be significantly higher for multisensory cues than unisensory ones in 5–10 year-old schoolchildren[24]. This is consistent with literature in adults reporting evidence for links between processes subserving multisensory integration on the one hand and cognitive functions, including recognition memory, on the other hand. For example, Thelen *et al.*[25] showed that individual performance on a continuous object recognition task could be predicted by brain responses to multisensory, but not unisensory, stimuli at initial encounters. Likewise, healthy elderly and those with mild cognitive impairment can be classified based on performance on a simple multisensory detection task, but not from unisensory performance alone, highlighting functional links between multisensory processes and memory (dys)functions[26].

While there is evidence that children (and, later, adults) indeed garner benefits from multisensory contexts when performing memory tasks, the links between benefits of multisensory information during stimulus processing and measures of intelligence remain to be firmly established. For example, one study of 95 school children aged 6–11 years old[27] compared performance on an auditory-visual simple detection task with scores from Raven's Coloured Progressive Matrices[28] and the Neal Analysis of Reading Ability[29]. In this work, there was no evidence of a statistically reliable link between multisensory facilitation of behaviour and these measures of cognitive function. Instead, those results provide evidence that multisensory processes, at least those indexed by violations of Miller's race model inequality, remain immature in this age group. In a later study of 88 school children, Barutchu and colleagues observed a significant difference in full-scale IQ between those children whose facilitation of reaction times exceeded probability summation and those children whose multisensory facilitation could be explained by probability summation[30]. An additional more recent study of 38 8–11 years-old children reported no correlation between (absolute) multisensory reaction time facilitation and IQ scores. Instead, there was a significant positive correlation between raw multisensory reaction time and their working memory index[31]. It should be noted, though, that there was no evidence for a systematic correlation between measures of multisensory facilitation and IQ scores (In fact, there were positive correlations between IQ and unisensory RTs[27,32]; a pattern somewhat at odds with the notion of IQ being coupled with processing speed or with facilitation under multisensory conditions).

That multisensory processing capabilities are related in some manner or another to cognitive ones is certain, as is the evidence that this relationship develops (and perhaps modulates in its nature) over childhood and adolescence. As this relationship could potentially offer a long-term scaffold to improve a child's scholastic outcomes, both in the case of typical development as well as in cases of neurodevelopmental disorders[15,33,34], clarification seems important. Our prior work in adults would indeed suggest that the manner in which an individual detects multisensory stimuli in their environment is predictive of how well multisensory contexts will be beneficial for recognition memory functions[25,35–41]. One implication is that low-level multisensory processes may be predictive of higher-level cognitive functions, be it multisensory or more traditional and unisensory, and that such relationships may be formed during childhood (and perhaps earlier). To better understand the nature of these interactions, here, we collected data from both a simple detection task and a continuous recognition memory task, which we have used extensively in our research in adults[25,35–41], together with standardised neuropsychological measures of working memory and fluid intelligence in school-age children.

Materials and Methods

Participants. In total, seventy-seven children (36 girls) from 4.6 to 15.5 years old ($M_{age} = 8.1$ years, $SD = 3.0$ years) partook in the experiment. All children had normal or corrected vision and reported no hearing loss. Moreover, Swiss children are all screened at age of 4 for sensory and learning disabilities. Any child with a reported suspicion of such disabilities was excluded from participating in our study. These individuals are the subset of participants from another study comparing pedagogical settings, and so information about schooling was also collected (Montessori and traditional). Nine schoolchildren were excluded from the study due to poor performance on the detection task ($N = 3$), defined as an accuracy rate lower than 30%, or due to missing data from technical issues ($N = 6$). The final sample included 68 children (32 girls), aged 4.6–15.5 years ($M_{age} = 7.9$ years, Median = 6.4 years, SD = 3.0 years). The study was conducted in accordance with the Declaration of Helsinki, and all parents provided written informed consent for their child to participate. The experimental procedures were approved by the Vaudois Cantonal Ethics Committee (Commission cantonale d'éthique de la recherche sur l'être humain).

Tasks and procedure. All experiments took place within Swiss French-speaking schools, and a separate room was set up for testing of individual children. Two different examiners collected the data, and task order was randomized. For computerized tasks, children were seated in front of a 20"-screen laptop. The auditory stimuli were presented over headphones (model: CASIO LK-260), and the volume was adjusted to a comfortable level (~60 dB, as measured with the Decibel meter from the laptop)[42]. Both tasks were presented and controlled electronically using the E-Prime 2.0 Professional software (Psychology Software Tools, Pittsburgh, PA), and the behavioural data were recorded through the laptop's keyboard.

Simple detection task. Children were presented with either visual (V), auditory (A) or audiovisual (AV) stimuli. Each child was presented with a total of 60 trials with a pseudo-randomised presentation, and equal distribution of the V, A, and AV conditions (i.e. 20 per condition). The visual stimuli were white drawings (cloud or star) presented on a black background, and the auditory stimuli were two different tones (44100 Hz digitisation; 16 bit stereo) that differed in their spectral composition to create two "opposite" types of sounds (the first one ranged from 20 Hz to 21000 Hz and the second one - from 18700 Hz to 19600 Hz). Stimuli were intermixed within blocks to maintain a high level of attention and unpredictability (in terms of which specific sensory modality would be stimulated). The audiovisual (AV) stimuli were the simultaneous and synchronous presentation of a visual and auditory stimulus. This type of detection paradigm is highly similar to that used by Fort and colleagues[43] in their seminal work in adults. Stimulus duration was 500 ms and was followed by a randomised inter-stimulus interval (ISI) ranging from 1500 to 1900 ms, during which time a central, white fixation cross was presented. Children were asked to press a button (the keyboard spacebar) as fast as possible when they perceived any type of stimulus. Both accuracy and reaction time were recorded.

Continuous recognition task. Children performed a continuous recognition task, adapted from Thelen *et al.*[25] The task was a 2-alterative forced choice that required the discrimination of initial (i.e., 'first') from repeated (i.e., 'second') instances of line drawings of common objects presented in a series of trials within a block (i.e., an "old/new" task) by pressing one of two buttons. The visual objects were black drawings presented centrally on a white background. The sounds were also selected from Thelen *et al.* (16 bit stereo; 44100 Hz digitization; 10 ms rise/fall to avoid clicks, they differed in their spectral composition, ranging from 100 Hz to 4700 Hz, and sometimes were modulated in terms of amplitude envelopes and/or waveform types). Trials were pseudo-randomised within a block of 60 trials (30 different drawings). On each trial a single image (selected from the original study) was presented alone (V) or with a congruent (AVc) or meaningless (AVm) sound (equal distribution of the three conditions; 10 trials per condition). Images were controlled to equate spatial frequency spectra and luminance between image groups (AV vs. V), according to the original task. Stimuli were presented for 500 ms, followed by a randomised inter-stimulus interval (IS) ranging from 900 to 1500 ms, where a fixation cross was shown. The mean number of trials between the initial and the repeated presentation was 5 ± 1 pictures for both V and AV conditions. Altogether, children performed four different blocks with new drawings each time (only two presentations of each drawing over all the experiment). The second presentation being always unisensory (V). Emphasis was put on both speed and accuracy. Supplementary Figure 1 illustrates the paradigm. Stimulus timing and synchrony across sensory modalities for both the simple detection task and the continuous recognition task were tested and verified using the EEG system in our laboratory as an "oscilloscope". Visual signals were converted to voltage with a photodiode, and auditory signals were directly taken from the output of the sound card. Simultaneous stimulus presentation has been reported to be perceived as synchronous both by adults and children (e.g.[44]).

Working memory. Children performed the Ascending Span task from the WISC-IV[45] to investigate the relationship between elementary multisensory processes and more complex cognitive abilities such as working memory[46]. The child was asked to listen and memorise a string of numbers spoken out loud by the experimenter and to repeat the string in an ascending order. The assessment started with a two digits string, and if the child successfully performed two trials in a row, an additional digit was added to the string. If the child missed a trial, a digit was removed from the string. If they missed three trials in a row the evaluation stopped. A final score was computed for the ascending digit task, based on the maximal number of correctly memorized and properly re-ordered digits, with a maximum of 7. No time limit was set for the answer; only accuracy was emphasized. These scores were then age-standardised based on mean span per year of age based on ref. [47].

Fluid intelligence (g factor). Children performed the black and white version[48] of Raven's Coloured Progressive Matrices[28] to assess abstract reasoning and non-verbal intelligence. It is a multiple-choice test composed of 36 items. For each item, an incomplete matrix was presented, and the child was asked to identify the missing element to complete the matrix. Participants had 15 minutes to complete as many matrices as possible. This test was conducted collectively (per small groups of maximally 5 children). Raw scores were based on the number of correct items (max. 36). The raw scores were then age-standardised using the calibration scale based on a sample of 1064 French schoolchildren following a traditional pedagogy (ECPA Pearson)[49].

Analysis design. As mentioned above, participants who missed more than 30% of the trials at the Simple Detection Task (3 children; mean age = 6.53, SD = 2.15), or with missing data due to technical issues (6 children; mean age = 10.28, SD = 1.10) were excluded from the analyses. Computerized data were pre-processed using

Excel; correct trials with a valid RT (*subject' smean RT* ± 3*SD*) were considered in analyses. Statistical analyses were run with Jamovi open-access software (retrieved from https://www.jamovi.org) as well as SPSS version 26 (IBM Corporation). Statistical significance criterion was set at $p \leq 0.05$. For all tests, the effect size is reported (either partial eta squared or Cohen's d). The full correlation matrix of the measures used in this study are provided in Supplemental Table 1.

First, to confirm multisensory benefits on a simple detection task, a repeated-measures analysis of covariance variance (ANCOVA) on mean RTs was performed with the within-subjects factor Condition (A, V, AV), and Age as the co-variate. We also performed this ANCOVA on detection rates. We also ran a repeated-measures ANCOVA on the accuracy rate [%] with the repetition conditions only from the continuous recognition task. The within subjects factor was Condition (V−, V + c, V + m) and Age was the co-variate. While previous results in adults has repeatedly indicated that RTs are not significantly modulated across conditions in this task[50], we nonetheless also analysed RTs from the continuous recognition task in a similar ANCOVA design as described above.

Second, in order to investigate how low-level multisensory gain (simple detection task) was related to high-level (continuous recognition task) multisensory gain as well as to both working memory and fluid intelligence scores, a *relative multisensory gain* was derived from the detection task for each subject as:

$$\Delta RT[\%] = \frac{\textit{faster unisensory Mean RT} - \textit{multisensory Mean RT}}{\textit{faster unisensory Mean RT}} \times 100$$

In addition, a *relative multisensory memory gain* was computed from the continuous recognition task for congruent AV recall condition as:

$$\Delta Accuracy[\%] = (\%Accuracy\ V + c) - (\%Accuracy\ V-)$$

In this study, we specifically addressed the relationship between low-level multisensory processes and higher-order cognitive abilities. First, the relative *multisensory gain* value of each subject was related to the *relative multisensory memory gain* from the continuous recognition task using a stepwise linear regression with the relative multisensory memory gain as the dependent variable and relative multisensory gain and age as the independent variables. Next, we related the relative multisensory gain and age-standardised working memory scores using a logistic regression model (given the fact that the working memory scores are discrete rather than continuous). Finally, we related the relative multisensory gain with age-standardised fluid intelligence scores using a stepwise linear regression with the fluid intelligence scores as the dependent variable and relative multisensory gain and age as the independent variables. For completion and despite our specific research questions regarding the relationship of relative multisensory gain to various global cognition measures, we also include a complete correlation table across all the measures in this study.

In addition, to control for the specificity of multisensory versus unisensory processes, we also computed the *relative unisensory gain* from the detection task, as:

$$\Delta RT[\%] = \frac{\textit{slower unisensory Mean RT} - \textit{faster unisensory Mean RT}}{\textit{slower unisensory Mean RT}} \times 100$$

We identified the slower and the faster sensory modality for each participant, separately. In 65 of the children, the visual modality was faster. In the remaining 3 children, the auditory modality was faster. This measure of relative unisensory gain was then related to (i) the *relative multisensory memory gain* from the continuous recognition task, (ii) age-standardised working memory scores, and age-standardised fluid intelligence scores in an analogous manner to what is described above.

Results

Simple detection task. The children performed the simple detection task with near-ceiling performance. Mean detection rates were 93.1%, 95.1%, and 96.5% for the visual, auditory, and multisensory conditions, respectively. These data were submitted to a one-way repeated-measures ANCOVA, with Condition as the within-subjects factor and Age as the co-variate (Greenhouse-Geisser corrected degrees of freedom are reported in cases of violation of assumptions of sphericity). There was a main effect of Condition ($F_{(1.813,119.639)} = 4.769$, $p = 0.012$, $\eta_p^2 = 0.07$) and a general increase in accuracy with age (i.e., significant covariation; $F_{(1,66)} = 14.11$, $p < 0.001$, $\eta_p^2 = 0.18$). However, this co-variation did not reliably differ across conditions ($F_{(1.813,119.639)} = 2.22$, $p = 0.12$, $\eta_p^2 = 0.03$). Detection rates for visual stimuli were significantly lower than those for multisensory stimuli ($p_{bonferroni} = 0.005$, $d = 0.36$). No other contrasts were statistically significant (p's > 0.17). Thus, and despite RTs being overall slower for A than V conditions (see below), there was no evidence that this slowing was matched by impaired detection rates.

Mean RTs were computed for each condition (AV, V, A) and subject (see Table 1 for group averages). Results of the one-way repeated-measures ANCOVA, with Condition as the within-subjects factor and Age as the co-variate, yielded a main effect of Condition ($F_{(2,132)} = 23.53$, $p < 0.001$, $\eta_p^2 = 0.26$) and a general decrease of RT with age (i.e., significant covariation; $F_{(1,66)} = 40.50$, $p < 0.001$, $\eta_p^2 = 0.38$). However, this co-variation did not reliably differ across conditions ($F_{(2,132)} = 2.27$, $p = 0.11$, $\eta_p^2 = 0.00$) (Fig. 1A). Post-hoc paired t-tests with a false-rate discovery (FDR) p-value correction at $q = 0.05$, showed participants had faster RTs on trials with AV stimuli than those with A stimuli ($t_{(67)} = 12.95$, $p_{FDR} = 0.002$, Cohen's $d = 1.57$) as well as those with V stimuli ($t_{(67)} = 2.14$, $p_{FDR} = 0.036$; Cohen's $d = 0.26$), and faster RTs for V than A condition ($t_{(67)} = 11.58$, $p_{FDR} = 0.002$, Cohen's $d = 1.40$).

Across participants, the average relative *multisensory gain* was 3.19%, SD = 10.2%. The average absolute multisensory gain in milliseconds was 17.78 ms, SD = 75.95 ms. These metrics were highly positively correlated, even when controlling for age (partial $r_{(65)} = 0.975$; $p < 0.001$). Across participants, the average relative unisensory gain

Tasks		Mean	SD
Detection RT [ms]	*A*	748	187
	V	635	146
	AV	615	173
Multisensory gain [%]		3.19	10.17
Unisensory gain [%]		14.98	6.98
Continuous recognition Accuracy [%]	*V−*	68.8	19.5
	V+c	70.2	18.2
	V+m	69.4	17.3
Relative multisensory memory gain [%]		1.40	13.0
Age-standardised Working Memory [%]		61	25
Age-standardised Fluid Intelligence [scale]		5.42	2.81

Table 1. Tasks' mean scores and standard deviations.

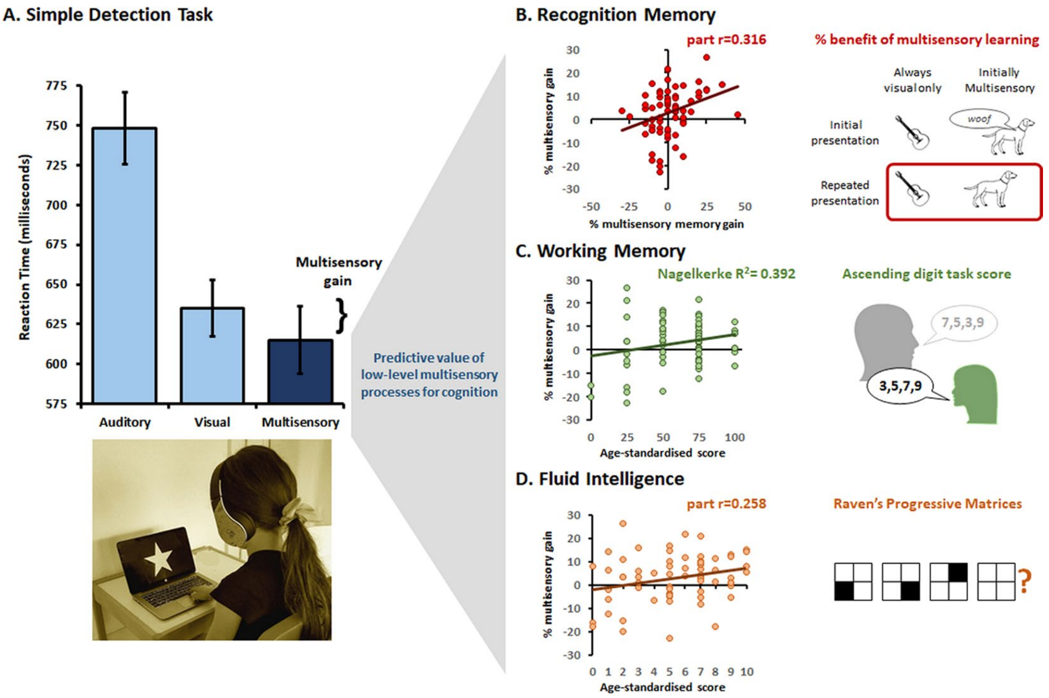


Figure 1. Multisensory gains in simple detection predict memory as well as fluid intelligence in schoolchildren. (**A**) Simple detection task; children were asked to press a button as fast as possible whenever a stimulus (auditory, visual or auditory-visual multisensory) was perceived. On average, reaction times were significantly faster for multisensory than for either auditory or visual stimuli ($p < 0.001$ and $p < 0.035$, respectively). For each child, a measure of multisensory gain was derived, equal to the relative difference in mean reaction time between the multisensory and the better unisensory condition. This percentage of multisensory gain (plotted on the y-axis in panels B–D) was linearly related to several measures of cognitive functioning, including recognition memory on a continuous old/new recognition task (**B**), working memory as assessed with the ascending digit task (**C**), and fluid intelligence as measured with Raven's Progressive Matrices (**D**). The images in panel B are from the Snodgrass and Vanderwart (1980) database[85].

was 14.98%, SD = 6.98%. The relative multisensory and unisensory gains (in percentages) were negatively correlated, when controlling for age (partial $r_{(65)} = -0.343$; $p = 0.004$).

It is important to mention that our paradigm, which entailed 2 visual stimuli, 2 auditory stimuli and their 4 multisensory combinations. It could be argued that one of the visual stimuli or auditory stimuli was more challenging to process, despite the task requirement of simple detection and the high performance rates of the participants. To assess this possibility, we compared mean RTs for the 2 visual stimuli, and there was no significant difference (635 vs. 626 ms; $p = 0.43$). We also compared mean RTs for the 2 auditory stimuli, and there was no significant difference (735 vs. 753 ms; $p = 0.20$). It could also be argued that participants established an implicit association between a given visual and auditory stimulus; a notion referred to as crossmodal correspondence[51]. While the fact that all multisensory combinations were equally probable provides one level of argument against this possibility, we also assessed this empirically by comparing mean RTs from what could arguably be labelled as the congruent and incongruent combinations[51]. There was no significant difference (548 vs. 562 ms; $p = 0.24$).

Continuous recognition task. Accuracy rates [%] were computed for each repetition condition per subject; initially visual [V−] (mean = 68.8%, SD = 19.5%), initially paired with a meaningless sound [V + m] (mean = 69.4%, SD = 17.3%), and initially paired with a semantically congruent sound [V + c] (mean = 70.2%, SD = 17.3%). A repeated-measures ANCOVA, with Condition as the within-subjects factor and Age as the co-variate, yielded a significant covariation between accuracy and age ($F_{(1,132)} = 23.0$, $p < 0.001$, $\eta_p^2 = 0.26$). Neither the main effect of Condition ($F_{(2,132)} = 1.39$, $p = 0.25$, $\eta_p^2 = 0.20$), nor the interaction term of age co-varying differently across Condition ($F_{(2,132)} = 2.03$, $p = 0.14$, $\eta_p^2 = 0.03$) were reliable. Across participants, the average *relative multisensory memory gain* was 1.40%, SD = 13.0%, range −30% to 45%. The ANCOVA using RTs as the dependent measure did not yield a reliable main effect of Condition ($F_{(2,132)} < 1$) or any reliable covariation with age ($F_{(2,132)} < 1$).

Predictive value of gains in simple detection for memory and global cognitive functions. We first conducted a stepwise linear regression, using the relative multisensory memory gain as the dependent, outcome variable and relative multisensory gain on the detection task as well as age as independent variables. The regression model was statistically significant ($R = 0.316$; $F_{(1,66)} = 7.296$, $p = 0.009$). Only the relative multisensory gain on the detection task was identified as a significant predictor of relative multisensory memory gain, accounting for 10% of the unique variance (part $r = 0.316$). Age did not significantly increase the performance of the model, $p = 0.456$). Figure 1B shows a scatterplot relating the relative multisensory gain on the detection task with that on the continuous recognition memory task.

Next, we conducted a multinomial logistic regression, using the age-standardised working memory scores as the dependent, outcome variable and relative multisensory gain on the detection task as well as age as covariates. Addition of the relative multisensory gain on the detection task and age to a model that contained only the intercept significantly improved the fit between the model and data, $\chi^2(8, N = 68) = 30.90$, Nagelkerke $R^2 = 0.392$, $p < 0.001$. Significant unique contributions were made by both the relative multisensory gain on the detection task [$\chi^2(4, N = 68) = 11.381$; $p = 0.023$] and age [$\chi^2(4, N = 68) = 19.724$; $p = 0.001$]. Goodness of fit was explored by using the Pearson chi-square statistic, which was not statistically significant ($p = 0.98$). Figure 1C shows a scatterplot relating the relative multisensory gain on the detection task with age-standardised working memory scores.

Finally, we conducted a stepwise linear regression, using the age-standardised fluid intelligence scores as the dependent, outcome variable and relative multisensory gain on the detection task as well as age as independent variables. This regression model was statistically significant ($R = 0.258$; $F_{(1,66)} = 4.718$, $p = 0.033$). Only the relative multisensory gain on the detection task was identified as a significant predictor of the age-standardised fluid intelligence scores, accounting for 6.7% of the unique variance (part $r = 0.258$). Age did not significantly increase the performance of the model, $p = 0.392$). Figure 1D shows a scatterplot relating the relative multisensory gain on the detection task with age-standardised fluid intelligence scores scores.

To assess the specificity of the relative multisensory gain on the detection task as a predictor of global cognitive functions, we performed the abovementioned regressions with the relative unisensory gain on the detection task. In the case of relative multisensory memory gain, the model including age and unisensory gain as predictors did not result in a significant improvement ($R = 0.175$; $F_{(2,65)} = 1.31$; $p = 0.362$). In the case of age-standardised working memory scores, addition of the unisensory gain and age to a model that contained only the intercept significantly improved the fit between the model and data, $\chi^2(8, N = 68) = 20.537$, Nagelkerke $R^2 = 0.280$, $p = 0.008$. Significant unique contributions were made only by age [$\chi^2(4, N = 68) = 18.966$; $p = 0.001$], but not by the unisensory gain [$\chi^2(4, N = 68) = 1.016$; $p = 0.907$]. Goodness of fit was explored by using the Pearson chi-square statistic, which was not significant ($p = 0.94$). In the case of age-standardised fluid intelligence scores, the model including age and unisensory gain as predictors did not result in a significant improvement ($R = 0.221$; $F_{(2,65)} = 1.67$; $p = 0.196$).

Discussion

In this study, we investigated the relationship between multisensory gain in a simple detection task and global cognitive measures such as memory, working memory and fluid intelligence. Our principal finding is the statistically significant and selective link between low-level multisensory processes and multiple measures of higher-order cognitive performance in schoolchildren. These links were observed not only with laboratory-based tasks, for which the contribution of age was controlled, but also with age-standardized clinical evaluation tools that index working memory and fluid intelligence. Such links did not generalize to unisensory processes, suggestive of a certain degree of specificity of the studied constructs. These collective findings reinforce the hypothesis that multisensory perceptual processes provide a crucial scaffolding for cognition throughout the lifespan[1].

A long history of research has reported links between unisensory as well as multisensory stimulus processing and measures of cognition and intelligence in infants and pre-school children[5,10,16,17]. However, the majority of these studies involved tasks that required matching information (e.g. shape or temporal pattern) across the senses rather than simple detection of the stimuli[19,20], obfuscating the ability to claim that it is specifically the low-level stimulus processing mediating such links, rather than a common higher-level cognitive function contributing to both tasks. In fact, we do know that children and adults allocate attention differently to unisensory and multisensory stimuli[52,53]. In a series of studies in schoolchildren, Barutchu and colleagues did not observe a linear correlation between indices of low-level multisensory processing and intelligence scores. Specifically, in their 2009 study, they found no evidence for associations between simple reaction times to multisensory stimuli (or absolute measures of multisensory facilitation as derived from mean RTs between multisensory condition and the better unisensory condition) and non-verbal IQ (as measured with Raven's progressive matrices) or reading abilities (as measured with the Neal Analysis of Reading Ability). Likewise, no correlation was observed in a subsequent 2011 paper that used the WISC-IV as a measure of IQ. The 2009 study focused on age-related differences in the extent to which simple reaction times were facilitated under multisensory conditions beyond what could be explained by probability summation, using Miller's race model inequality[54]. The 2011 study revealed IQ differences between sub-groups of children according to whether or not the child's multisensory facilitation of RTs exceeded predictions of probability summation. The present results and those of Barutchu *et al*. are consistent to the extent that they both indicate that the degree to which a child benefits from multisensory stimuli is related to their global cognition. Here, in our view, age-related differences in violations of Miller's race model inequality, even if repeatedly reported, ought to be considered with some caution. For one, there are examples in adults where such violation have not been systematically observed[55], raising the possibility that this metric is not fully indicative of the maturity of multisensory processes. Second, biases in the use of Miller's inequality can be observed when numbers of trials are low[56,57], which is often the case in studies of children. Third, non-linear neural response interactions can be decoupled from violations in Miller's inequality[58,59]. Additional research is clearly required to fully ascertain which multisensory or integrative processes are indexed by violations of Miller's inequality. Such notwithstanding, evidence from multiple laboratories would indeed indicate that some forms of multisensory integration remain immature in children as old as 10–11 years old. Our findings thus provide an important extension beyond what has been previously reported by Barutchu and colleagues. By showing a correlation between the percentage of multisensory facilitation of simple reaction-time and age-standardised measures of IQ, we provide evidence that multisensory processes are sufficiently mature already in school-aged children to be informative of their global cognitive abilities. Our findings also extend the studies showing that cross-modal matching is linked with cognitive functions[19,20]. We show that multisensory gains during simple detection, which arguably rely on more rudimentary processes than those involved in temporal matching, can reliably predict several measures of global cognitive functions.

It is also worth mentioning that our data are consistent with a rich literature characterising links between unisensory processing speed (as measured on either simple or choice reaction time tasks) and intelligence measures (e.g.[11,60]; reviewed in[61]). This can be gleaned from the correlation matrix (Supplementary Table 1), which generally shows a negative correlation between unisensory reaction times and working memory as well as fluid intelligence scores, though not with performance on the continuous recognition task. However, multiple regression analyses that included both RTs and age did not lead to a consistent pattern of results. While unisensory RTs reliably predicted age-standardised fluid intelligence scores, they did not predict either age-standardised working memory scores or relative multisensory memory gains on the continuous recognition task (Supplementary Table 2). Such notwithstanding, the claim in these prior studies is that simple reaction times, in general, reflect processing speed of basic cognitive operations. One prominent hypothesis focuses on the notion of neural efficiency (reviewed in[62]). More efficient processing, as in the case of individuals with higher intelligence or cognitive abilities, is paralleled by faster reaction times. It has also be shown that multisensory conditions result in less variable reaction times[63], which may be a further contributing factor as to why multisensory processes may be particularly informative of cognitive functions. Here, our findings only reinforce the idea that low-level stimulus processing – and specifically the ability to garner benefits from multisensory contexts during low-level stimulus processing – are tightly related to higher-level cognitive processes (i.e., memory) and intellectual abilities in schoolchildren. Nonetheless, additional research will be required to not only determine to what extent multisensory processes are innate and/or experience-dependent (see[10] for discussion), but also to what extent genetic factors contribute to multisensory stimulus processing, particularly given some evidence for genetic contributions to speed of information processing and its link to IQ (e.g.[60]). Regarding the former aspect, an ongoing clinical trial by our group is investigating multisensory processes in prematurely born infants and children as well as their predictive value for cognitive functions and scholastic achievement[64]. It is likewise important to consider the extent to which our findings are indicative of links between multisensory processes and a common (and perhaps general) cognitive construct or multiple such constructs. One access point to this issue is the pattern of relationships across the continuous recognition task, working memory task, and fluid intelligence. There was no reliable association between the relative multisensory memory gain from the continuous recognition task and age-standardised working memory scores ($\eta = 0.291$; $p = 0.225$) nor a correlation between such and age-standardised fluid intelligence scores ($r_{(66)} = -0.027$; $p = 0.828$). By contrast, and unsurprisingly (see[65]), age-standardised working memory and fluid intelligence measures were reliably associated ($\eta = 0.482$; $p < 0.001$) (see Supplementary Table 1). This overall pattern would suggest that the measures of working memory and fluid intelligence may be indexing a common cognitive construct. By contrast, the continuous recognition task is likely gauging a distinct construct. As such, it would thus appear that a child's ability to garner multisensory benefits on a simple detection task provides an indicator of the integrity of at least two distinct cognitive systems.

There is a particularly straightforward and exciting implication of our findings; it suggests that multisensory learning, which is arguably more reflective of the sensory environment a child confronts and acts upon from birth onwards, could potentially empower cognitive development[66]. Linked to the above idea is the question of the extent to which multisensory processing abilities can be trained. This is a burgeoning field of empiric research. On the one hand, there are data showing that the so-called temporal binding window over which multisensory signals are perceptually bound is flexible and subject to learning[67,68]. This is important as the temporal binding window has been reported to be altered in a number of neurodevelopmental disorders[69], as well as in aging[70], and also to scale across tasks from simple detection to speech processing[71]. On the other hand, there are data showing that multisensory contexts are particularly effective for recognition memory not only in adults (reviewed in[41]), but also in schoolchildren[21–24]. That is, if multisensory processes can be trained[67], preschool and school years may be ideal to facilitate early learning and basic perceptual skills through multisensory programs (discussed in[72]). This was already acknowledged by some educational approaches, such as Montessori Education where children work mainly through the manipulation of sensory materials[73]. Interestingly, the few quantitative existing studies in this area do report scholastic advantages for schoolchildren following a Montessori versus traditional system[74–76]. An additional, exploratory analysis of the current dataset was thus run on the children according to their schooling background and their multisensory gain on the detection task. The results revealed that Montessori scholars (half of the participants) were more likely to exhibit a multisensory gain than their peers in traditional schooling ($\chi^2(1) = 5.7$, $p = 0.02$) (Fig. S2). These preliminary results call for further investigation of the topic of pedagogical tools, but already support increased efforts in multisensory enrichment targeting learning and memory processes. More generally, our results are in line with the cognitive cascade model as proposed by Rose *et al.*[16] that focuses on the relationship between low-level multisensory processes and higher-order cognitive skills. While our sample size was modest, there was the added value of readily controlling for many demographic factors in a setting such as Switzerland. However, replication and multi-cultural studies will be required to establish the potential utility of multisensory tasks as a screening tool and multisensory enrichment as a learning aide.

It is important to mention some limitations of the present study. First, our study included a wider range of ages than other comparable studies[27,30,31]. The fact that we included younger children may be one contributing factor to the smaller average relative multisensory gain on the detection task that we observed here versus that observed in works by Barutchu and colleagues. When we considered an age range restricted to the 6–11 year-old range in Barutchu *et al.*[27], the average relative multisensory gain on the detection task was 5%. That said, it is perhaps important to note that relative multisensory gains in studies of adults exhibit considerable inter-individual as well as between-study variability[55,77–80]. More importantly, our results provide no evidence that age was significantly contributing to any of the models using relative multisensory gain on the detection task as a predictor. Second, our study made no effort to calibrate the stimuli used in the simple detection task; RTs to auditory stimuli were slower than those to visual stimuli. Other similar work in children has used stimuli that resulted in equivalent mean RTs to both unisensory conditions[27,81]. Larger multisensory gains are obtained when the distributions to the unisensory conditions are closer to each other[63,82]. Interestingly and consistently with our prior work in adults[26], we observed a strong negative correlation between relative multisensory gains and relative unisensory gains. While there was no evidence here that relative unisensory gains were reliable predictors of cognitive abilities, it may be that a combined metric of relative multisensory and unisensory processes may prove particularly effective should a multisensory detection task be used as a screening tool for neurodevelopmental disorders (cf.[26] for a similar tactic in the case of mild cognitive impairment). Third, our study used a detection task that was somewhat different from that used in prior works. While prior studies used a single visual stimulus, a single auditory stimulus and their multisensory combination, the present study used 2 visual stimuli, 2 auditory stimuli, and all 4 multisensory combinations thereof. Prior studies in adults have used a detection task with multiple stimuli and did not observe differences between specific items[43]. Similarly, we found no such differences here nor any evidence for implicit crossmodal correspondences (at least with the stimulus set we used). Nonetheless, it would be informative for future research to determine what might constitute an optimal detection task design both in terms of predictive value for cognitive (dys)function and in terms of ease-of-use in schoolchildren, but also in preschoolers and infants.

The present findings of reliable links between multisensory processes and higher-level cognition cannot directly speak to their causality. Nonetheless, our results would indeed suggest that low-level multisensory processes may constitute an effective access point for the assessment of children and their cognitive development. They reinforce the possible applicability of multisensory processes to public health screening in schoolchildren. In fact, our group has already demonstrated such in the case of screening for mild cognitive impairment in the elderly based on a similar multisensory simple detection task[26]. In that study, a combined measure of sensory dominance and multisensory gain on performance reliably classified healthy elderly from those with mild cognitive impairment at level comparable with a standard clinical tool (i.e. the Hopkins Verbal Learning Task). It would be particularly promising to apply the present results in screening of (pre)school children, particularly given that multisensory processing has been shown to be selectively impaired in dyslexia (e.g.[33,83]) as well as autism (e.g.[84], reviewed in[10]). Moreover, the detection task *per* se circumvents some of the major limitations of current screening batteries (e.g. parental report, socio-economic bias, requirement of literacy/numeracy skills). Combined with a prompt administration time, a simple detection task makes an attractive potential screening tool for pre-schoolers or pre-linguistic children.

Received: 10 May 2019; Accepted: 14 January 2020;
Published online: 04 February 2020

References

1. Murray, M. M., Lewkowicz, D. J., Amedi, A. & Wallace, M. T. Multisensory Processes: A Balancing Act across the Lifespan. *Trends Neurosci.* **39**, 567–579 (2016).
2. Amso, D. & Scerif, G. The attentive brain: insights from developmental cognitive neuroscience. *Nat. Rev.* **16**, 606–619 (2015).
3. Stein, B. E. & Meredith, M. A. *The Merging of the Senses*. (MIT Press, 1993).
4. Murray, M. M. & Wallace, M. T. *The Neural Bases of Multisensory Processes*. (CRC Press, 2012).
5. Bahrick, L. E. & Lickliter, R. Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Dev. Psychol.* **36**, 190–201 (2000).
6. Bahrick, L. E. & Lickliter, R. The role of intersensory redundancy in early perceptual, cognitive, and social development. In *Multisensory development* (eds. Bremner, A. J., Lewkowicz, D. J. & Spence, C.) 183–206 (Oxford University Press, 2012).

7. Lewkowicz, D. J. & King, A. J. The developmental and evolutionary emergence of multisensory processing: From single cells to behavior. In *The New Handbook of* Multisensory *Processing* (ed. Stein, B. E.) (MIT Press, 2012).
8. Lewkowicz, D. J. Early experience and multisensory perceptual narrowing. *Dev. Psychobiol.* **56**, 292–315 (2014).
9. Matusz, P. J., Dikker, S., Huth, A. G. & Perrodin, C. Are We Ready for Real-world Neuroscience? *J. Cogn. Neurosci*. 1–13. https://doi.org/10.1162/jocn_e_01276 (2018).
10. Dionne-Dostie, E., Paquette, N., Lassonde, M. & Gallagher, A. Multisensory Integration and Child Neurodevelopment. *Brain Sci.* **5**, 32–57 (2015).
11. Baker, L. A., Vernon, P. A. & Ho, H. Z. The genetic correlation between intelligence and speed of information processing. *Behav. Genet.* **21**, 351–367 (1991).
12. Vernon, P. A., Nador, S. & Kantor, L. Reaction times and speed-of-processing: Their relationship to timed and untimed measures of intelligence. *Intelligence* **9**, 357–374 (1985).
13. Sheppard, L. D. & Vernon, P. A. Intelligence and speed of information-processing: A review of 50 years of research. *Pers. Individ. Dif.* **44**, 535–551 (2008).
14. Vernon, P. A. & Weese, S. E. Predicting intelligence with multiple speed of information-processing tests. *Pers. Individ. Dif.* **14**, 413–419 (1993).
15. Park, J., Mainela-Arnold, E. & Miller, C. A. Information processing speed as a predictor of IQ in children with and without specific language impairment in grades 3 and 8. *J. Commun. Disord.* **53**, 57–69 (2015).
16. Rose, S. A. *et al*. Cognitive Cascade in Infancy: Pathways from Prematurity to Later Mental Development. *Intelligence* **36**, 367–378 (2008).
17. Birch, H. G., Belmont, L., Ph, D., Belmont, L. & Ph, D. Auditory visual integration in normal and retarded readers. *Bull. Ort. Soc.* **15**, 48–96 (1965).
18. Birch, H. G. & Belmont, L. Auditory-Visual Integration, Intelligence and Reading Ability in School Children. *Percept. Mot. Skills* **20**, 295–305 (1965).
19. Birch, H. G. & Belmont, L. Auditory visual integration in normal and retarded readers. *Bull. Ort. Soc.* **15**, 48–96 (1965).
20. Rose, S. A., Feldman, J. F., Jankowski, J. J. & Futterweit, L. R. Visual and auditory temporal processing, cross-modal transfer, and reading. *J. Learn. Disabil.* **32**, 256–66.
21. Heikkilä, J. & Tiippana, K. School-aged children can benefit from audiovisual semantic congruency during memory encoding. *Exp. brain Res.* **62**, 123–130 (2015).
22. Heikkilä, J., Alho, K., Hyvönen, H. & Tiippana, K. Audiovisual semantic congruency during encoding enhances memory performance. *Exp. Psychol.* **62**, 123–30 (2015).
23. Broadbent, H. J., White, H., Mareschal, D. & Kirkham, N. Z. Incidental learning in a multisensory environment across childhood. *Dev. Sci*. **21**, (2018).
24. Broadbent, H. J., Osborne, T., Mareschal, D. & Kirkham, N. Z. Withstanding the test of time: Multisensory cues improve the delayed retention of incidental learning. *Dev. Sci.* 1–7. https://doi.org/10.1111/desc.12726 (2018).
25. Thelen, A., Matusz, P. J. P. J. & Murray, M. M. Multisensory context portends object memory. *Curr. Biol.* **24**, R734–R735 (2014).
26. Murray, M. M. *et al*. Sensory dominance and multisensory integration as screening tools in aging. *Sci. Rep.* **8**, 8901 (2018).
27. Barutchu, A., Crewther, D. P. & Crewther, S. G. The race that precedes coactivation: Development of multisensory facilitation in children. *Dev. Sci.* **12**, 464–473 (2009).
28. John & Raven, J. Raven Progressive Matrices. in *Handbook of Nonverbal Assessment* 223–237 (Springer US, 2003). https://doi.org/10.1007/978-1-4615-0153-4_11.
29. Neale, M. D., McKay, M. F. & Childs, G. H. The Neale Analysis of Reading Ability - Revised. *Br. J. Educ. Psychol.* **56**, 346–356 (1986).
30. Barutchu, A. *et al*. The relationship between multisensory integration and IQ in children. *Dev. Psychol.* **47**, 877–885 (2011).
31. Barutchu, A., Fifer, J. M., Shivdasani, M. N., Crewther, S. G. & Paolini, A. G. The Interplay Between Multisensory Associative Learning and IQ in Children. *Child Dev*., https://doi.org/10.1111/cdev.13210 (2019).
32. Barutchu, A., Sahu, A., Humphreys, G. W. & Spence, C. Multisensory processing in event-based prospective memory. *Acta Psychol. (Amst).* **192**, 23–30 (2019).
33. Harrar, V. *et al*. Multisensory integration and attention in developmental dyslexia. *Curr. Biol.* **24**, 531–535 (2014).
34. Wallace, M. T. & Stevenson, R. A. The construct of the multisensory temporal binding window and its Dysregulation in developmental Disabilities. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2014.08.005 (2014).
35. Murray, M. M. *et al*. Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage* **21**, 125–135 (2004).
36. Lehmann, S. & Murray, M. M. The role of multisensory memories in unisensory object discrimination. *Cogn. Brain Res*. **24**, (2005).
37. Murray, M. M., Foxe, J. J. & Wylie, G. R. The brain uses single-trial multisensory memories to discriminate without awareness. *Neuroimage* **27**, 473–8 (2005).
38. Thelen, A., Cappe, C. & Murray, M. M. Electrical neuroimaging of memory discrimination based on single-trial multisensory learning. *Neuroimage* **62**, 1478–1488 (2012).
39. Matusz, P. J. P. J. *et al*. The role of auditory cortices in the retrieval of single-trial auditory-visual object memories. *Eur. J. Neurosci.* **41**, 699–708 (2015).
40. Thelen, A., Talsma, D. & Murray, M. M. Single-trial multisensory memories affect later auditory and visual object discrimination. *Cognition* **138**, (2015).
41. Matusz, P. J., Wallace, M. T. & Murray, M. M. A multisensory perspective on object memory. *Neuropsychologia* **105**, (2017).
42. Huth, M. E., Popelka, G. R. & Blevins, N. H. Comprehensive measures of sound exposures in cinemas using smart phones. *Ear Hear*. **35**, 680–6.
43. Fort, A., Delpuech, C., Pernier, J. & Giard, M.-H. Dynamics of cortico-subcortical cross-modal operations involved in audio-visual object detection in humans. *Cereb. Cortex* **12**, 1031–9 (2002).
44. Hillock, A. R., Powers, A. R. & Wallace, M. T. Binding of sights and sounds: Age-related changes in multisensory temporal processing. *Neuropsychologia* **49**, 461–467 (2011).
45. Wechsler, D. Wechsler Intelligence Scale for Children and Adolescents. (2003).
46. Rose, S. A., Feldman, J. F. & Jankowski, J. J. The building blocks of cognition. *J. Pediatr.* **143**, S54–61 (2003).
47. Fournier, M. & Albaret, J.-M. Étalonnage des blocs de Corsi sur une population d'enfants scolarisés du CP à la 6e. *Développements* **16–17**, 76 (2013).
48. Tuddenham, R. D., Davis, L., Davison, L. & Schindler, R. An experimental group version for school children of the progressive matrices. *J. Consult. Psychol.* **22**, 30 (1958).
49. Raven, J., Raven, J. & Court, H. *Manual for Raven's Progressive Matrices*. (Harcourt Assessment, 2003).
50. Murray, M. M. & Thelen, A. The Efficacy of Single-Trial Multisensory Memories. *Multisens. Res.* **26**, 483–502 (2013).
51. Spence, C. & Deroy, O. How automatic are crossmodal correspondences? *Conscious. Cogn.* **22**, 245–260 (2013).
52. Matusz, P. J. *et al*. Multi-modal distraction: Insights from children's limited attention. *Cognition* **136**, 156–165 (2015).
53. Matusz, P. J., Merkley, R., Faure, M. & Scerif, G. Expert attention: Attentional allocation depends on the differential development of multisensory number representations. *Cognition* **186**, 171–177 (2019).
54. Miller, J. Divided attention: evidence for coactivation with redundant signals. *Cogn. Psychol.* **14**, 247–79 (1982).

55. Hughes, H. C., Reuter-Lorenz, P. A., Nozawa, G. & Fendrich, R. Visual-auditory interactions in sensorimotor processing: saccades versus manual responses. *J. Exp. Psychol. Hum. Percept. Perform.* **20**, 131–53 (1994).
56. Kiesel, A., Miller, J. & Ulrich, R. Systematic biases and Type i error accumulation in tests of the race model inequality. *Behav. Res. Methods* **39**, 539–551 (2007).
57. Gondan, M. & Minakata, K. A tutorial on testing the race model inequality. *Attention, Perception, Psychophys.* **78**, 723–735 (2016).
58. Sperdin, H. F., Cappe, C., Foxe, J. J. & Murray, M. M. Early, low-level auditory-somatosensory multisensory interactions impact reaction time speed. *Front. Integr. Neurosci.* **3**, 2 (2009).
59. Murray, M. M., Foxe, J. J., Higgins, B. A., Javitt, D. C. & Schroeder, C. E. Visuo-spatial neural response interactions in early cortical processing during a simple reaction time task: a high-density electrical mapping study. *Neuropsychologia* **39**, 828–44 (2001).
60. Rijsdijk, F. V., Vernon, P. A. & Boomsma, D. I. The genetic basis of the relation between speed-of-information- processing and IQ. *Behav. Brain Res.* **95**, 77–84 (1998).
61. Vernon, P. A. Speed of Information Processing and Intelligence. (Ablex Publishing, 1987).
62. Jensen, A. R. Reaction Time and Psychometric g. In *A Model for Intelligence* 93–132 (Springer Berlin Heidelberg, 1982). https://doi.org/10.1007/978-3-642-68664-1_4.
63. Murray, M. M., Thelen, A., Ionta, S. & Wallace, M. T. Contributions of intra- and inter- individual differences to multisensory processes. *J. Cogn. Neurosci.* (2018).
64. Neel, M. L. *et al.* Randomized controlled trial protocol to improve multisensory neural processing, language and motor outcomes in preterm infants. *BMC Pediatr.* **19**, 81 (2019).
65. Fry, A. F. & Hale, S. Processing speed, working memory, and fluid intelligence. *Psychol. Sci.* **7**, 237–241 (1996).
66. Shams, L. & Seitz, A. R. Benefits of multisensory learning. *Trends Cogn. Sci.* **12**, 411–7 (2008).
67. Powers, A. R., Hillock, A. R. & Wallace, M. T. Perceptual training narrows the temporal window of multisensory binding. *J. Neurosci.* **29**, 12265–74 (2009).
68. Powers, A. R., Hillock-Dunn, A. & Wallace, M. T. Generalization of multisensory perceptual learning. *Sci. Rep.* **6**, 1–9 (2016).
69. Stevenson, R. A. *et al.* Multisensory temporal integration in autism spectrum disorders. *J Neurosci* **34**, 691–697 (2014).
70. Poliakoff, E., Shore, D. I., Lowe, C. & Spence, C. Visuotactile temporal order judgments in ageing. *Neurosci. Lett.* **396**, 207–11 (2006).
71. Stevenson, R. A., Wallace, M. T. & Altieri, N. The interaction between stimulus factors and cognitive factors during multisensory integration of audiovisual speech. *Front. Psychol.* **5**, 352 (2014).
72. Murray, M. M., Matusz, P. J. & Amedi, A. Neuroplasticity: Unexpected Consequences of Early Blindness. *Curr. Biol.* **25**, (2015).
73. Montessori, M. *The Secret of Childhood*. (Ballantine Books, 1982).
74. Lillard, A. S. *et al.* Montessori preschool elevates and equalizes child outcomes: A longitudinal study. *Front. Psychol.* **8**, 1–19 (2017).
75. Lillard, A. THE EARLY YEARS: Evaluating Montessori Education. *Science (80-.).* **313**, 1893–1894 (2006).
76. Denervaud, S., Knebel, J.F., Hagmann, P. & Gentaz, E. Beyond executive functions, creativity skills benefit academic outcomes: Insights from Montessori education. *PLoS One* **14**, e0225319 (2019).
77. Hecht, D., Reiner, M. & Karni, A. Multisensory enhancement: gains in choice and in simple response times. *Exp. Brain Res.* **189**, 133–43 (2008).
78. Romei, V., Murray, M. M., Merabet, L. B. & Thut, G. Occipital transcranial magnetic stimulation has opposing effects on visual and auditory stimulus detection: implications for multisensory interactions. *J. Neurosci.* **27**, 11465–72 (2007).
79. Martuzzi, R. *et al.* Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cereb. Cortex* **17**, 1672–9 (2007).
80. Molholm, S. *et al.* Multisensory auditory-visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Brain Res. Cogn. Brain Res.* **14**, 115–28 (2002).
81. Brandwein, A. B. *et al.* The Development of Multisensory Integration in High-Functioning Autism: High-Density Electrical Mapping and Psychophysical Measures Reveal Impairments in the Processing of Audiovisual Inputs. *Cereb. Cortex*. https://doi.org/10.1093/cercor/bhs109 (2012).
82. Otto, T. U., Dassy, B. & Mamassian, P. Principles of multisensory behavior. *J. Neurosci.* **33**, 7463–74 (2013).
83. Hahn, N., Foxe, J. J. & Molholm, S. Impairments of multisensory integration and cross-sensory learning as pathways to dyslexia. *Neurosci. Biobehav. Rev.* **47**, 384–392 (2014).
84. Stevenson, R. A. *et al.* The cascading influence of multisensory processing on speech perception in autism. *Autism* **22**, 609–624 (2018).
85. Snodgrass, J. G. & Vanderwart, M. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol. Hum. Learn.* **6**, 174–215 (1980).

Acknowledgements

This study was supported by the Swiss National Sciences Foundation (Grants 320030-149982 and 320030-169206 to MMM, and grant PZ00P1_174150 to PJM) as well as by the Fondation Asile des aveugles (232933 to MMM), the Pierre Mercier Foundation (RSVS30349 to PJM) and a grantor advised by Carigest SA (232930 to MMM). We thank all the children and parents for their participation. We are likewise grateful to Professor David Sander for his kind support of S.D.

Author contributions

S.D. and M.M.M are responsible for the study concept and design. S.D. acquired the data. The analysis and interpretation of data were carried out by S.D., P.J.M. and M.M.M. The manuscript was drafted by S.D., P.J.M. and M.M.M. E.G. provided input on revisions to the manuscript. All authors approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-58329-4.

Correspondence and requests for materials should be addressed to M.M.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Multisensory gains in simple detection predict global cognition in schoolchildren

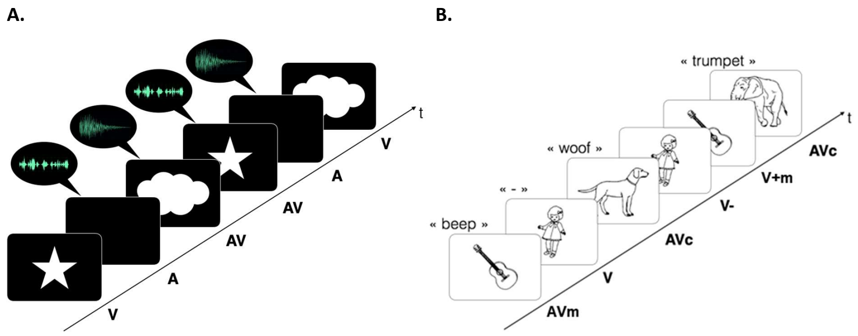
Solange Denervaud, Edouard Gentaz, Pawel J. Matusz*, Micah M. Murray*

	1. Age	2. Det-V	3. Det-A	4. Det-AV	5. Abs_AV	6. %_AV	7. %_uni	8. %Mem	9. WM	10. FI
1. Age(years)										
2. Detection task: Visual RT (ms)	-0.612**									
3. Detection task: Auditory RT (ms)	-0.590**	0.911**								
4. Detection task: Multisensory RT (ms)	-0.587**	0.897**	0.892**							
5. Detection task: Absolute Multisensory gain (ms)	0.186	-0.173	-0.302*	-0.573**						
6. Detection task: Relative multisensory gain (%)	0.254*	-0.163	-0.287*	-0.565**	0.938**					
7. Detection task: Relative unisensory gain (%)	-0.125	0.019	0.404**	0.181	-0.331**	-0.361**				
8. Continuous recognition task: Multisensory memory gain (%)	0.165	-0.013	-0.085	-0.159	0.353**	0.316**	-0.080			
9. Age-standardised working memory score	0.357**	-0.504**	-0.476**	-0.519**	0.270*	0.220	-0.035	-0.002		
10. Age-standardised fluid intelligence score	0.165	-0.278*	-0.192	-0.325**	0.189	0.258*	0.125	-0.027	0.422**	

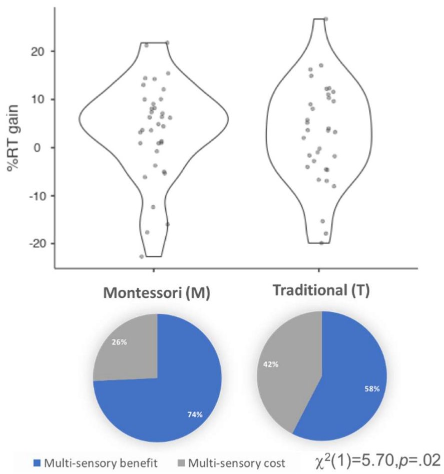
Supplementary Table 1. Correlation matrix reporting Pearson's correlation coefficients between pairs of measures. * indicates p<0.05; ** indicated p<0.01

	Predictors		
	Visual RTs + age	Auditory RTs + age	Multisensory RTs + age
Dependent variable			
Continuous recognition task	Neither variable improved the model. ($F_{(2,65)}$=1.34; p=0.27)	Neither variable improved the model. ($F_{(2,65)}$=0.918; p=0.404)	Neither variable improved the model. ($F_{(2,65)}$=1.116; p=0.334)
Age-standardised working memory score	Significantly improved the fit between the model and data, χ^2(8, N=68) = 27.22, Nagelkerke R^2= 0.354, p< 0.001). However, there were no significant unique contributions by either visual RTs [χ^2(4, N=68) = 7.702; p=0.103] or age [χ^2(4, N= 68) = 7.841; p=0.098]. Goodness of fit was explored by using the Pearson chi-square statistic, which was not significant (p=0.978).	Significantly improved the fit between the model and data, χ^2(8, N=68) = 24.96, Nagelkerke R^2= 0.330, p= 0.002). However, there were no significant unique contributions by either auditory RTs [χ^2(4, N=68) = 5.434; p=0.246] or age [χ^2(4, N= 68) = 8.091; p=0.088]. Goodness of fit was explored by using the Pearson chi-square statistic, which was not significant (p=0.995).	Significantly improved the fit between the model and data, χ^2(8, N=68) = 26.947, Nagelkerke R^2= 0.351, p< 0.001). However, there were no significant unique contributions by either multisensory RTs [χ^2(4, N=68) = 7.425; p=0.115] or age [χ^2(4, N= 68) = 6.200; p=0.185]. Goodness of fit was explored by using the Pearson chi-square statistic, which was not significant (p=0.985).
Age-standardised fluid intelligence score	($F_{(1,66)}$=5.54; p=0.022), though age did not significantly contribute to the model (p=0.953)	Neither varibale improved the model. ($F_{(2,65)}$=1.383; p=0.258)	($F_{(1,66)}$=7.81; p=0.007), though age did not significantly contribute to the model (p=0.786)

Supplementary Table 2. Results of multiple regression analyses using mean reaction times on the detection task and age as predictors of relative multisensory memory gain on the continuous recognition task, age-standardised working memory scores, or age-standardised fluid intelligence scores.



Supplementary Figure S1. Experimental design. A. Simple Detection task; V stands for visual-only sitmuli, A for auditory-only and AV for audio-visual ones. **B.** Continuous Recognition task; lines of black and white drawings taken from the Snodgrass and Vanderwart (1980) database. The first encounters were combined with meaningless (AVm), congruent (AVc) auditory sounds, or not (V). The recall conditions were always unisensory, visual only (V+m, V+c, and V-, respectively).



Supplementary Figure S2. The reaction time gain (*top*) according to scholastic background of the children (Montessori or traditional). By setting a threshold at zero, we could classify the scholars showing a multisensory gain (%RT>0ms) or cost (%RT<0ms). Here, the subsequent percentages per system are displayed, and statistically tested as significantly different.

MIND, BRAIN, AND EDUCATION

Effects of Traditional Versus Montessori Schooling on 4- to 15-Year Old children's Performance Monitoring

Solange Denervaud[1,2,3], Jean-François Knebel[2], Mary Helen Immordino-Yang[4,5,6], and Patric Hagmann[3,†]

ABSTRACT— Through performance monitoring individuals detect and learn from unexpected outcomes, indexed by post-error slowing and post-error improvement in accuracy. Although performance monitoring is essential for academic learning and improves across childhood, its susceptibility to educational influences has not been studied. Here we compared performance monitoring on a flanker task in 234 children aged 4 through 15, from traditional or Montessori classrooms. While traditional classrooms emphasize that students learn from teachers' feedback, Montessori classrooms encourage students to work independently with materials specially designed to support learners discovering errors for themselves. We found that Montessori students paused longer post-error in early childhood and, by adolescence, were more likely to self-correct. We also found that a developmental shift from longer to shorter pauses post-error being associated with self-correction happened younger in the Montessori group. Our findings provide preliminary evidence that educational experience influences performance monitoring, with implications for neural development, learning, and pedagogy.

[1]The Center for Affective Sciences (CISA), Campus Biotech, University of Geneva
[2]The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology and Department of Clinical Neurosciences, Lausanne University Hospital and University of Lausanne (CHUV-UNIL)
[3]Connectomics Lab, Department of Radiology, Lausanne University Hospital and University of Lausanne (CHUV-UNIL)
[4]Brain and Creativity Institute, University of Southern California
[5]Rossier School of Education, University of Southern California
[6]Neuroscience Graduate Program, University of Southern California

Address correspondence to Solange Denervaud, The Center for Affective Sciences (CISA), Campus Biotech, University of Geneva, 1202 Geneva, Switzerland; e-mail: solange.denervaud@chuv.ch

[†] Both authors contributed equally.

To be successful in school, children must learn to distinguish correct from incorrect responses. To do this, they must monitor their own performance, noticing when their work is flawed and efficiently adapting their behavior. That is, they must notice, learn from and correct errors.

Performance monitoring comprises the set of behavioral and neuronal responses that individuals show in reaction to unexpected outcomes. When individuals notice something unexpected, such as an error, they tend to pause, a phenomenon known as post-error slowing (PES; Ullsperger, Danielmeier, & Jocham, 2014). To learn from the discrepant event, they must adapt their behavior accordingly (Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004), a phenomenon measured as post-error improvement in accuracy (PIA; Danielmeier & Ullsperger, 2011; Schroder & Moser, 2014).

The response to own error is measured through conflict tasks (e.g., Flanker task, Simon task, Go/noGo task), where congruent and incongruent conditions require task-specific response monitoring. PES can be computed as the reaction time (RT) difference between post-error and post-correct responses. While a multitude of infant looking time and electrophysiological studies show that by 2–3 months children detect and respond to deviant events (Dehaene-Lambertz, 2000; Dehaene-Lambertz & Gliga, 2004), PES has not been measured in children until about age 3 years (the Simple Simon Task; Jones, Rothbart, & Posner, 2003). However, developmental data are inconsistent, sometimes showing invariant (Davies, Segalowitz, & Gavin, 2004; Ladouceur, Dahl, & Carter, 2007; Santesso, Segalowitz, & Schmidt, 2006), and sometimes increasing (Santesso & Segalowitz, 2008) or decreasing (Carrasco et al., 2013; Hajcak, Franklin, Foa, & Simons, 2008; Smulders, Soetens, & van der Molen, 2016) PES across age. Neural developmental studies have also reported inconsistent outcomes in error-related brain components and their trajectories (Tamnes, Walhovd, Torstveit, Sells, & Fjell, 2013). These contradictory results may reflect task design features such as interstimulus timing or difficulty level (Dutilh et al., 2012), but they have also been proposed to reflect the instability of children's error monitoring (Smulders et al., 2016). Taken together with a behavioral study reporting that PIA increases with age (Overbye et al., 2019), such data suggest the possibility that the relationship between PES and PIA changes with development (Brewer & Smith, 1989). Perhaps longer post-error pauses are adaptive in early childhood, when children require time to effortfully stop and redirect, while shorter pauses post-error may be the hallmark of efficient and automatic self-correction by adolescence. This would be in line with the literature reporting a shift in how children process learning signal saliency (van den Bos, Guroglu, van den Bulk, Rombouts, & Crone, 2009; van Duijvenvoorde, Zanolie, Rombouts, Raijmakers, & Crone, 2008).

In addition to the possibility of a developmental shift in adaptive use of post-error pauses, these apparently contradictory results could reflect variability of experience, particularly in school. Such environmental influences would be consistent with known experience-dependent effects on the development of executive functioning (Davidson, Amso, Anderson, & Diamond, 2006), for example in children experiencing different school curricula (Diamond, 2012). If true for executive functions, related competencies could be impacted as well. Here, we hypothesize that even among children from relatively privileged backgrounds, school pedagogical approach may modulate the processes by which children come to notice and respond to their own errors.

Pedagogical traditions differ on how they teach children to learn from feedback, though school-based influences on the development of performance monitoring have not been investigated to our knowledge. For example, traditional education provides children from one age group with opportunities to engage in work, and then to learn about and correct their performance later based on a teacher's feedback. By contrast, Montessori education focuses on supporting children in self-correcting in real time. It utilizes specialized materials that encourage children's self-discovery of relevant concepts, and multiage classes in which children discuss correct and incorrect answers as they work.

Given these open developmental and experience-related questions, the purposes of this study are: (1) to examine PES and PIA development across childhood and early-middle adolescence; and (2) to evaluate whether children from equivalently high-quality pedagogical environments with systematically different pedagogical approaches differ in their developmental trajectories of performance monitoring. Accordingly, we measured PES and PIA among groups of Montessori and traditionally schooled children during the completion of a child-friendly Flanker Task (Eriksen & Eriksen, 1974). We hypothesized that: (a) PES would decrease with age; (b) PIA would increase with age; (c) there would be a developmental shift in the relationship between PES and PIA such that young children's longer pauses, and adolescents' shorter pauses, would be associated with more effective self-correction; (d) these developmental effects would be stronger in the Montessori group; (e) by adolescence, Montessori students would show lower PES and higher PIA than their traditionally schooled peers.

The study was conducted in Switzerland, where alternative education is only found within private schools. As a consequence, the Montessori schools in this study were private. Accordingly, we had no option for a lottery design study. Given these constraints, we selected Montessori and traditional schools from the same neighborhoods and of similarly high academic quality, and confirmed that participating groups of students did not differ on the basis of family socioeconomic status (SES), fluid intelligence (FI), or age.

METHODS

The study was conducted in accordance with the Declaration of Helsinki and the research protocol was approved by the University ethics committee. The study is part of a larger project investigating effects school pedagogy on children's brain development. Teachers' and pupils' participation were voluntary.

Recruitment Site Selection

Schools were selected to include children from affluent areas only, using official government data on mean salary. Traditional and Montessori schools coexisted in the same neighborhoods and were of similar academic quality as judged by adherence to either official Montessori qualifications or government specifications for public schools.

Traditional schools were public and were identified based on their adherence to the official local study plan. This plan strictly implies, as from 7 years old, that children are provided:

(i) lecture-style, interactive teaching;
(ii) feedback in the form of grades and summative assessments;
(iii) breaks every hour;

(iv) age-specific class groupings (i.e., one grade level per class).

Prior to 7 years old, traditionally schooled children are provided play-oriented curricula with emphasis on building gross and fine motor skills and social skills.

Montessori schools were private and were selected following the criteria set by the International Montessori Association (https://montessori-ami.org) to ensure high fidelity implementation of the curriculum (Lillard, 2012), such that:

(i) all teachers were Montessori trained and there was a complete set of Montessori materials in each classroom;
(ii) children were not provided formal grades or summative assessments on work;
(iii) children had 3 hr of continuous work time each day;
(iv) classes contained at least three different grade levels;

Prior to 7 years old, Montessori curricula emphasize independent learning from specialized materials designed to refine the senses and teach early academic skills.

Participants

Ten traditional and 13 Montessori schools were selected to participate. Consent letters were sent home to all parents of children ages 4 to 15, and 234 children's parents consented. An additional four children were consented but excluded due to having made no errors (accuracy rate of 100%; $n = 2$), or being outside the target age range (older than 15.5 years; $n = 2$). The final sample consisted of 234 children from 4.4 to 15.3 years ($M_{age} = 9.02$ years old, $SD = 2.43$ years, 114 girls and 120 boys); 111 participants were enrolled in the traditional educational system (59 girls), and 123 were schooled in the Montessori educational system (55 girls).

Study Protocol

The experiment took place in the children's schools. Participants were seated in a separate room outside their classroom. Assessments of both groups were conducted by four trained experimenters (including the first author).

Group Comparison

To ensure between-group homogeneity, we tested for pedagogical group differences in: (i) age; (ii) gender; (iii) SES (measured through a parental questionnaire (Genoud, 2011) that 78.6% filled out, 84.7% in the traditionally schooled group, and 73.2% in the Montessori group); and (iv) FI, evaluated using the black-and-white version of the Raven's Progressive Matrices task (Raven, Raven, & Court, 1998).

PES and PIA were assessed using RT and accuracy from a child-friendly version of the Flanker task (Eriksen & Eriksen, 1974), presented via Presentation® software. The child was asked to "feed hungry fish" by pressing the key facing the same direction as the target fish in the display (five fish in line). Each child completed three blocks, the first two of which were preceded by training sessions. In the first block, the child was instructed to focus on the fish at the center of the line (17 trials). In this block, all stimuli (fishes) were blue. In the second block, the child was instructed to focus on the fishes flanking the central one (17 trials). In this block, all stimuli were pink. In the final block, the child was instructed to focus either on the inside fish or on the outside, flanker fish, depending on their color (45 trials). Due to RT change across age, the response time limit was 2000 ms up to 6 years old, and 1,500 ms for older children.

Dependent Variables

We examined PES and PIA measures across age and tested for interactions by pedagogical group.

Statistical analyses were done with the Jamovi open-source software (Version 0.9; The jamovi project (2019). Retrieved from https://www.jamovi.org).

RESULTS

Pedagogical Group Demographic Comparison

Independent samples t-tests showed no pedagogical group differences in age, SES background or FI scores (all $p > .4$; see Table 1). A χ^2 test showed no group differences in gender distribution.

Pedagogical Group Task Performance Comparison

One-way analyses of covariance (ANCOVAs) showed no statistically significant differences between Montessori and traditional students on either mean RT or error rate, controlling for age; RT: $F(1{,}230) = 0.34$, $p = .56$, $\eta^2 = 0.001$; error rate: $F(1{,}230) = 1.29$, $p = .26$, $\eta_p^2 = 0.006$. As expected, RT and Error rate significantly decreased with age; RT: $F(1{,}230) = 295.52$, $p < .001$, $\eta_p^2 = 0.56$; error rate: $F(1{,}230) = 167.10$, $p < .001$, $\eta_p^2 = 0.42$.

Post-error Slowing

The detection of errors was investigated by comparing the RT after correct and incorrect responses. Following previous studies' protocol (Eriksen & Eriksen, 1974; McDermott, Perez-Edgar, & Fox, 2007), we computed differences in valid RT[1] (RT > 250 ms and within 2 SD) between post-error and post-correct trials ($RT_{posterror} - RT_{postcorrect}$). We computed a one-way ANCOVA with age as the covariate and pedagogy as a fixed factor, and the interaction term (pedagogy*age). The analysis revealed that PES increased with age, $F(1{,}230) = 4.99$, $p = .026$, $\eta_p^2 = 0.02$. There was no effect of pedagogy on PES, $F(1{,}230) = 2.26$, $p = .134$, $\eta_p^2 = 0.01$, and no

Table 1
Participant Demographics and Control Variables

	Group		
Control Variables	*Montessori (*n = *123)*	*Traditional (*n = *111)*	χ^2 *or* t*-test*
Age, mean (*SD*)	8.89 (2.34)	9.17 (2.53)	$t(232) = -0.87$, $p = .39$, ns
Age min, max	4.40, 15.2	4.50, 15.3	ns
Gender (*N* girls)	52	59	$\chi^2(1{,}231) = 2.23$, $p = .14$,ns
Socioeconomic status (*SD*)	0.70 (0.11)	0.70 (0.12)	$t(182) = -0.006$, $p = .99$, ns
Fluid intelligence, mean (*SD*)	30.1 (7.81)	29.6 (6.61)	$t(209) = 0.56$, $p = .58$, ns
Mean RT (*SD*)	878.28 (199.52)	878.09 (191.28)	$t(232) = 0.007$, $p = .99$, ns
Error rate (*SD*)	20.95 (14.72)	20.94 (12.79)	$t(232) = 0.08$, $p = .99$, ns

Table 2
Groups Per Pedagogy

Pedagogy	*Age Group (Years)*	*Mean Age*	SE
Montessori	4.5–6	5.35	0.159
	6–9	7.75	0.125
	9–11.5	10.19	0.107
	11.5–15	12.36	0.187
Traditional	4.5–6	5.31	0.167
	6–9	7.29	0.167
	9–11.5	10.04	0.106
	11.5–15	12.56	0.163

Notes. To complement our continuous analysis, we additionally ran an age-group discrete analysis. Accordingly, children were divided into four groups. A two-way analysis of variance confirmed that mean ages did not differ by pedagogical group, $F(3{,}226) = 1.49$, $p = .217$. *SE* stands for standard error.

significant interaction between pedagogy and age on PES, $F(1{,}230) = 1.76$, $p = .186$, $\eta_p^2 = 0.001$. To accommodate the possibility that the developmental change would not be linear, we ran an additional age-group analysis. We divided the children into four age groups with cut points corresponding as closely as possible to school transitions, i.e., from Kindergarten to grade school; from early to middle grade school; from middle grade school to preadolescent classrooms; from early adolescence onward. Groups spanned an average of 2.5 years (see Table 2). To test differences in developmental pattern in PES, we computed a two-way analysis of variance (ANOVA) with age group, pedagogy, and the interaction term (age group*pedagogy) as fixed factors and confirmed an age-group difference, $F(1{,}226) = 2.67$, $p = .048$, $\eta_p^2 = 0.034$. A post-hoc t-test revealed that pedagogical groups differed in PES from age 4.5 to 9 years; traditional students showed significantly less PES in this age range, $t(226) = -3.16$, $p_{\text{tukey}} = 0.038$; see Figure 1a.

Post-error Improvement in Accuracy

Self-correction post-error was computed as the ratio of error trials that were immediately followed by a correct trial, divided by the total number of error trials. The ANCOVA analysis revealed that PIA increased with age, $F(1{,}230) = 104.76$, $p < .001$, $\eta_p^2 = 0.31$. Pedagogy was associated with PIA at the trend level, $F(1{,}230) = 3.09$, $p = .08$, $\eta_p^2 = 0.013$, with overall self-correction post-error slightly higher in Montessori students. As hypothesized, there was a significant interaction between pedagogy and age on PIA, $F(1{,}230) = 4.08$, $p = .045$, $\eta_p^2 = 0.02$, such that Montessori students showed a stronger developmental increase in self-correction. A two-way ANOVA on PIA by age group, pedagogy, and the interaction term (age group*pedagogy) further confirmed a significant increase across age groups, $F(1{,}226) = 32.39$, $p < .001$, $\eta_p^2 = 0.30$, an effect that was stronger in the Montessori old schoolers ($p_{\text{tukey}} < 0.025$; see Figure 1b).

The relationship between PES, PIA, and Pedagogy shifted with development

In a linear regression model, PES, age, pedagogy, and the interaction terms (age*PES, age*pedagogy, pedagogy*PES) were all predictors of PIA ($R^2 = 0.34$; all $p < .035$), such that as students grew older, the more PES they showed, the greater PIA they showed. Pedagogy was nearly a significant predictor of PIA in this model ($p = .052$). Furthermore, there was a significant interaction between pedagogy and age such that Montessori students showed a stronger and earlier developmental effect, $F(1{,}227) = 5.72$, $p = .018$; see Figure 2. By adolescence (age 11–15 years), Montessori students showed lower PES and greater PIA than their traditionally schooled peers, PES: $t(46) = 2.13$, $p = .039$; PIA: $t(46) = 2.14$, $p = .037$.

DISCUSSION

In this study, we compared PES and PIA in traditional and Montessori schoolers across age, with the aim of mapping developmental differences in performance monitoring in the flanker task, and observing how developmental changes in performance monitoring may be influenced by pedagogy.

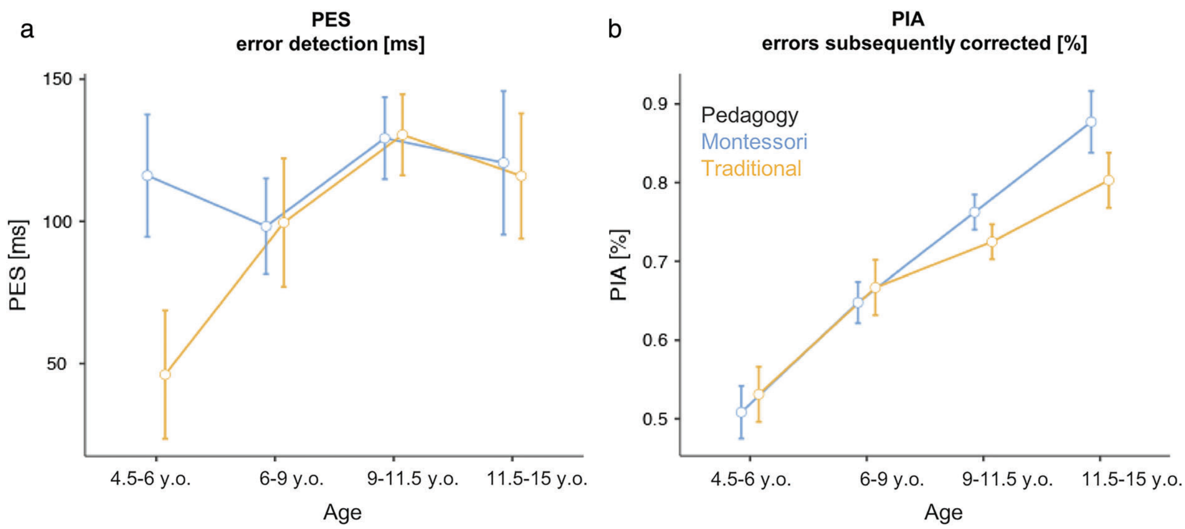


Fig. 1. (A) Post-error slowing (PES) as a measure of error detection. ΔRT (M_{RT} of trials$_{incorrect}$ − trials$_{correct}$[ms] over all valid trials) as a function of age group showing that young children differ in their PES pattern according to their pedagogical experience, while older students do not. (B) Post-error improvement in accuracy (PIA) as the percentage of errors that were subsequently corrected. Post-error correct trials over the total number of errors as a function of age-group showing that 9- to 15-year-old students differ in PIA pattern according to their pedagogical experience, while younger students do not. For both graphs, error bars display standard error (*SE*).

The age range included in our study, from 4 to 15 years old, corresponds to the broad developmental period in which the brain regions responsible for conflict monitoring and error recognition and correction are maturing (Kelly et al., 2009; Velanova, Wheeler, & Luna, 2008). It is also the age range in which schooling is known to have a profound effect on cognitive skills (e.g., on intelligence quotient; Falch & Massih, 2011).

We found that, whether attending Montessori or traditional school, as children grew older both their PES and their PIA increased, partially corroborating previous studies (Overbye et al., 2019; Smulders et al., 2016). Interestingly, the developmental increase in PES we observed was largely driven by the younger children (<9 years old) enrolled in traditional schools. In our study, the youngest Montessori students' PES was similar to that of the oldest students, suggesting an earlier maturation of the capacity to detect response errors. These findings may possibly explain inconsistencies between previously reported findings (Smulders et al., 2016); our data suggest that the ability to detect response errors is especially malleable up to 9 years old, and may be trained by schooling. Of note, the number of errors did not differ between pedagogical groups at any age; instead, it was the capacity for self-correction that showed pedagogical effects.

The relationship between PES and PIA also changed with age: in young children, pausing after an error was associated with a subsequent self-correction. This was no longer the case among adolescents, for whom shorter pauses were associated with more effective self-correction. The developmental shift from longer to shorter pauses being associated with more effective self-correction happened at a younger age among Montessori students (approximately at age 8 vs. at age 10). Together these results shed light on the developmental trajectory of performance monitoring across childhood and early mid adolescence, and suggest that error recognition may be modulated by early school experience (<9 years old), with implications for self-correction in adolescence. Future longitudinal work is needed to uncover the developmental processes that undergird our findings, and to better understand the implications of these findings for children's development and learning more broadly.

Over the course of schooling, children learn to associate salient events like task outcomes with context-dependent feedback, and to adjust behavior. Our results could possibly reflect the Montessori curriculum's relative emphasis on students building early awareness of the sensory properties of materials (i.e., learning to discriminate forms, colors, textures, temperatures, etc.), without waiting for a teacher's feedback or external reinforcers (e.g., grades, rewards, etc.; Dolan & Dayan, 2013; Glascher, Daw, Dayan, & O'Doherty, 2010). This may serve to orient Montessori students toward directly perceiving information about outcomes in academic tasks and may teach them to more effectively self-monitor in academic tasks. Whereas the free play orientation used early in traditional schools may benefit children in other ways (Lillard, 2017), it may not orient children to notice errors on academic-style tasks as effectively.

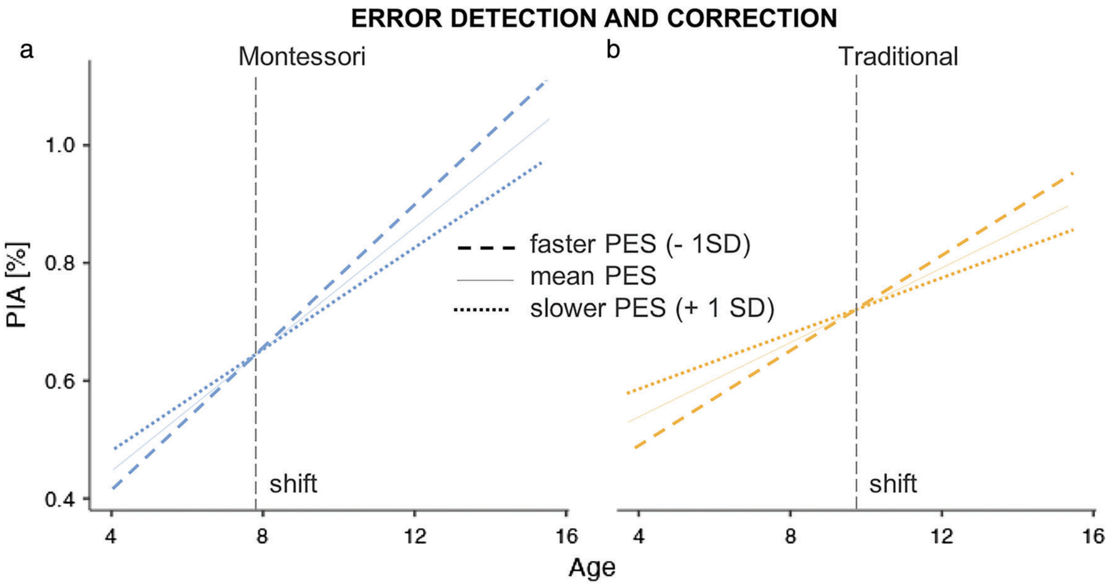


Fig. 2. Pedagogical influence on the relationship between error detection (post-error slowing [PES]) and error correction (post-error improvement in accuracy [PIA]). PIA was reliably predicted by age and age*PES, such that younger children who slowed more after errors (slower PES; mean reaction time [RT] + 1 *SD*) were more likely to self-correct, while older children who slowed less after errors (faster PES; mean RT − 1 SD) were more likely to self-correct. Pedagogy effected this developmental shift, which happened earlier in Montessori than in traditionally schooled students.

The social-affective orientation toward errors in Montessori and traditional classrooms my further contribute to our effects. Montessori students' reliance on direct sensory perceptions as feedback may be reinforced by the culture of the Montessori classroom, which emphasizes non-competitive peer-to-peer teaching within classes of students of different ages. Previous research has demonstrated that cooperative environments lead to more effective shared learning from errors, and hence to better transfer of knowledge (Koban, Pourtois, Vocat, & Vuilleumier, 2010). By contrast, the delayed, teacher-provided evaluative feedback in traditional classrooms may lead students to become increasingly reactive to errors, which come to be negatively valenced, privately conveyed from the teacher to the student, and socially stigmatized. As formal instruction is instituted and the grading system for external valuation of work takes hold, traditionally schooled children learn to value correct answers provided by their teachers and to avoid, rather than productively engage with, errors (Hayek, Toma, Guidotti, Oberlé, & Butera, 2017; Hayek, Toma, Oberle, & Butera, 2014, 2015). Indeed, a study suggests that correct answers are selectively associated with positive valence in traditionally schooled children, while Montessori students show no such effect (Denervaud et al., in review).

Our interpretation builds from the basic distinction between internally derived, sensory learning and externally derived, value-based learning that is fundamental to learning theories, including for example to work on extrinsic versus intrinsic motivation (Oudeyer, Gottlieb, & Lopes, 2016; Ryan & Deci, 2000) and reinforcement learning (Dolan & Dayan, 2013; Frank, Woroch, & Curran, 2005; Glascher et al., 2010; Worthy, Cooper, Byrne, Gorlick, & Maddox, 2014). If we are correct, it would suggest that many types of experiences that support children in building skills for safe and adaptive self-directed learning should support children's performance monitoring development. It is possible that, for example, the strong outcomes found with well-designed and supported project-based learning (Condliffe et al., 2017; Knecht, Gannon, & Yaffe, 2016) could have a similar effect, or, for that matter, increased proportions of productive self-directed time during childhood (Barker et al., 2014).

This study has two main limitations. First, the study is cross-sectional. Though the study includes a continuous age range of students, longitudinal work with children experiencing different styles of pedagogy would increase confidence in the developmental trajectories described here. Second, the study has a nonrandomized design. It is possible that parents who value self-directed behavior themselves are more likely not only to enroll their children into Montessori schools but also to encourage their child to err and autonomously correct themselves on academic tasks at home. That said, we find it unlikely that the effects we report here are entirely due to selection biases. A study of scholastic, social–emotional, and creativity measures in students from the schools included here produced similar effect sizes as have existing randomized studies in other Montessori and traditional schools (Lillard & Else-Quest, 2006; Lillard et al., 2017; Denervaud, Knebel, Hagmann, & Gentaz, 2019).

In conclusion, our findings suggest that children from 4 to 15 years of age develop performance monitoring, and that this development may be influenced by the pedagogical approach they experience at school. Our findings provide preliminary evidence that inconsistencies in previous studies may be due in part to classroom experience. Our results also suggest the possibility that early and strong development of the capacity to discriminate errors, though this requires time in young children, may have downstream developmental effects on the ability to efficiently self-correct in adolescence.

Debates in education often focus on the most appropriate ways to optimally provide content-area knowledge to students. Our findings reiterate the importance of understanding how pedagogical orientations toward the learning process, and not simply curricular content itself, are important factors shaping children's development. The findings also point to the need for nuanced developmental studies of children's cognitive, affective, and social capacities and their interactions in various school contexts (Immordino-Yang, Darling-Hammond, & Krone, 2019). Given the rapid pace of societal change youth face, it is of utmost importance that educational experiences equip children not simply with knowledge but with abilities to effectively and flexibly learn independently. Our study provides a small step toward connecting research on performance monitoring with the development of mental processes in children's educational environments.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interests for this article.

Acknowledgment—We thank the children and their parents for their participation. We are also grateful to Prof. Kerstin von Plessen for comments on an earlier version of this manuscript. This research was supported by the National Center of Competence in Research (NCCR) for Affective sciences financed by the Swiss National Science Foundation (nº51NF40-104897) and hosted by the University of Geneva. SD was funded by the Boninchi Foundation in Geneva and the Department of Radiology of Lausanne University Hospital (CHUV).

ENDNOTE

1 *Analysis with raw RTs strengthens the findings. However, to avoid that outlier trials drive the significant results, RTs outside 2 SD were filtered out*.

REFERENCES

Barker, J. E., Semenov, A. D., Michaelson, L., Provan, L. S., Snyder, H. R., & Munakata, Y. (2014). Less-structured time in children's daily lives predicts self-directed executive functioning. *Front Psychol*, *5*, 593. doi:10.3389/fpsyg.2014.00593

Brewer, N., & Smith, G. A. (1989). Developmental-changes in processing speed—Influence of speed accuracy regulation. *Journal of Experimental Psychology-General*, *118*(3), 298–310. https://doi.org/10.1037/0096-3445.118.3.298

Carrasco, M., Hong, C., Nienhuis, J. K., Harbin, S. M., Fitzgerald, K. D., Gehring, W. J., & Hanna, G. L. (2013). Increased error-related brain activity in youth with obsessive-compulsive disorder and other anxiety disorders. *Neuroscience Letters*, *541*, 214–218. https://doi.org/10.1016/j.neulet.2013.02.017

Condliffe, B., Quint, J., Visher, M. G., Bangser, M. R., Drohojowska, S., Saco, L., & Nelson, E. (2017) Project-Based Learning; A Literature Review. Retrieved from https://files.eric.ed.gov/fulltext/ED578933.pdf

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology*, *2*, 233. https://doi.org/10.3389/fpsyg.2011.00233

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

Davies, P. L., Segalowitz, S. J., & Gavin, W. J. (2004). Development of response-monitoring ERPs in 7- to 25-year-olds. *Developmental Neuropsychology*, *25*(3), 355–376. https://doi.org/10.1207/s15326942dn2503_6

Dehaene-Lambertz, G. (2000). Development of phonological perception in children: Electrophysiological studies. *Revue de Neuropsychologie*, *10*(4), 519–533.

Dehaene-Lambertz, G., & Gliga, T. (2004). Common neural basis for phoneme processing in infants and adults. *Journal of Cognitive Neuroscience*, *16*(8), 1375–1387. https://doi.org/10.1162/0898929042304714

Denervaud, S., Knebel, J. F., Hagmann, P., & Gentaz, E. (2019). Beyond executive functions, creativity skills benefit academic outcomes: Insights from Montessori education. *PLoS One*, *14*(11), e0225319. https://doi.org/10.1371/journal.pone.0225319

Diamond, A. (2012). Activities and programs that improve children's executive functions. *Current Directions in Psychological Science*, *21*(5), 335–341. https://doi.org/10.1177/0963721412453722

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*(2), 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., & Wagenmakers, E. J. (2012). Testing theories of post-error slowing. *Attention, Perception, & Psychophysics*, *74*(2), 454–465. https://doi.org/10.3758/s13414-011-0243-2

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a non- search task. *Perception & Psychophysics*, *16*, 143–149.

Falch, T., & Massih, S. S. (2011). The effect of education on cognitive ability. *Economic Inquiry*, *49*(3), 838–856.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, *47*(4), 495–501. https://doi.org/10.1016/j.neuron.2005.06.020

Genoud, P. (2011). Indice de position socioéconomique (IPSE): Un calcul simplifié. Retrieved from https://www3.unifr.ch/cerf/fr/assets/public/pdf/pages%20personnelles/Genoud%20Philippe/IPSE.pdf

Glascher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595. https://doi.org/10.1016/j.neuron.2010.04.016

Hajcak, G., Franklin, M. E., Foa, E. B., & Simons, R. F. (2008). Increased error-related brain activity in pediatric obsessive-compulsive disorder before and after treatment. *American Journal of Psychiatry*, *165*(1), 116–123. https://doi.org/10.1176/appi.ajp.2007.07010143

Hayek, A. S., Toma, C., Guidotti, S., Oberlé, D., & Butera, F. (2017). Grades degrade group coordination: Deteriorated interactions and performance in a cooperative motor task. *European Journal of Psychology of Education*, *32*, 97–112.

Hayek, A. S., Toma, C., Oberle, D., & Butera, F. (2014). The effect of grades on the preference effect: Grading reduces consideration of disconfirming evidence. *Basic and Applied Social Psychology*, *36*(6), 544–552. https://doi.org/10.1080/01973533.2014.969840

Hayek, A. S., Toma, C., Oberle, D., & Butera, F. (2015). Grading hampers cooperative information sharing in group problem solving. *Social Psychology*, *46*(3), 121–131. https://doi.org/10.1027/1864-9335/a000232

Immordino-Yang, M.-H., Darling-Hammond, L., & Krone, C. R. (2019). Nurturing nature: How brain development is inherently social and emotional, and what this means for education. *Educational Psychologist*, *54*, 1–20. https://doi.org/10.1080/00461520.2019.1633924

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science*, *6*(5), 498–504. https://doi.org/10.1111/1467-7687.00307

Kelly, A. M., Di Martino, A., Uddin, L. Q., Shehzad, Z., Gee, D. G., Reiss, P. T., … Milham, M. P. (2009). Development of anterior cingulate functional connectivity from late childhood to early adulthood. *Cerebral Cortex*, *19*(3), 640–657. https://doi.org/10.1093/cercor/bhn117

Knecht, D., Gannon, N., & Yaffe, C. (2016). Across classrooms: School quality reviews as a progressive educational policy. Bank Street Occasional Paper Series 35.

Koban, L., Pourtois, G., Vocat, R., & Vuilleumier, P. (2010). When your errors make me lose or win: Event-related potentials to observed errors of cooperators and competitors. *Social Neuroscience*, *5*(4), 360–374. https://doi.org/10.1080/17470911003651547

Ladouceur, C. D., Dahl, R. E., & Carter, C. S. (2007). Development of action monitoring through adolescence into adulthood: ERP and source localization. *Developmental Science*, *10*(6), 874–891. https://doi.org/10.1111/j.1467-7687.2007.00639.x

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

Lillard, A. S. (2017). Why do the children (pretend) play? *Trends in Cognitive Sciences*, *21*(11), 826–834. https://doi.org/10.1016/j.tics.2017.08.001

Lillard, A., & Else-Quest, N. (2006). The early years. Evaluating Montessori education. *Science*, *313*(5795), 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. *Front Psychol*, *8*, 1783. doi:10.3389/fpsyg.2017.01783

McDermott, J. M., Perez-Edgar, K., & Fox, N. A. (2007). Variations of the flanker paradigm: Assessing selective attention in young children. *Behavior Research Methods*, *39*(1), 62–70.

Oudeyer, P. Y., Gottlieb, J., & Lopes, M. (2016). Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. *Changing Brains Applying Brain Plasticity to Advance and Recover Human Ability*, *229*, 257–284. https://doi.org/10.1016/bs.pbr.2016.05.005

Overbye, K., Walhovd, K. B., Paus, T., Fjell, A. M., Huster, R. J., & Tamnes, C. K. (2019). Error processing in the adolescent brain: Age-related differences in electrophysiology, behavioral adaptation, and brain morphology. *Dev Cogn Neurosci*, *38*, 100665. doi:10.1016/j.dcn.2019.100665

Raven, J., Raven, J. C., & Court, J. H. (1998) *Manual for Raven's progressive matrices and vocabulary scales. Section 2: The coloured progressive matrices*. San Antonio, TX: Harcourt Assessment.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443–447. https://doi.org/10.1126/science.1100301

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Santesso, D. L., & Segalowitz, S. J. (2008). Developmental differences in effor-related ERPs in middle- to late-adolescent males. *Developmental Psychology*, *44*(1), 205–217. https://doi.org/10.1037/0012-1649.44.1.205

Santesso, D. L., Segalowitz, S. J., & Schmidt, L. A. (2006). Error-related electrocortical responses in 10-year-old children and young adults. *Developmental Science*, *9*(5), 473–481. https://doi.org/10.1111/j.1467-7687.2006.00514.x

Schroder, H. S., & Moser, J. S. (2014). Improving the study of error monitoring with consideration of behavioral performance measures. *Frontiers in Human Neuroscience*, *8*, 178. https://doi.org/10.3389/fnhum.2014.00178

Smulders, S. F., Soetens, E., & van der Molen, M. W. (2016). What happens when children encounter an error? *Brain and Cognition*, *104*, 34–47. https://doi.org/10.1016/j.bandc.2016.02.004

Tamnes, C. K., Walhovd, K. B., Torstveit, M., Sells, V. T., & Fjell, A. M. (2013). Performance monitoring in children and adolescents: A review of developmental changes in the error-related negativity and brain maturation. *Developmental Cognitive Neuroscience*, *6*, 1–13. https://doi.org/10.1016/j.dcn.2013.05.001

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, *94*(1), 35–79. https://doi.org/10.1152/physrev.00041.2012

van den Bos, W., Guroglu, B., van den Bulk, B. G., Rombouts, S. A. R. B., & Crone, E. A. (2009). Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. *Frontiers in Human Neuroscience*, *3*, 1–11. https://doi.org/10.3389/neuro.09.052.2009

van Duijvenvoorde, A. C., Zanolie, K., Rombouts, S. A., Raijmakers, M. E., & Crone, E. A. (2008). Evaluating the negative or valuing the positive? Neural mechanisms supporting feedback-based learning across development. *The Journal of Neuroscience*, *28*(38), 9495–9503. https://doi.org/10.1523/JNEUROSCI.1485-08.2008

Velanova, K., Wheeler, M. E., & Luna, B. (2008). Maturational changes in anterior cingulate and frontoparietal recruitment support the development of error processing and inhibitory control. *Cerebral Cortex*, *18*(11), 2505–2522. https://doi.org/10.1093/cercor/bhn012

Worthy, D. A., Cooper, J. A., Byrne, K. A., Gorlick, M. A., & Maddox, W. T. (2014). State-based versus reward-based motivation in younger and older adults. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(4), 1208–1220. https://doi.org/10.3758/s13415-014-0293-8

Error-Monitoring is Modulated by School Pedagogy

Solange Denervaud[1-3*], Jean-François Knebel[1,4*], Emeline Mullier[5], Patric Hagmann[5], Micah M. Murray [1,4,6,7,†]

[1]The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), 1011 Lausanne, Switzerland

[2]The Center for Affective Sciences, Campus Biotech, University of Geneva, 1202 Geneva, Switzerland

[3]Faculty of Psychology and Educational Sciences (FAPSE), University of Geneva, 1204 Geneva, Switzerland

[4]Electroencephalography Brain Mapping Core, Center for Biomedical Imaging (CIBM) of Lausanne and Geneva, Lausanne, Switzerland

[5]Department of Radiology, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), 1011 Lausanne, Switzerland

[6]Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA

[7]Department of Ophthalmology, Fondation Asile des aveugles and University of Lausanne, Switzerland

*Equal contributions

†E-mail: micah.murray@chuv.ch

Abstract

Within an inherently dynamic environment, unexpected outcomes are part of daily life. Performance monitoring allows us to detect these events and adjust behavior accordingly. The necessity of such an optimal functioning has made error-monitoring a prominent topic of research over the last decades. Event-related potentials (ERPs) have differentiated between two brain components involved in error-monitoring: the error-related negativity (ERN) and error-related positivity (Pe) that are thought to reflect detection vs. emotional/motivational processing of errors, respectively. Both ERN and Pe depend on the protracted maturation of the frontal cortices and anterior cingulate through adolescence. To our knowledge, the impact of schooling pedagogy on error-monitoring and its brain mechanisms remains unknown and was the focus of the present study. Swiss schoolchildren completed a continuous recognition task while 64-channel EEG was recorded and later analyzed within an electrical neuroimaging framework. They were enrolled either in a Montessori curriculum (N=13), consisting of self-directed learning through trial-and-error activities with sensory materials, or a traditional curriculum (N=14), focused on externally driven activities mainly based on reward feedback. The two groups were controlled for age, gender, socio-economic status, parental educational style, and scores of fluid intelligence. The ERN was significantly enhanced in Montessori schoolchildren (driven by a larger response to errors), with source estimation differences localized to the cuneus and precuneus. In contrast, the Pe was enhanced in traditional schoolchildren (driven by a larger response to correct trials), with source estimation differences localized to the ventral anterior cingulate. Receiver operating characteristic (ROC) analysis demonstrated that the ERN and Pe could reliably classify if a child was following a Montessori or traditional curriculum. Brain activity subserving error-monitoring is modulated differently according to school pedagogy.

Key words: Error-monitoring, executive functions, development, pedagogy, event-related potential (ERP)

Introduction

The ability to adjust our actions in response to a constantly changing environment is necessary for optimal goal achievement. This process is called *performance monitoring* and signals the necessity, type and amplitude of adaptation required in the face of *unexpected events*, being adversity, conflict, error, etc. It allows any agent to (i) quickly detect outcomes that deviate from expectation, (ii) integrate feedback and (iii) fine-tune behavior and/or internal models for future decisions (Ullsperger, Danielmeier, & Jocham, 2014). Consequently, human beings can avoid repeating the same mistakes, and adopt a more appropriate behavior. The necessity of such an optimal functioning has made it a topic of interest over the last decades of research. While many developmental studies focused on performance monitoring characterization (i.e. specific features, personality traits, etc.) or individual trait or state differences in healthy or clinical populations (anxiety, mindsets, obsessive-compulsive disorders, etc.), inconsistencies in reported outcomes preclude establishing a clear developmental trajectory. We hypothesized that error-monitoring development is subject to environmental influences, such as learning strategies reinforced within schooling pedagogies. Here, we evaluated the extent to which school-pedagogy modulates children's *error-monitoring.*

There is no direct measure of performance monitoring, but indirect informative markers have been widely studied and reported in the context of erroneous actions. At a behavioral level, studies have generally focused on reaction times on correct compared to incorrect conditions (referred as *post-error slowing*), providing a measure of individual reactivity when facing *errors*. On the other hand, there is a long history of EEG studies examining the time course of events following error commission. Despite differences in the protocols used, a similar sequence of brain responses has been reported in the case of tasks where both speed and accuracy were emphasized. In adults, response-locked brain potentials include two components: an early frontocentral negativity (typically peaking 50-100ms post-response onset), dubbed the *error-related negativity* (ERN) and a later and slower response (~200-400ms post-response onset) with a more central scalp distribution, called the *error positivity* (Pe). They are thought to respectively reflect an early task-unspecific detection of need for adjustments (ERN), and the late task-specific selective attention for orientation and learning (Pe) (Ullsperger & Danielmeier, 2016) or conscious evaluation (Ullsperger et al., 2014). ERN and Pe together reflect a built-in error-detection system (Elton, Band, & Falkenstein, 2000) of respectively low-level perceptive ability for fast detection of incongruency (or mismatch) (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Murphy, van Moort, & Nieuwenhuis, 2016), and for later top-down processes of adaptation/evaluation (Falkenstein, 2000; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001; Overbeek, Nieuwenhuis, & Ridderinkhof, 2005). While these markers are well-studied in adults, developmental data are inconsistent.

Both components are reported to be present in children of preschool age (Brooker, Buss, & Dennis, 2011). For example, the ERN was recorded in children as young as 3-4 years old, suggesting that discrepancy is detected even in young children (Grammer, Carrasco, Gehring, & Morrison, 2014; Smulders, Soetens, & van der Molen, 2016). This process seems to gain in efficiency with age. Indeed, a majority of the pediatric studies on ERN report an increase in amplitude with age, reaching adult-like levels at late adolescence (Tamnes, Walhovd, Torstveit, Sells, & Fjell, 2013). However, there is also evidence that the trajectory of this maturation is non-linear and inconsistent across ERN and Pe components (Tamnes et al., 2013). For example, there are reports of reduced ERN during adolescence (between 10-13 y.o.), despite constant Pe. Such fluctuations are thought to reflect protracted maturation of the underlying related brain structures, such as the anterior cingulate cortex (ACC) (Velanova, Wheeler, & Luna, 2008), and subject to developmental changes.

In fact, part of these variations is related to impaired behavior. Indeed, pediatric clinical populations have been studied, and showed differences in PES or ERN amplitude when compared to healthy controls. For example, the ERN of clinically anxious children differ from the ERN of healthy children (studies from 6-18 y.o.) (Meyer & Gawlowska, 2017; Meyer et al. 2013). The relationship between ERN and anxiety scores appears to be age-dependent; positively correlated in older children and negatively correlated in younger children (Meyer et al., 2012). ERN amplitude was also reported to be altered in obsessive-compulsive disorder patients (Gehring, Himle, & Nisenson, 2000; Hajcak & Simons, 2002), and decreased in ADHD children (Plessen et al., 2016; Liotti, Pliszka, Perez, Kothmann, & Woldorff, 2005; Pliszka, Liotti, & Woldorff, 2000). However, clinical states do not fully explain inter-individual variabilities in ERN and Pe amongst same-age children, nor inconsistencies across studies.

So far, developmental studies on ERN and/or Pe have focused on characterization, personality or clinical comparisons. However, it may be that interaction with the environment plays a key role in the development of these processes and in turn contribute to inter-individual variability. More specifically, the extent to which experience with school pedagogy modulates young children *error-monitoring* responses remains unexplored. In this study, we specifically investigated the possibility of a modulatory effect in 27 schoolchildren experiencing different learning environments; namely Montessori and Swiss traditional pedagogies. Indeed, both approaches address error-monitoring differently. On the one hand, the Montessori pedagogy is based on a self-corrective learning strategy through the use of sensory materials within multi-age classes (Montessori, 1936). On the other hand, the traditional pedagogy is mainly based on externally-driven reward feedback (i.e. grades). Given the different emphasis that Montessori and traditional Swiss pedagogies place on self-corrective vs. externally-driven reward, we predicted that ERN and Pe components would differ across schoolchildren; with early error-detection emphasized in Montessori children. We had no a-priori hypothesis regarding the Pe amplitude.

Methods

Ethics. The experiment was conducted in accordance with the Declaration of Helsinki. Written parental consents were obtained for each child and informed assent was provided by each participant. The experimental procedures were approved by the Ethics Committee of the Hospital Center and University of Lausanne (Vaud CER). No subject had a history of a neurological or psychiatric illness.

Participants. In total, 31 children completed the experiment as part of a larger study, including neuroimaging and behavioral measures aiming at evaluating the impact of school environment on development. Selection criteria were age (5 to 14 y.o.) and school system (children had to be enrolled in a Swiss Montessori or traditional school system from the early years on or for at least 3 years). Children with an error rate lower than 10% were excluded from the study (N=4), and recordings with excessive movement, technical difficulties with EEG recording were removed (N=5). Finally, data from 27 children (mean age ±SD = 9.0±1.9 years; all right-handed with normal or corrected-to-normal vision) were included in the current analyses.

Behavioral control variables. Control variables were collected online post-recording to evaluate between-group homogeneity. Furthermore, we controlled for trait anxiety (STAI-Y2) (Spielberger & Vagg, 1984) and mindsets (implicit theories of intelligence) (Blackwell, Trzesniewski, & Dweck, 2007), given their impact on ERN and Pe measures (Moser, Moran, Schroder, Donnellan, & Yeung, 2013; Moser, Schroder, Heeter, Moran, & Lee, 2011). Finally, to control, as best as possible, for the selection bias of including Montessori scholars from private school systems and children from local traditional schools, we measured the socio-economic status (Largo et al., 1989) and evaluated the parental education style to ensure home environments to be similar (see Supplementary Materials). Finally, children were controlled for fluid intelligence as well (Raven, Raven, & Court, 2003).

Multiple t-tests (independent or Wilcoxon according to the preliminary data check with Q-Q plots and Levene's test) with a 95% confidence interval (CI) for the mean difference were run on the different control variable scores to statistically determine significant differences between the two groups of schoolers (Montessori vs control), with a false-rate discovery (FDR) p-value correction at q=0.05. None of the multiple t-tests were found to be significant (p>0.05), revealing comparable groups in term of age, gender, fluid intelligence, trait anxiety, and mindsets. Furthermore, parental SES and education style at home were similar as well (**Table 1**).

EEG task and procedure. Children performed a continuous recognition task, modified for children from Thelen et al. (2014) and which required the discrimination of initial (i.e., 'new') from repeated (i.e., 'old') presentations of line drawings. Initial and repeated presentations of an item were pseudo-randomized within a block of trials. Each object was repeated only once within each experimental block. The line drawings were taken from a standardized set or obtained from an online library (dgl.microsoft.com). The experiment took place in a sound-attenuated chamber (MDL 102126E from Whisperroom Inc.), where children were seated centrally in front of a 20" LCD computer monitor that was located about 80cm away from them to produce a visual angle of ~4°. The task was designed and controlled by PsychoPy 3.0 (Peirce et al., 2019), and all behavioural data were recorded in conjunction with a serial response box (Psychology Software Tools, Inc.; www.pstnet.com).

EEG acquisition and pre-processing. Continuous EEG was acquired at 1024Hz through a 64-channel Biosemi ActiveTwo AD-box (http://www.biosemi.com) referenced to the common mode sense (CMS; active electrode) and grounded to the driven right leg (DRL; passive electrode), which functions as a feedback loop driving the average potential across the electrode montage to the amplifier zero (full details, including a diagram of this circuitry, can be found at http://www.biosemi.com/faq/cms&drl.htm).

Pre-processing and analyses were performed using both homemade python scripts using Anaconda distribution (Python Software Foundation. Python Language Reference, version 2.7. Available at http://www.python.org) and the Cartool freeware (Brunet et al. 2011; cartoolcommunity.unige.ch). Data were first filtered with a 2nd order Butterworth filter (12dB/octave roll-off; 01.Hz high-pass; 60Hz low-pass; 50Hz notch). The filters were computed linearly in both forward and backward directions to eliminate phase shifts. Then EEG epochs were time-locked to the motor response and spanned 500ms pre-response and 500ms post-response. Epochs with amplitude deviations in excess of ±100µV at any channel, with the exception of those labeled as 'bad' due to poor electrode-skin contact or damage, were considered artifacts and were excluded. Data from 'bad' channels ($M_{electrodes}$=2.04, SD=2.49) were interpolated using 3D splines (Perrin et al., 1987). Prior to group-averaging response-locked potentials were baseline-corrected (-500 to -400ms from the response).

Error-related components. Two error-related components were extracted: the ERN and the Pe. The time periods for these components were selected based on hierarchical topographic clustering analyses of our dataset (Murray et al., 2008). Briefly, topographic clustering is a data-driven way to identify, at the group-level, time periods of stable electrical field configurations on the scalp surface. These stable configurations are characterized by a topographic map or so-called template map (since they are based on group-averaged data). These hypothetical segments were then statistically tested with a fitting procedure and resulting maps were used as relevant physiological markers of periods of interest. For the ERN, stable scalp topography was identified from -25ms to 25ms, and for the Pe stable topography was observed from 200ms to 500ms (see Supplementary Materials).

Statistical analyses - Behavioral data. First, independent t-tests were computed on accuracy and error rates, to statistically evaluate group differences (Montessori vs. traditional). Second, a mixed model analysis of variance (ANOVA) with response condition (correct, incorrect) as the within-subject factor and group (Montessori, traditional) as the between-subjects factor, was run on RTs (with ?<0.05).

Electrical neuroimaging analyses. The analyses of response-locked data were based on the hypothesis that school pedagogy (Montessori vs. traditional) would affect the differential neural response between correct and incorrect trials; the latter of which would nonetheless were anticipated to differ based on prior observations of ERN and Pe in schoolchildren (e.g. Hogan et al., 2005). Based on previous literature (e.g.Aarts, De Houwer, & Pourtois, 2013; Grammer et al., 2014; Meyer, Weinberg, Klein, & Hajcak, 2012) and according to our dataset (topographic clustering), we selected scalp locations (Fpz and Cz) and the time period of interest (-25 to 25ms and 200-500ms) respectively for the ERN and Pe components (**Figure 1**). We performed the statistical analysis on the area under the curve. For each component, a mixed model ANOVA was conducted with response condition (correct, incorrect) as the within-subjects factor and group (Montessori, traditional) as the between-subjects factor. Post-hoc Tukey tests were performed when appropriate. Additionally, we applied the local auto-regressive average distributed linear inverse solution (LAURA) to these response-locked data to visualize and statistically contrast the likely underlying sources of the effects identified during the preceding steps of analysis of the surface-recorded response-locked potentials, with use of the software brain template for 7.5-13.5 year-old children (provided by Cartool). Statistical analysis entailed the same mixed model design as above and was performed using STEN software (Knebel & Notter, 2012). A spatial extension criterion of at least 10 contiguous significant nodes ($p<0.05$) was applied. F-maps thresholded by significant points was displayed.

Binomial logistic regression analysis. We tested the specificity and sensitivity with which the two brain components (ΔERN and ΔPe) could classify the schooling system in which a child was enrolled (Montessori vs. traditional). The results of this analysis are reported as the area under the receiver operating characteristic (ROC) curve (AUC).

Results

Behavioral outcomes. Overall, participants committed an average of 17% (SD=12%) errors, with an accuracy rate of 74%(SD=20%). Age was unrelated to accuracy ($r_{(27)}=0.10$, $p=0.63$). Both error and accuracy rates were comparable in the two experimental groups (respectively $t_{(25)}=0.56$, $p=0.58$, Cohen's d=0.21 and $t_{(25)}=0.46$, $p=0.65$, Cohen's d=0.18). As shown in previous studies, participants were faster on incorrect than correct trials ($M_{RT}\pm SD$ = 926±360ms vs. 1040±436ms; $F_{(1,25)}=7.39$, $p=0.012$, $\eta_p^2=0.23$). There was no evidence of a significant effect of group ($F_{(1,25)}<1$, $p=0.84$, $\eta_p^2=0.00$) nor of an interaction between these two factors ($F_{(1,25)}<1$, $p=0.93$, $\eta_p^2=0.00$). As intended, we obtained comparable behavioral outcomes in both groups (**Table 2**).

Electrophysiological results. We calculated the summed amplitude of the ERP during the ERN and Pe time windows for both correct and incorrect trials as well as both groups of schoolers (**Figure 2A**). The ERN summed amplitude over the time period was significantly more negative on incorrect than correct trials (-173±584μV vs. 104±522μV; $F_{(1,25)} = 15.81$, $p < 0.001$, $\eta_p^2=0.34$). There was a significant interaction of action (correct, incorrect) on group (Montessori, traditional), $F_{(1,25)} = 7.19$, $p=0.013$, $\eta_p^2=0.22$. A post-hoc Tukey test revealed that Montessori schoolers differed significantly at $p_{\text{tukey}}<.001$ between the incorrect and correct conditions, while there was no reliable difference for the traditional schoolers ($p_{\text{tukey}}=0.788$). Analysis of the Pe amplitude revealed a significant main effect of action (correct, incorrect)($F_{(1,25)} = 4.53$, $p =0.043$, $\eta_p^2=0.15$) as well as a non-significant trend for an interaction of action (correct, incorrect) on group (Montessori, traditional), $F_{(1,25)}=3.08$, $p=0.09$, $\eta_p^2=0.11$. A post-hoc Tukey test revealed that traditional schoolchildren presented a significant difference at $p_{\text{tukey}}=0.045$, while there was no reliable difference for the Montessori schoolers ($p_{\text{tukey}}=0.994$). There was no evidence that the magnitude of the Pe and ERN were correlated ($r=-0.05$, $p=0.80$). The results of analyses of distributed source estimations are displayed in **Figure 2B**. For both the ERN and Pe there was a significant interaction between action and pedagogy. For the ERN period, loci exhibiting a significant interaction were limited to the precuneus. The Cartesian coordinates (Talairach & Tournoux, 1988) of the maximal F-value was -17, -70, 19 mm, which is situated within Brodmann's Area 18. Significant voxels extended into the cuneus. For the Pe, voxels exhibiting a significant interaction were limited to the anterior cingulate cortex, with a maximal F-value at the coordinates 10, 44, 7 mm, which is situated within Brodmann's Area 32.

Binomial logistic regression analysis. The area under the ROC-curve (AUC) is related to the overall ability of the test to correctly identify schoolers based on multivariate logistic regression using the predicted probability from the combined ΔERN and ΔPe values (**Figure 2C**). The AUC 0.824 (95% CI =0.667-0.982) and was significantly above 0.5 change levels (p=0.004).

Discussion

The current study demonstrated the effect of schooling environment on the processing of errors. Using standardized electrophysiological measures, we showed that there was a differential expression of both when and where in the brain children in Montessori vs. traditional schooling processed errors, despite no reliable differences in their behavior.

Children were controlled for their level of fluid intelligence, trait anxiety, mindsets, socio-economic background and education at home. While outcomes were similar in these external measures, this was also apparent with the task-related behavioral outcomes. Reaction time and performance analyses revealed no reliable differences between schooling systems, suggestive of well-controlled groups on multiple levels, minimizing contributions of a selection bias.

Regarding the electrophysiological data, children enrolled in Montessori schools exhibited a larger early component of error-detection. Montessori curricula offer pedagogical tools in the form of sensory materials that are self-corrective, so that children, from early years on, are trained to explore and solve unexpected outcomes on their own (Montessori, 1936). Not only is the feedback direct, but is also conveyed through the senses (Lillard, 2017). Their error-monitoring may thus be shaped or reinforced in regard to this early perceptive detection response, in line with the recent framework suggesting ERN to be an orienting reflex (Ullsperger & Danielmeier, 2016). While an increased ERN is often related to anxiety or obsessive-compulsive disorders (e.g. Meyer et al., 2012;Endrass, Riesel, Kathmann, & Buhlmann, 2014; Gehring et al., 2000; Hajcak & Simons, 2002), students with higher academic outcomes also show strong responses to deviant actions. In such cases, increased ERN was argued to reflect higher cognitive resources during cognitive control tasks and was found to be predictive of academic performance in undergraduate students (Hirsh & Inzlicht, 2010). However, in college students, the ERN has been linked to mood and affect (Luu, Collins, & Tucker, 2000). Here, we would argue that beyond greater cognitive control or affective state, Montessori children appear prone to early perception of discrepancies. In fact, the ERN may be related to sensory prediction error, as framed by the reinforcement learning theory. Indeed, a study in 5-7 y.o. (N=18) young children revealed that ERN was not modulated by error value while error-detection was still present (Torpey, Hajcak, Klein, 2009). A study in adults reported the ERN to be predictive of reinforcement learning, with larger ERNs in better learning participants (Frank, Woroch, & Curran, 2005). In our study, the distinction between correct and incorrect actions seems clear in Montessori students, while not reliably so for students in traditional schooling. By contrast, the later component Pe, is often reported to reflect more value-based processes (e.g. motivation, evaluation)(Overbeek et al., 2005). In adults, its amplitude is modulated by a happy mood (positivity) and mindsets (Moser et al., 2011), suggesting motivational salience and/or conscious appraisal of errors (Paul, Walentowska, Bakic, Dondaine, & Pourtois, 2017). However, much less is known about the Pe, and its function is currently debated (Nieuwenhuis, Ridderinkhof, Blow, Band, & Kok, 2001; Overbeek et al., 2005). Here, the Pe was been found to be larger in traditional schoolers, while not in Montessori schooling children. Interestingly, it was the correct condition that elicited higher amplitude. Together, our results suggest development of error-monitoring competencies of different kinds according to the school pedagogy, revealing modulatory effect in school age children. This suggests that, underlying maturation of the related brain structures are prone to modulations by daily-life experience in general, and at school in particular.

Source estimations in the present study confirmed differences in underlying active brain networks, with two main regions exhibiting statistically reliable interactions. First, Montessori and traditional scholars differ during the early stage of error-monitoring (ERN) within both the cuneus and precuneus. The cuneus is related to visual processes and may reflect an attention-related function (Simpson et al., 2011). It may be that errors elicit a subsequent increase in selective attention. Here we would contend that early precuneus coupled with cuneus activities are related to sensory consciousness of discrepancy. Indeed, the cuneus is tightly related to the precuneus, which in turn, is acknowledged to have a widespread connectivity with subcortical and cortical structures. It thus is thought to play a central role in many cognitive processes, but mainly those related to consciousness (Cavanna & Trimble, 2006). Furthermore, it was shown to be subdivided into an anterior part implying self-related functions and a posterior one tightly related to episodic memory retrieval (Cavanna & Trimble, 2006). We thus further hypothesize that visuo-spatial imagery (also reported to activate the precuneus) differences between expectation and outcome are related to internal model representation and embodiment (sensory learning) (Immordino-Yang, 2009). Further studies would be required to investigate whether early performance monitoring markers are related to internal model revision.

Source estimations during the later component (Pe) indicated that brain activity in the anterior cingulate cortex (ACC) was significantly different across Montessori and traditional schoolchildren. This is in close agreement with previous work on error or conflict-monitoring where the ACC is robustly activated (e.g. (Carter et al., 1998) (Kiehl, Liddle, & Hopfinger, 2000; van Veen & Carter, 2002), and linked to greater prefrontal activity (Botvinick, Cohen, & Carter, 2004). In the current study, differences were most robust in the ventral region of the ACC. While the dorsal part of the ACC is related to cognitive processes, the ventral one is thought to reflect emotional components (Bush, Luu, & Posner, 2000). The latter is thought to assess the saliency of emotional and motivational information, and is connected with the amygdala, nucleus accumbens, hippocampus, and anterior insula. It may thus be that error evaluation and motivation are dissimilar in Montessori and traditional schoolchildren.

Here, we show that error-monitoring is modulated by daily-life experience such as the school learning environment. As depicted by the ROC analysis, children could be classified based on their two error-related brain components significantly above chance levels. On the one hand, Montessori scholars seem to be trained at early and sensory error-detection (A. Lillard, 2013; A. S. Lillard, 2005). On the other hand, traditional schoolchildren show positive reinforcement in motivational later component, potentially reflecting value-based reward learning. In conclusion, we would posit that different pedagogies effectively emphasize different aspects of error-monitoring; in the cases here perception vs. evaluation of errors. Our results call for further research to unveil how ERN and Pe develops in the context of school pedagogy and raises the question of the relation between error-monitoring and reinforcement learning.

Acknowledgements

We are grateful to both Prof. Gilles Pourtois and Prof. David Sander for their insights and helpful advice. This research was supported by the National Center of Competence in Research (NCCR) Affective sciences financed by the Swiss National Science Foundation (n°51NF40-104897) and hosted by the University of Geneva, the Boninchi Foundation. MMM is supported by the Swiss National Science Foundation (n°169206), The Fondation Asile des aveulges (n°232933), and a grantor advised by Carigest SA.

Author contributions statement

S.D., J.F.K and M.M.M. conceived the experiment. S.D. conducted the experiment. S.D. and J.F.K. analyzed the data. S.D. and M.M.M. drafted the manuscript, and all authors provided essential input on subsequent versions prior to approval for submission.

Group					
	Montessori	Traditional	t or X^2	p-values FDR corrected	Cohen's d
N (girls)	13 (4)	14 (8)	1.90	0.45	
Age [years]	9.24 (2.19)	8.79 (1.72)	0.60	0.75	0.23
min, max	6.33-14.3	5.75-12.3			
SES [au]	7.25 (0.87)	7.50 (1.24)	-0.59	0.75	-0.23
Parental education style [score]	26.42 (4.54)	26.86 (4.20)	-0.26	0.91	-0.11
Fluid intelligence [score]	34.08 (2.25)	32.36 (3.27)	1.58	0.45	0.61
STAI [score]	13.38 (6.20)	13.14 (4.88)	0.11	0.91	0.04
Growth mindset [score]	12.80 (6.28)	15.60 (4.99)	-1.23	0.46	-0.49
Fixed mindset [score]	13.00 (5.44)	9.92 (3.88)	1.64	0.45	0.66

Note. Mean and SD indicated. au = arbitrary units.

Table 1. Population and control variables.

	Group				
	Montessori	Traditional	Statistics	p	Cohen's d/ η_p^2
Behavior					
Accuracy (%)	75.8 (16.0)	72.3 (23.2)	t(25)=0.46	0.65	0.18
Error rate (%)	18.1 (13.9)	15.6 (0.1)	t(25)=0.56	0.58	0.21
RT Correct (ms)	1054 (477)	1019 (412)	F(1,25)=0.009	0.93	
RT Incorrect (ms)	941 (382)	913 (352)			
ERP components (µV)					
ERN_Correct (Fpz)	129.7 (542.6)	81.0 (520.7)			
ERN_Incorrect (Fpz)	-346.8 (527.8)	-11.6 (605.6)	F(1,25)=7.19	0.013	
ΔERN	**-476 (411)**	**-92.7 (331)**	**-2.68**	**0.01**	**-1.03**
Pe_Correct (Cz)	161 (2529)	755 (3760)	F(1,25)=3.08	0.09	
Pe_Incorrect (Cz)	313 (2386)	2336 (3576)			
ΔPe	**-151 (1659)**	**-1581 (2461)**	**1.76**	**0.09**	**0.676**

Table 2. Means (SD) of behavioral performance and ERP amplitudes at electrode sites.

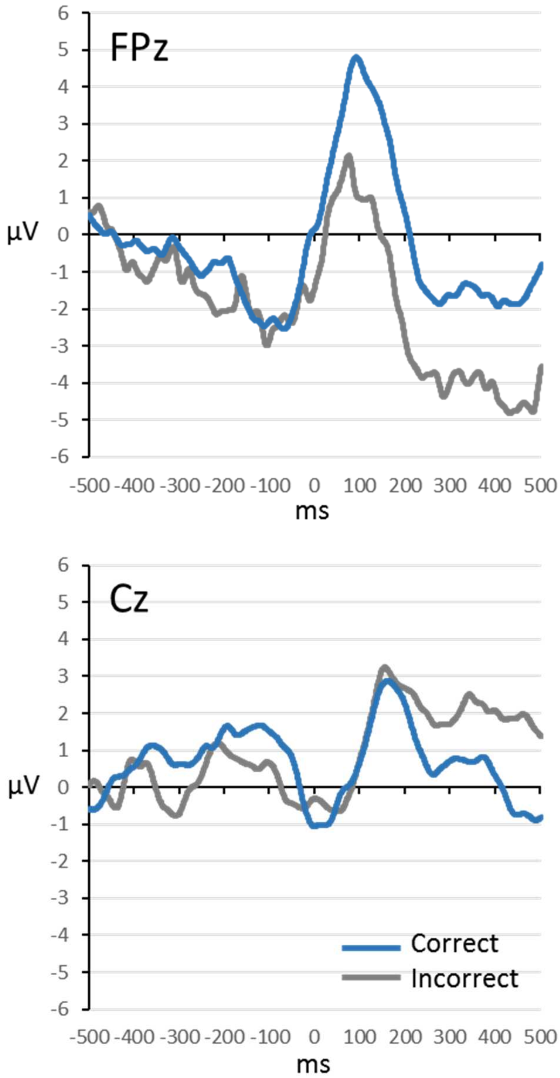


Figure 1. Mean response-locked event-related potentials at frontocentral (Fpz) and vertex (Cz) scalp locations for correct (blue) and incorrect (grey) actions, collapsed across all participants.

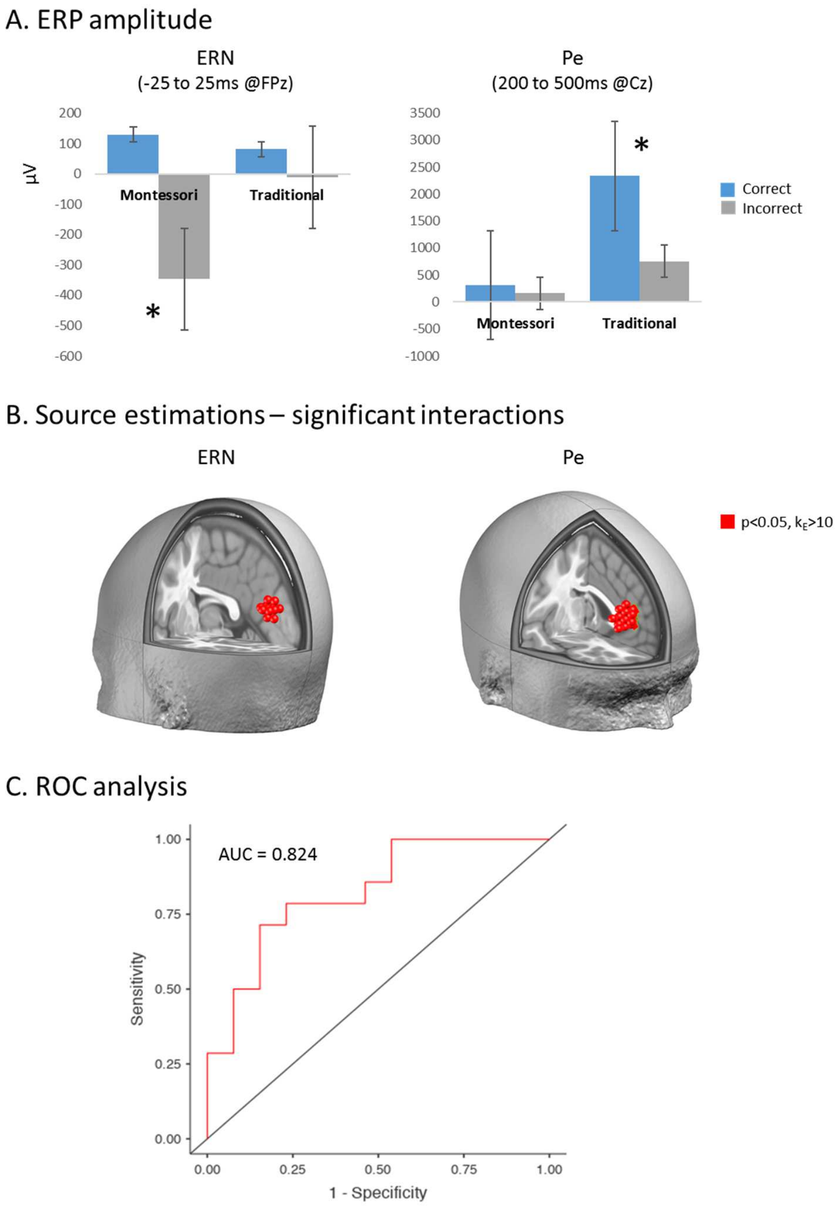


Figure 2. Electrical Neuroimaging and ROC Analyses. **A.** ANOVAs were performed on the sum under the curve for correct and error trials from the response-locked ERPs of schoolchildren in Montessori and Traditional systems at Fpz for the ERN (left) and at Cz for the Pe (left). Asterisks indicate a significant interaction. **B.** The images display the interaction term of the 2×2 (Condition × Group) mixed model repeated measures ANOVA using scalar values from the LAURA source estimations over the ERN and Pe time periods (left and right images, respectively). Results are displayed on the surface of the 7.5-13.5 y.o. Cartool brain template with red representing sources with maximum F-Values. Only solution points exhibiting a significant interaction are shown (p≤0.05 and a spatial extent of at least 10 significant points). All other points were set to zero. The maximum of activity differences for ERN and Pe were observed in the cuneus/precuneus and in the anterior cingulate cortex, respectively. **C.** Based on the combined measure of ΔERN and ΔPe, the schooling system of a child (Montessori vs. Traditional) was reliably classified using a binomial logistic regression analysis. The resulting area under the ROC curve was 0.824 and reliably discriminated between Montessori (69.2%) and Traditional schoolchildren (78.6%).

References

Aarts, K., De Houwer, J., & Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: evidence from ERPs. *Emotion, 13*(5), 960-973. doi:10.1037/a0032808

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention. *Child Dev, 78*(1), 246-263. doi:10.1111/j.1467-8624.2007.00995.x

Brooker, R. J., Buss, K. A., & Dennis, T. A. (2011). Error-monitoring brain activity is associated with affective behaviors in young children. *Dev Cogn Neurosci, 1*(2), 141-151. doi:10.1016/j.dcn.2010.12.002

Carter, C. S., Braver, T., Barch, D. M., Botvinick, M., Noll, D., & Cohen, J. D. (1998). The role of the anterior cingulate cortex in error detection and the on-line monitoring of performance: An event related fMRI study. *Biological Psychiatry, 43*, 13s-13s. doi:Doi 10.1016/S0006-3223(98)90491-7

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain, 129*(Pt 3), 564-583. doi:10.1093/brain/awl004

Elton, M., Band, G., & Falkenstein, M. (2000). To err is human. *Biological Psychology, 51*(2-3), 83-85. doi:Doi 10.1016/S0301-0511(99)00035-6

Endrass, T., Riesel, A., Kathmann, N., & Buhlmann, U. (2014). Performance monitoring in obsessive-compulsive disorder and social anxiety disorder. *J Abnorm Psychol, 123*(4), 705-714. doi:10.1037/abn0000012

Falkenstein, M. (2000). Action monitoring, evaluation, and control. *Psychophysiology, 37*, S9-S9.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of Crossmodal Divided Attention on Late Erp Components .2. Error Processing in Choice Reaction Tasks. *Electroencephalography and Clinical Neurophysiology, 78*(6), 447-455. doi:Doi 10.1016/0013-4694(91)90062-9

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron, 47*(4), 495-501. doi:10.1016/j.neuron.2005.06.020

Gehring, W. J., Himle, J., & Nisenson, L. G. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychol Sci, 11*(1), 1-6. doi:10.1111/1467-9280.00206

Grammer, J. K., Carrasco, M., Gehring, W. J., & Morrison, F. J. (2014). Age-related changes in error processing in young children: a school-based investigation. *Dev Cogn Neurosci, 9*, 93-105. doi:10.1016/j.dcn.2014.02.001

Hajcak, G., & Simons, R. F. (2002). Error-related brain activity in obsessive-compulsive undergraduates. *Psychiatry Res, 110*(1), 63-72.

Hirsh, J. B., & Inzlicht, M. (2010). Error-related negativity predicts academic performance. *Psychophysiology, 47*(1), 192-196. doi:10.1111/j.1469-8986.2009.00877.x

Kiehl, K. A., Liddle, P. F., & Hopfinger, J. B. (2000). Error processing and the rostral anterior cingulate: An event-related fMRI study. *Psychophysiology, 37*(2), 216-223. doi:Doi 10.1017/S0048577200990231

Knebel, J.-F., & Notter, M. P. (2012). STEN 1.0: Statistical Toolbox for Electrical Neuroimaging (Version 1.0). Zenodo.

Largo, R. H., Pfister, D., Molinari, L., Kundu, S., Lipp, A., & Duc, G. (1989). Significance of prenatal, perinatal and postnatal factors in the development of AGA preterm infants at five to seven years. *Dev Med Child Neurol, 31*(4), 440-456.

Lillard, A. (2013). Playful Learning and Montessori Education. *American Journal of Play, 5*(2), 157-186.

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

Lillard, A. S. (2017). *Montessori : the science behind the genius* (Third Edition. ed.). New York: Oxford University Press.

Liotti, M., Pliszka, S. R., Perez, R., Kothmann, D., & Woldorff, M. G. (2005). Abnormal brain activity related to performance monitoring and error detection in children with ADHD. *Cortex, 41*(3), 377-388.

Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-monitoring: negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *J Exp Psychol Gen, 129*(1), 43-60. doi:10.1037/0096-3445.129.1.43

Meyer, A., & Gawlowska, M. (2017). Evidence for specificity of the impact of punishment on error-related brain activity in high versus low trait anxious individuals. *Int J Psychophysiol, 120*, 157-163. doi:10.1016/j.ijpsycho.2017.08.001

Meyer, A., Weinberg, A., Klein, D. N., & Hajcak, G. (2012). The development of the error-related negativity (ERN) and its relationship with anxiety: evidence from 8 to 13 year-olds. *Dev Cogn Neurosci, 2*(1), 152-161. doi:10.1016/j.dcn.2011.09.005

Montessori, M. (1936). *The Secret of Childhood* (1981 ed.). New York: Ballantine.

Moser, J. S., Moran, T. P., Schroder, H. S., Donnellan, M. B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: a meta-analysis and conceptual framework. *Front Hum Neurosci, 7*, 466. doi:10.3389/fnhum.2013.00466

Moser, J. S., Schroder, H. S., Heeter, C., Moran, T. P., & Lee, Y. H. (2011). Mind your errors: evidence for a neural mechanism linking growth mind-set to adaptive posterror adjustments. *Psychol Sci, 22*(12), 1484-1489. doi:10.1177/0956797611419520

Murphy, P. R., van Moort, M. L., & Nieuwenhuis, S. (2016). The Pupillary Orienting Response Predicts Adaptive Behavioral Adjustment after Errors. *PLoS One, 11*(3), e0151763. doi:10.1371/journal.pone.0151763

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: evidence from an antisaccade task. *Psychophysiology, 38*(5), 752-760.

Nieuwenhuis, S., Ridderinkhof, K. R., Blow, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology, 38*(5), 752-760. doi:Doi 10.1111/1469-8986.3850752

Overbeek, T. J. M., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing - On the functional significance of the Pe Vis-a-vis the ERN/Ne. *Journal of Psychophysiology, 19*(4), 319-329. doi:10.1027/0269-8803.19.4.319

Paul, K., Walentowska, W., Bakic, J., Dondaine, T., & Pourtois, G. (2017). Modulatory effects of happy mood on performance monitoring: Insights from error-related brain potentials. *Cogn Affect Behav Neurosci, 17*(1), 106-123. doi:10.3758/s13415-016-0466-8

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Hochenberger, R., Sogo, H., . . . Lindelov, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behav Res Methods*. doi:10.3758/s13428-018-01193-y

Plessen, K. J., Allen, E. A., Eichele, H., van Wageningen, H., Hovik, M. F., Sorensen, L., . . . Eichele, T. (2016). Reduced error signalling in medication-naive children with ADHD: associations with behavioural variability and post-error adaptations. *J Psychiatry Neurosci, 41*(2), 77-87.

Pliszka, S. R., Liotti, M., & Woldorff, M. G. (2000). Inhibitory control in children with attention-deficit/hyperactivity disorder: event-related potentials identify the processing component and timing of an impaired right-frontal response-inhibition mechanism. *Biol Psychiatry, 48*(3), 238-246.

Raven, J., Raven, J. C., & Court, J. H. (2003). Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview. . San Antonio, TX: Harcourt Assessment.

Simpson, G. V., Weber, D. L., Dale, C. L., Pantazis, D., Bressler, S. L., Leahy, R. M., & Luks, T. L. (2011). Dynamic activation of frontal, parietal, and sensory regions underlying anticipatory visual spatial attention. *J Neurosci, 31*(39), 13880-13889. doi:10.1523/JNEUROSCI.1519-10.2011

Smulders, S. F., Soetens, E., & van der Molen, M. W. (2016). What happens when children encounter an error? *Brain Cogn, 104*, 34-47. doi:10.1016/j.bandc.2016.02.004

Spielberger, C. D., & Vagg, P. R. (1984). Psychometric properties of the STAI: a reply to Ramanaiah, Franzen, and Schill. *J Pers Assess, 48*(1), 95-97. doi:10.1207/s15327752jpa4801_16
Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain : 3-dimensional proportional system : an approach to cerebral imaging*. Stuttgart ; New York: Georg Thieme.
Tamnes, C. K., Walhovd, K. B., Torstveit, M., Sells, V. T., & Fjell, A. M. (2013). Performance monitoring in children and adolescents: a review of developmental changes in the error-related negativity and brain maturation. *Dev Cogn Neurosci, 6*, 1-13. doi:10.1016/j.dcn.2013.05.001
Ullsperger, M., & Danielmeier, C. (2016). Reducing Speed and Sight: How Adaptive Is Post-Error Slowing? *Neuron, 89*(3), 430-432. doi:10.1016/j.neuron.2016.01.035
Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiol Rev, 94*(1), 35-79. doi:10.1152/physrev.00041.2012
van Veen, V., & Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiol Behav, 77*(4-5), 477-482.
Velanova, K., Wheeler, M. E., & Luna, B. (2008). Maturational Changes in Anterior Cingulate and Frontoparietal Recruitment Support the Development of Error Processing and Inhibitory Control. *Cerebral Cortex, 18*(11), 2505-2522. doi:10.1093/cercor/bhn012

Is error-monitoring influenced by school pedagogy?
An fMRI study with Montessori and traditionally-schooled students

Solange Denervaud[1-3*], Eleonora Fornari[4], Xiao-Fei Yang[5], Patric Hagmann[2], Mary Helen Immordino-Yang[5†] & David Sander[1,6†*]

[1] Swiss Center for Affective Sciences (CISA), Campus Biotech, University of Geneva, 1202 Geneva, Switzerland

[2] Connectomics Lab, Department of Radiology, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), 1011 Lausanne, Switzerland

[3] The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology and Department of Clinical Neurosciences, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), 1011 Lausanne, Switzerland

[4]Biomedical Imaging Center (CIBM), Lausanne University Hospital and University of Lausanne (CHUV-UNIL), 1011 Lausanne, Switzerland

[5] Brain and Creativity Institute; Rossier School of Education; Neuroscience Graduate Program, University of Southern California, USA

[6] Laboratory for the study of Emotion Elicitation and Expression, Department of Psychology, Faculty of Psychology and Educational Sciences (FPSE), University of Geneva

†both authors contributed equally

*Corresponding author

E-mail: david.sander@unige.ch, solange.denervaud@chuv.ch

Abstract

The development of error-monitoring is central to learning and academic achievement. However, few studies exist on the neural correlates of children's error-monitoring, and no studies have examined its susceptibility to educational influences. Pedagogical methods differ on how they teach children to learn from errors. Here, 32 students (aged 8-12 years) from high quality Swiss traditional or Montessori schools performed a math task with feedback

during fMRI. Although the groups' accuracies were similar, Montessori students skipped fewer trials, responded faster and showed more neural activity in right parietal and frontal regions involved in math processing. While traditionally-schooled students showed greater functional connectivity between the ACC, involved in error-monitoring, and hippocampus following correct trials, Montessori students showed greater functional connectivity between the ACC and frontal regions following incorrect trials. The findings suggest that pedagogical experience influences the development of error-monitoring and its neural correlates, with implications for neurodevelopment and education.

Introduction

Given the changing landscape of work and the ease of acquiring factual information via technology (*New Vision for Education; Unlocking the Potential of Technology*, 2015), there is an active debate around how pedagogical approaches can support students not simply in memorizing facts and becoming proficient at procedures but in developing abilities for evaluating their ongoing learning processes (Soderstrom & Bjork, 2015); in essence, for learning how to learn efficiently. Central to this enterprise is fostering a self-directed, process-oriented approach to learning, in which children learn to recognize and utilize information about incorrect responses to iteratively improve their skills (Benedek et al., 2014; Melby-Lervag, Redick, & Hulme, 2016; Oudeyer, Gottlieb, & Lopes, 2016; Redick, Shipstead, Wiemers, Melby-Lervag, & Hulme, 2015).

Error-monitoring refers to the intrinsic ability to detect and evaluate outcomes that violate expectation and to adapt in response (Ullsperger, Danielmeier, & Jocham, 2014). Existing work suggests that error-monitoring shares processing features with surprise or violation of expectations and serves as a basic orienting mechanism for subsequent behavioral adaptation and learning (Wessel, Danielmeier, Morton, & Ullsperger, 2012; Danielmeier, Eichele, Forstmann, Tittgemeyer, & Ullsperger, 2011; Laubach, Caetano, & Narayanan, 2015). Individuals' error-monitoring competencies are tightly related to self-regulatory and flexible goal-directed behaviors (Posner, Rothbart, Sheese, & Tang, 2007; Ullsperger, Danielmeier, & Jocham, 2014). Along with executive functions, error-monitoring improves across development, with adult-like responses by mid-adolescence (Buzzell et al., 2017; Grammer, Carrasco, Gehring, & Morrison, 2014; Tamnes, Walhovd, Torstveit, Sells, & Fjell, 2013). Its developmental trajectory is known to depend upon underpinning brain networks, most notably involving the cingulate gyrus, including the anterior cingulate cortex (ACC), and developmental changes in both behavior and brain activity have been described (Smulders, Soetens, & van der Molen, 2016; Tamnes et al., 2013; Velanova, Wheeler, & Luna, 2008; Posner, Rothbart, Sheese, & Tang, 2007). Most notably, these include increased capacity to detect errors quickly and self-correct accurately, and corresponding shifts in functional connectivity of the ACC.

It is possible that error-monitoring and its neural correlates are shaped by developmental experience, and in particular by schooling. Pedagogical traditions differ in the ways they support children in recognizing and reacting to incorrect responses. While traditional pedagogy typically teaches children by providing them information about when they have made a mistake, feedback that is often delayed (Dihoff, Brosvic, & Epstein, 2003; Epstein, Epstein, & Brosvic, 2001), so that they can avoid such mistakes in future (Metcalfe, 2017), Montessori students are typically not given direct information about the correctness of their answers. Instead, Montessori teachers encourage children to notice their own incorrect thinking or to help peers identify incorrect thinking in a pro-social manner (Montessori, 1936). This method was built from the thinking of Maria Montessori in the 1920's based on her systematic observations of children's development in educational settings and in interactions (Montessori, 1936). The global aim was to facilitate children actively organizing their own understanding in a socially cooperative context in which children of mixed ages observe and sometimes teach each other (Montessori, 1936). This could implicitly teach Montessori-schooled children to engage with errors and self-correction in a more autonomous, process-oriented and constructive way, while also helping them leverage social skills. This is in contrast to methods that focus children through testing on memory and recall (Karpicke & Roediger, 2007; Marsh, Roediger, Bjork, & Bjork, 2007), which are emphasized in traditional schooling.

Montessori students have been reported to achieve higher scores on academic tasks (Lillard & Else-Quest, 2006; Lillard et al., 2017; Rathunde, 2001), on tests of socio-emotional skills (Rathunde, 2001; Rathunde & Csikszentmihalyi, 2005) and on creativity tests (Besançon & Lubart, 2008; Denervaud, Knebel, Hagmann, & Gentaz, 2019). These outcomes are thought to reflect children's experiences with the pedagogical strategies (Lillard, 2012), but the cognitive origins of these effects have not been studied. Given the focus on independent recognition of errors in Montessori pedagogy, one reasonable hypothesis is that Montessori schooling may be effective in part because it impacts children's development of error-monitoring.

Our previous work supports the hypothesis that error-monitoring may differ in children exposed to Montessori versus traditional pedagogy (Denervaud, Knebel, Immordino-Yang, & Hagmann, 2020). Compared with Montessori students, traditionally-schooled students were found to react more strongly when detecting an incorrect response, as measured by strength of global field power in EEG, suggesting lower self-regulated error-monitoring ability (Denervaud et al., *in prep.*). These first results suggest that pedagogical practice influences the way young students learn to perceive and respond to errors, with traditional teaching methods potentially teaching children to strive to remember and produce only correct responses (Denervaud et al., *in revision*). These results align well with research on adults demonstrating that incorrect responses are typically affectively tagged as negative and aversive (Aarts, De Houwer, & Pourtois, 2012, 2013). Of note, the adults participating in the existing studies had likely almost exclusively been traditionally-schooled.

Human studies consistently implicate cingulate regions in general, and the ACC in particular, in the process of error-monitoring. This is consistent with functional connectivity and brain activation studies on neurodevelopmental maturation of the cingulate gyrus. These studies report an age-related caudal-ventral gradient of developmental change (Kelly et al., 2009), that is evident with regard to error-monitoring (Velanova, Wheeler, & Luna, 2008), especially between 8-12 years of age. However, to date, functional neuroimaging studies of error-monitoring have mainly been conducted with adults. These studies shed light on the spatio-temporal specificity of error-monitoring. They make clear its distinction from conflict responses and reward processes (Iannaccone et al., 2015), and highlight its role not only in behavioral change but in social adaptation (Bediou et al., 2012; Danielmeier et al., 2011; Izuma et al., 2015; Klucharev, Munneke, Smidts, & Fernandez, 2011). That is, error-monitoring processes are not only invoked for our own mistakes, but also for mistakes that we monitor in others (de Bruijn, de Lange, von Cramon, & Ullsperger, 2009; Kang, Hirsh, & Chasteen, 2010). Competitive or cooperative social settings have been shown to differently influence error perception and vicarious learning, such that cooperative settings heighten error-monitoring responses to others' mistakes, and increase subsequent learning (Koban, Pourtois, Vocat, & Vuilleumier, 2010).

Here, we asked 8-12 year old students from Swiss traditional and Montessori schools to judge whether solutions to straightforward math problems were right or wrong during fMRI scanning, and studied their brain activation patterns when their responses were correct or incorrect using blood oxygen-level dependent (BOLD) activation and functional connectivity analyses. Public traditional and private Montessori schools in Switzerland both provide high quality education, but differ systematically in their pedagogical approaches (according to local educational policies and the Montessori application standards). To control as best as possible for selection bias, we collected information on demographic factors, families' reported educational practices and beliefs, as well as on fluid intelligence and math ability, and found students in the two pedagogical groups to be comparable (Table 1).

We hypothesized that both groups of children would show neural responses to their self-generated incorrect *versus* correct responses, and that this neural activity and connectivity would differ in children from these two types of schools. Specifically, we hypothesized that: 1) regardless of pedagogical experience, students would show increased activity along the cingulate gyrus to incorrect versus correct answers, consistent with data from adults; and 2) Montessori and traditionally-schooled students would show relatively different patterns of brain activation and connectivity in trials corresponding to correct and incorrect responses; Montessori students would show higher brain activation and connectivity in brain regions implied in error-monitoring (ACC, medial frontal cortex), whereas traditionally-schooled students would show effects in brain regions involved in memory (e.g., hippocampus).

Materials and Methods

Participants. Thirty-seven healthy children (18 females; aged 8-12.3 years, mean ± SD = 9.95 ± 1.25) completed the experiment as part of a larger study including neuroimaging and behavioral measures aiming at evaluating the impact of school environment on cognitive and emotional development of error monitoring. Selection criteria were age (8 to 12 years of age) and school enrollment (participants had to be enrolled in a Montessori or in a traditional school system from the Kindergarten on, or for at least 3 years in the case of the youngest children). All but one participant in each group were right-handed. One Montessori participant stopped the task half-way due to sickness in the scanner, and four others were excluded due to dental braces interference on the fMRI scan (n = 1 Traditional student), high dyslexia and dyscalculia (n = 1 Montessori student), or motion greater than 3 mm exceeding a rate of 20% of the slices collected (n = 2, one from each group), leaving 32 subjects (17 female; aged 8-12.3 years, mean ± SD = 9.98 ± 1.25) available for analyses (half from each schooling system). This study was approved by the local ethics committee (CER-Vaud). Written informed consent to take part in the study was obtained from parents and oral consent from subjects; participants acknowledged that they were free to withdraw at any time without penalty. Participants were compensated with a voucher and received the personalized gift that had been displayed during the fMRI task.

Group variables.

Data used to evaluate between-group homogeneity were:

(i) non-verbal intelligence (black and white version of the Progressive Matrices; Raven et al., 1998); the child had to choose from amongst 6 items one pattern that would fit within a matrix. There were 36 matrices to be completed and each correct answer granted a point (maximum 36).

(ii) self-reported anxiety (STAI-Y2; Spielberger & Vagg, 1984); on a 3-point Likert scale, the child responded to questions about their state anxiety. Responses were summed (scores ranging from 0 to 40; higher scores denote higher anxiety).

(iii) working memory (Digit-Letter span tasks; Wechsler, 2005); the child listened to and memorized a string of mixed digit-letters and repeated them in ascending order (that is, they mentally reorganized the information). The score was age-standardized (higher scores denote greater working memory capacity).

(iv) mathematical skills (Simonart, 2008); the child solved standardized math problems including arithmetic, logic and geometric paper-based tasks. The maximum score was 100% correct answers.

Finally, to control as best as possible for the possibility of selection bias stemming from recruiting Montessori students from private schools and traditional students from public schools, to ensure home environments were similar, and to ensure equivalence on math anxiety (which is known to be impacted by parents; Maloney, Ramirez, Gunderson, Levine, & Beilock, 2015) we measured :

(i) family's socio-economic status, including both parents' education levels and current job (higher scores denote higher socio-economic status).

(ii) parental report of child math affect (parents' reports of their child's level of math anxiety and math enjoyment) using a 5-point Likert scale (higher scores denote more positive affect toward math).

(iii) home physical environment, using a questionnaire about whether there is a yard, the number of rooms, etc. (higher scores denote more enriched home environment).

(iv) parents' perceived life stress (higher scores denote higher perceived life stress);

(v) home pedagogical environment, including questions about parents' interest in education and pedagogy (e.g., how many books on education they have at home), and style of parenting (e.g., number of meals shared with the child per week on average, frequency of museum visits together, type of feedback given when the child succeeds); (higher scores denote increased knowledge about pedagogy and more parental involvement in their child's intellectual development).

Control variables were partially collected online (parental questionnaires; children's fluid intelligence and self-reported anxiety) and through a behavioral assessment that took place after fMRI scanning.

A chi-square test was performed to determine whether the gender ratios differed between Montessori and traditional students. In addition, multiple t-tests (independent or Welch's according to the preliminary data check with Q-Q plots and Levene's test) with a 95% confidence interval (CI) for the mean difference were run on the control variables to test for significant differences between the groups (Montessori vs traditional), with a false-rate discovery (FDR) *p*-value correction at q=0.05.

Task and procedure

Students were individually assessed on a novel math proofreading fMRI paradigm that was designed to evoke a school-related task. During scanning, participants were asked to respond as to whether the solution of the math problem they viewed was right or wrong using a response-box in their right hand. Instructions emphasized the need for both speed and accuracy. Each trial consisted of (i) a start cue displayed for 1000 ms, followed by (ii) a simple addition or subtraction problem with a suggested solution that could be correct or incorrect (retrieved from a standardized age-normalized task, presented in random order; von Aster, 2005) displayed for 3000 ms, during which the participant had to respond, (iii) the feedback (words "correct" or "incorrect") displayed for 2000 ms, and (iv) a fixation cross as inter-trial jitter lasting between 2000-3000 ms, in steps of 500ms, varying randomly to provide adequate temporal sampling of the blood oxygen level dependent (BOLD) response (Figure 1). In total, 64 trials were divided evenly into eight blocks with an inter-block interval of 14000 ms.

Stimulus delivery and recording of behavioral data (reaction time and accuracy) were controlled by E-prime and a serial response box (www.pstnet.com; Psychology Software Tools). Button presses occurring more than 3000 ms after stimulus the presentation of the math problem were labeled as a miss, and were excluded from neuroimaging analysis.

The data were collected as part of a larger study examining the differential effects of various rewards versus no reward or no feedback on correct versus incorrect responses. Because of the short experiment duration necessary for child participants, we were unable to include sufficient trials within each reward condition to test condition-specific reward effects on error-monitoring. The current study therefore focused exclusively on the comparison between "correct *versus* incorrect" trials. To avoid confounding with reward, reward images were presented concurrent with the feedback for both correct and incorrect trials.

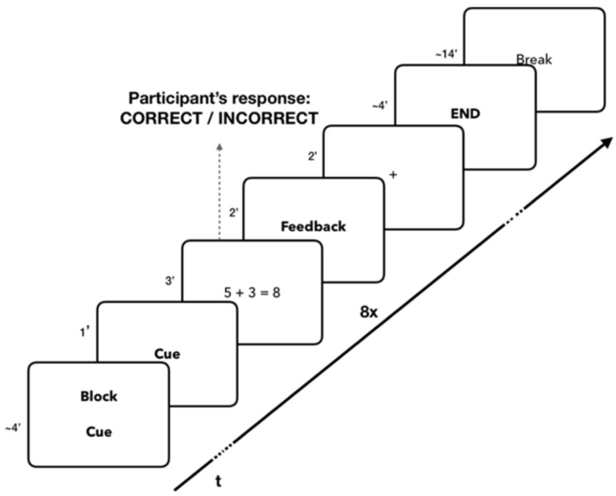


Figure 1. The fMRI task design: after the cue presentation (e.g., "If you perform well, you will gain points"), the participant had to determine whether the math problem and its suggested solution (e.g. "3+10 = 12") was right or wrong. Feedback was given, based on his/her real performance.

Data acquisition

Structural and functional images were collected at the Lemanic Biomedical Imaging Center (CIBM) of the University Hospital Lausanne (CHUV), on a Siemens 3T Prisma-Fit

MR scanner, with a 64-channel head-coil. For each participant, a 3-dimensional high-resolution isotropic T1-weighted sequence (MPRAGE) was acquired (TR = 2000ms, TE = 2.47 ms, 208 slices; voxel size= 1x1x1, flip angle=8°) as anatomical individual reference and basis for surface reconstruction. The functional scans were continuously acquired using a standard echo-planar gradient echo sequence acquired by simultaneous multislice (SMS) imaging technique and covering the whole brain with an isotropic voxel size of 2mm ([TR] = 1000 ms; echo time [TE] = 30 ms; 64 axial slices; slice thickness = 2 mm, no gap between slices, flip angle = 80°, matrix size = 100 x 100, field of view [FOV] = 200 mm, sms factor=4, parallel imaging acceleration factor=2). For each subject one session of 740 volumes was recorded, including seven "dummy" scans that were then discarded by the scanner, for a total acquisition time of 12m26s. Foam pads were placed around the subject's head inside the coil to prevent head's motion.

Behavioral analysis

Behavioral data were analyzed using the statistical R software jamovi (Jamovi Project, 2018). First, to validate the fMRI task, Pearson correlation coefficients were computed to assess the relationship between correct response scores and the participant's standardized math task performance and parental report of affect toward math. Second, main effects and their interaction on accuracy were analyzed using a 3-by-2 ANOVA (response type -correct, incorrect, missed-, as within-subject factor; Montessori *versus* traditional as between- subject factor). Main effects and their interaction on response time were analyzed using a 2-by-2 ANOVA (response type -correct, incorrect-, as within-subject factor; Montessori *versus* traditional as between-subject factor), with $\alpha<0.05$. Post-hoc Tukey tests were computed when relevant. Finally, we computed participants' efficiency as their reaction time divided by their proportion of correct responses (Bruyer & Brysbaert, 2011) and an independent t-test was used to statistically evaluate group differences (Montessori *versus* traditional students).

Neural activation analyses

Imaging data processing and analyses were carried out with Matlab (Mathworks, Natick, MA, USA Version 7.13) using the software SPM12 (Wellcome Department of Cognitive Neurology, London, UK) and the results were visualized using xjview Toolbox for SPM (http://www.alivelearn.net/xjview) and MRIcroGL (http://www.cabiatl.com/mricrogl/). Anatomical locations were labeled and described with the help of the aal atlas.

Preprocessing. Functional images were motion corrected with reference to the first scan, using a 6-parameter rigid-body realignment. Then, slice timing correction was performed on these realigned images. The functional images were then co-registered to the high-resolution T1 anatomical image of the participant, using mutual information. Finally, images were normalized (by estimation based on the anatomical images and then applied to the functional images) to the MNI template and spatially smoothed using an 8 mm Gaussian filter. Visual inspection of estimated motion parameters was conducted on a subject-by-subject basis, and subjects demonstrating a rate of motion-corrupted scans (> 3 mm, >3°) exceeding 20% were excluded (n=2).

Brain activation analysis. First-level statistics were performed for each subject using a general linear model as implemented in SPM12. Brain activity of interest comprised the 4-sec stimulus presentation that followed the start cue, including task and feedback time. Contrasts of the participant's correct vs. incorrect responses were computed. The realignment parameters were included in the model as a nuisance variable, and the highpass filter cut-off was set to 128 sec. The generated maps were then used as input values for the group-level analysis. Second-level random effects were analyzed using general mixed-design ANOVA including the factors *Response* (participant's correct or incorrect responses) as within-subjects factor, and *Pedagogy* (Montessori, traditional) as between-subjects factor. All activation maps were thresholded at $p<0.05$ for cluster level FDR correction, which correspond to a voxel-wise threshold of $p<0.001$ and a cluster size threshold of greater then 30 voxels per cluster.

Functional connectivity analysis. To investigate effects of response type and pedagogical group on functional connectivity, selected seed regions of interest (ROIs) identified in the activation analysis were used as seeds in a Psycho-Physiological Interaction (PPI) analysis implemented in SPM12. ROIs in the middle prefrontal cortex (x=-30, y=28, z=44) and ventral anterior precuneus (x=-10, y=-46, z=46) were identified from the main effect of correct versus incorrect responses; ROIs in the right prefrontal cortex (x=6, y=64, z=8), anterior middle cingulate (x=-2, y=52, z=-2) and cuneus (x=22, y=-90, z=8) were identified from the main effect of pedagogical group. Seeds were defined as 8mm-radius spheres. The average connectivity maps were computed for each subject by response type (correct and incorrect). We conducted a second-level group analysis using a t-test for each response type (correct versus incorrect), or pedagogical group (Montessori versus traditional). As the connectivity analyses were exploratory, the threshold was set at $p<0.001$ uncorrected at the voxel-wise level.

Results

Group variables

No significant differences were found ($p>0.3$) for age, fluid intelligence, trait anxiety, working memory, mathematical competency, and affect toward math, revealing comparable groups on these measures; parents' SES, parenting style, perceived life stress and pedagogical approach at home did not differ between the groups (Table 1). There were marginally more girls in the traditionally-schooled group (*p*=0.08). We thus examined the effect of gender and founded to be not significant (*p*=0.18).

	Group				
	M	T	X^2 or t-test	*p*-values FDR corrected	Cohen's *d*
N (girls)	16 (6)	16 (11)	3.14	0.08	
Age [years]	10.1 (1.24)	9.90 (1.29)	0.36	0.78	0.13
min, max	8.34-12.2	8.00-12.3			
Non-verbal Intelligence [score]	34.7 (1.35)	32.7 (3.24)	2.28	0.30	0.81
Self-report Anxiety [score]	10.1 (5.17)	13.5 (5.18)	-1.88	0.30	-0.66
Working Memory [score]	11.3 (2.77)	10.3 (1.44)	1.28	0.53	0.45
Mathematical Skills [score]	56.1 (16.4)	51.6 (20.0)	0.69	0.69	0.24
Family SES [score]	7.19 (0.70)	6.50 (1.40)	1.75	0.30	0.62
Parent-report Math Affect [au]	34.4 (16.7)	30.2 (15.5)	0.73	0.69	0.26
Home Phys. Environment [au]	25.63 (4.56)	26.50 (3.61)	-0.60	0.69	-0.21
Parents' life-stress [%]	45.81 (23.41)	54.50 (19.95)	-1.13	0.54	-0.40
Home Ped. Environment [au]	7.13 (1.31)	7.25 (1.24)	-0.28	0.78	-0.10

Note. Mean and SD. Au = arbitrary unit. "Phys.": Physical; "Ped.": Pedagogical.

Table 1 Demographic and control variables for the Montessori (M) and traditionally-schooled (T) groups.

Behavioral analysis

All participants performed at above 60% accuracy and no ceiling effect was observed, suggesting that the fMRI math task's difficulty level was calibrated well for this age. Validating the fMRI experimental task, as shown in Figure 2, performance was positively correlated with performance on the standardized math task completed outside the scanner (r = 0.56, n = 32, p <0.001) and with parental report of the child's math affect (r = 0.38, n = 32, p = 0.038).

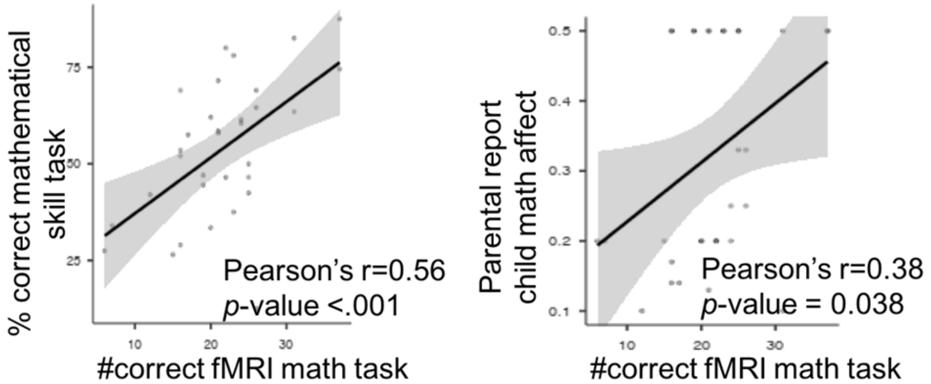


Figure 2. Participants' performance on the fMRI math task correlated with their mathematical skills score (standardized task; *left*) and with parental report of child's math affect (*right*).

As expected, participant's *correct* responses were more frequent than *incorrect* or *missed* (no response) responses, $F(60,2) = 22.88$, $p<0.001$, $\eta_p^2=0.43$. There was an interaction by group in response patterns, $F(60,2) = 3.78$, $p=0.028$, $\eta_p^2=0.11$. While the groups did not differ on rate of correct responses, $t(60)=-0.03$, $p_{tukey}=1.0$, Montessori participants had higher *incorrect* rates, while traditionally-schooled students *missed* more trials, $t(60)=3.05$, $p_{tukey}=.038$ (see Figure 3, *left*).

We checked whether *right* and *wrong* math problems (in the stimulus presentation) interacted with students' *correct* and *incorrect* responses and found no interaction (no differences in response patterns when presented with *right* and *wrong* math problems, $F(30,1) =1.46$, $p=0.237$, $\eta_p^2=0.05$). There was also no interaction by pedagogical group ($F(60,2) =0.23$, $p=0.638$, $\eta_p^2=0.001$). However, we did not have sufficient trials to power a test for condition-specific (right/wrong problem presented) response effects at the neural activation level.

Reaction time (RT) was on average faster in Montessori students ($M_{RT\_M}=1719$ ms, SD=405) than in traditionally-schooled students ($M_{RT\_T}=2060$, SD=351), $F(30,1)=6.55$, $p=0.016$, $\eta_p^2=0.18$. RT did not differ according to the *response type* (correct, incorrect),

$F(30,1)$ =0.94, p=0.34, η_p^2=0.03, and no interaction between *response type* and pedagogical group was found, $F(30,1)$=2.06, p=0.161, η_p^2=0.06 (see Figure 3, *center*).

Overall, *efficiency* (RT/Accuracy) was higher in Montessori students than in traditionally-schooled students; M_M=3.26, SD=1.31; M_T=2.46, SD=0.81; t(30)=2.09, p=0.046, Cohen's d=0.74 (see Figure 3, *right*).

Figure 3. Group averages for behavioral results; response-rate, reaction time and efficiency (computed as the ratio of reaction time to percentage correct responses). Error bars represent SEM.

Neural activation analyses

The two-way mixed-design ANOVA with *response* (student is correct or incorrect) as within-subject factor and *pedagogy* (Montessori, traditional) as between-subjects factor revealed main effects of both factors, but no interaction. First, relative to incorrect responses, correct responses elicited higher brain activity (Figure 4 *top panel*, table 2A) in the bilateral posterior cingulate cortex and left precuneus, as well as within the middle frontal gyrus, the left inferior temporal gyrus, and both left 1^{st} and 2^{nd} crus of the cerebellum. Second, regardless of whether the response was correct or incorrect, relative to the traditionally-schooled students, the Montessori students showed increased activation of: the left occipital cortex and right cuneus (Figure 5 *top panel*, and Table 2B), and contralateral regions at trend-level; the right superior parietal lobule; and the medial prefrontal cortex (with some activation extending into the left anterior cingulate cortex).

MNI coordinate of local peak in (x,y,z)	Cluster extent according to aal automatic labeling (aal)	Z-value	Cluster extent (in mm^3)
A) Main effect of *Response* *– voxel threshold at p<0.001 and cluster threshold at p<0.05 (FDR)*			
-10, -46, 44	Left precuneus	4.76, p<.000	1306
-2, -42, 32	Left PCC		
4, -48, 30	Right PCC		
-30, 28, 44	Left MFC	4.43, p<.000	569
-44, -78, -30	Left crus1 cerebellum	4.17, p<.000	693
-4, -80, -22	Left crus2 cerebellum		
-44, -46, -14	Left inferior temporal cortex	4.09, p=0.002	372
B) Main effect of *Pedagogy*			
20, -92, 10	Right cuneus	4.41, p=0.006	300
-28, -96, 10	Left middle occipital	4.31, p<.000	703
30, -56, 66	Right superior parietal	4.09, p=0.010	270
6, 64, 6	Right. MFC	4.08, p=0.019	234
-2. 52, -2	Left ACC		
C) Interaction term			
	No		

Table 2. Cluster-level information corresponding to the ANOVA fMRI activation results shown in Figures 4 and 5, top panels. PCC: posterior cingulate cortex; MFC: medial frontal cortex; ACC: anterior cingulate cortex.

We excluded the *miss* condition from the current analysis as a skipped response could be due to a range of situations, from fatigue to delayed button press to loss of concentration or uncertainty about the answer. Future research should differentiate brain activity underlying trials “missed” for these various reasons, especially given the group difference in frequency of missed responses.

Functional connectivity analyses

The functional connectivity analyses revealed that connectivity differed depending on whether students responded correctly or incorrectly. Independent of pedagogical group, the precuneus and the mid-frontal cortex seeds were more strongly connected to the left insula for *correct* responses compared to *incorrect* responses (Figure 4 *bottom panel*, table 3A). Connectivity patterns in *correct* compared to *incorrect* responses also differed by pedagogical group. For *correct* responses, the traditional pedagogy group showed stronger connectivity (higher correlation) between each seed (the left anterior cingulate cortex, the right medial prefrontal cortex, and the right cuneus) and the right hippocampus. Traditionally-schooled students also showed stronger connectivity during *correct* trials between the right medial prefrontal cortex seed and the right putamen. There were no regions that showed greater connectivity for *correct* responses for the Montessori group. For *incorrect* responses, Montessori students showed greater connectivity between the left anterior cingulate cortex seed and the right middle and superior frontal regions, and the left orbitofrontal cortex (Figure 5 *bottom panel*, table 3B). There were no regions that showed greater connectivity for *incorrect* responses for the traditionally-schooled group.

SEEDS	Condition	MNI coordinate of peak in (x,y,z)	Peak according to aal automatic labeling (aal)	T-value
A) Functional connectivity from the SEEDs selected from main effect of *correct vs incorrect response* *– voxel threshold at p<0.001 and peak-level threshold at p<0.001 (uncorr.)*				
Left MFC	**Correct > Incorrect**	-38, -16, 16	**Left insula**	3.77, $p_{\text{unc.}}<.001$
	Incorrect > Correct			
Left precuneus	**Correct > Incorrect**	-38, -18, 16	**Left insula**	3.72, $p_{\text{unc.}}<.001$
	Incorrect > Correct			
B) Functional connectivity from the SEEDs selected from main effect of *pedagogical group*				
Left ACC	M>T correct responses			
	T>M correct responses	24, -38, 4	**Right hippocampus**	4.44, $p_{\text{unc.}}<.001$
	M>T incorrect responses	34, 60, 28	**Right MFC**	3.70, $p_{\text{unc.}}<.001$
		14, 68, 26	**Right SFC**	3.17, $p_{\text{unc.}}=.001$
		-44, 42, -8	**Left inf. OFC**	3.26, $p_{\text{unc.}}=.001$
	T>M incorrect responses			
Right MFC	M>T correct responses			
	T>M correct responses	24, -38, 4	**Right hippocampus**	4.53, $p_{\text{unc.}}<.001$
		26, -6, 10	**Right putamen**	4.04, $p_{\text{unc.}}<.001$
	M>T incorrect responses			
	T>M incorrect responses			
Right cuneus	M>T correct responses			
	T>M correct responses	24, -38, 4	**Right hippocampus**	4.62, $p_{\text{unc.}}<.001$
	M>T incorrect responses			
	T>M incorrect responses			

Table 3. Peak-level information corresponding to the functional connectivity (PPI) results shown in Figures 4 and 5. MFC: medial frontal cortex; ACC: anterior cingulate cortex; SFC: superior frontal cortex; OFC: orbitofrontal cortex.

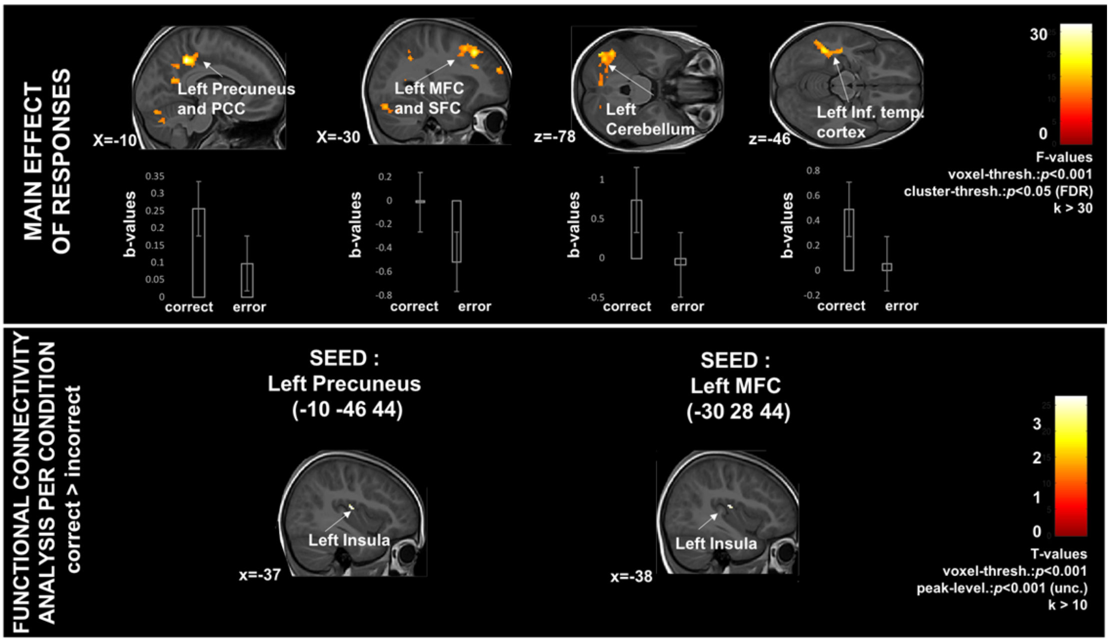


Figure 4. Relative to erroneous responses, correct responses elicited higher brain activity (*top panel*) and functional connectivity (*bottom panel*) across pedagogical groups. PCC: posterior cingulate cortex; MFC: medial frontal cortex; SFC: superior frontal cortex.

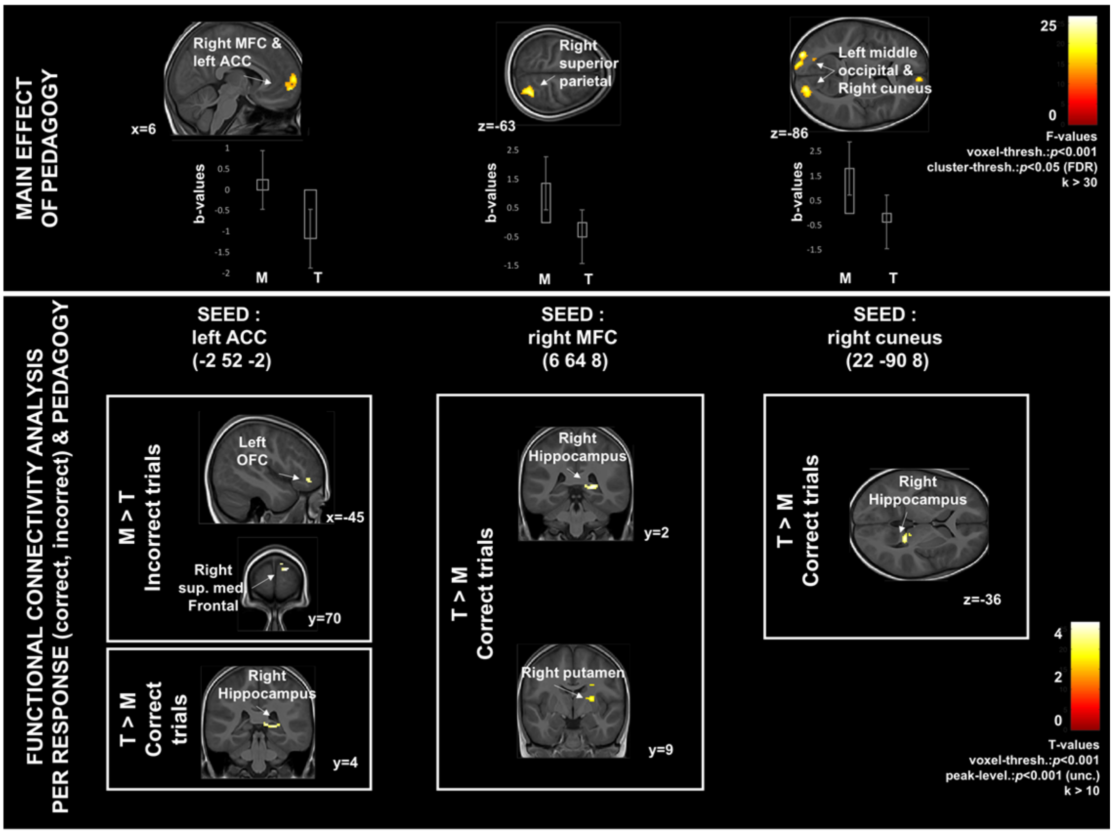


Figure 5. *Top panel*: Effect of pedagogy on neural activations during the math task. The relative activation was higher in Montessori students (M) compared to traditionally-schooled students (T). *Bottom panel*: Functional connectivity analyses by pedagogical group for trials with *correct* versus *incorrect* responses. Montessori students (M) showed stronger connectivity (higher correlation) with seed regions for *incorrect* responses, while traditionally-schooled students (T) showed stronger connectivity (higher correlation) for *correct* responses. MFC: medial frontal cortex; ACC: anterior cingulate cortex; OFC: orbitofrontal cortex.

Discussion

Though error-monitoring is fundamental to learning and develops across childhood and adolescence, to our knowledge, few studies have examined its neural correlates in children (Buzzell et al., 2017; Grammer et al., 2014; Hammerer, Muller, & Li, 2014) and no study has examined its susceptibility to pedagogical approaches. Accordingly, the goal of this study was to identify brain activity and functional connectivity during math-task error-monitoring in 8-12 year old schoolchildren from Montessori and traditional schools. Interestingly, we found that the schoolchildren in our samples showed greater neural activity for their correct than for their incorrect responses in various regions, independent of pedagogical experience. This was found in several regions, including regions implicated in the default mode and executive networks. This runs against previous studies with adults (Botvinick, Cohen, & Carter, 2004; Bush, Luu, & Posner, 2000; Carter et al., 1998; Danielmeier et al., 2011; Magno, Foxe, Molholm, Robertson, & Garavan, 2006), which report greater activity for incorrect responses.

We also found that pedagogical exposure was associated with the behavioral and neural correlates of error-monitoring in our samples. Even with similar levels of math competency to traditionally-schooled students, and similar proportions of correct answers in our task, we found that Montessori students reacted more quickly during the task and made more incorrect responses but missed fewer trials compared to traditionally-schooled students. Across conditions, Montessori participants showed higher brain activity than traditionally-schooled students in regions implicated in visual and math processing, as well as in regions related to attentional/executive control.

Most interestingly, the groups' functional connectivity patterns following their correct and incorrect responses differed. We used the neural activity contrast of correct versus incorrect participants' responses to identify seed regions in the ACC, the cuneus cortex and the right superior medial frontal cortex. These regions are interesting for our research question because they have been shown to be involved in self-monitoring (Jones, Rothbart, & Posner, 2003; Posner, Rothbart, Sheese, & Tang, 2007; Velanova, Wheeler, & Luna, 2008). Montessori students showed stronger connectivity (higher correlation) between these seed regions and the ventromedial prefrontal cortex in trials in which they had made errors. By contrast, traditionally-schooled students showed stronger connections between the seed regions and the hippocampus on trials in which they had answered correctly. Montessori students did not show significant changes in connectivity following correct responses, and traditionally-schooled students did not show significant changes following incorrect responses.

Higher brain activation for correct responses. Both the Montessori and traditionally-schooled groups showed greater neural activation during correct responses relative to incorrect responses in the precuneus, PCC, MFC, inferior temporal cortex and cerebellum. This is the opposite pattern than has generally been observed in adults (Ullsperger, Danielmeier, & Jocham, 2014), and deserves additional experimental attention in future work on the development of error-monitoring in children. One possible interpretation comes from the observation that our findings align with earlier work showing a stronger impact of positive than negative feedback on learning in late childhood (van Duijvenvoorde, Zanolie, Rombouts, Raijmakers, & Crone, 2008). This difference between adults' and children's responses to feedback has been proposed to reflect a change in the information that children find salient for learning, rather than a change in their perception of the affective value of the feedback (van den Bos, Guroglu, van den Bulk, Rombouts, & Crone, 2009; Denervaud, Knebel, Immordino-Yang, & Hagmann, 2019). Though further work is needed, it is possible that our neural finding corresponds to a developmental difference between children and adults around the saliency and utility of correct responses. It could be that prior to reaching a level of competence in math that allows for reliable error prediction, children may rely more heavily on associative learning and on integrating into their knowledge schemas the procedures that led them to correct information. This would make correct responses more relevant for learning and therefore more neurologically salient. At the same time, an experience-dependent shift in subsequent processing, given the known network plasticity of error-monitoring in children, could still be possible and would be reflected in pedagogical group differences in connectivity. A heavier reliance on integration of procedures leading to correct responses would also be consistent with the increased PCC and cerebellar activations that we observed to correct responses in both pedagogical groups. PCC is a highly anatomically connected hub central to the default mode network (Hagmann et al., 2008), and is involved in the switch from internally to externally directed attention (Yang, Pavarini, Schnall, & Immordino-Yang, 2018; Immordino-Yang, Christodoulou, & Singh, 2012), as well as in forming integrated memories (Leech & Sharp, 2014; Ryali et al., 2016). The activated sectors of cerebellum are involved in many cognitive functions involving associative and procedural learning (Keren-Happuch, Shen-Hsing, Moon-Ho, & Desmond, 2014).

Our findings open the question of whether these developmental processes are specific to math learning, or reflect the development of more general attentional and learning mechanisms. Here, the increase in activity in the ventral anterior sector of the precuneus (Cavanna & Trimble, 2006; Immordino-Yang, Christodoulou, & Singh, 2012), left middle frontal gyrus (Heitzeg et al., 2014) and Crus 1 and Crus 2 of the cerebellum (Keren-Happuch, Shen-Hsing, Moon-Ho, & Desmond, 2014; Marvel & Desmond, 2010; Schmahmann, 2010), could reflect greater executive control and external focus when responding correctly, and integration of executive control with regions known to be involved in mathematical cognition, including inferior temporal gyrus and the lateral frontal area (Amalric & Dehaene, 2017). Whether the student would be correct as a consequence of his or her cognitive control and engagement of mathematical processing regions, or whether being correct would elicit higher engagement, is a topic for future work.

Pedagogical influence. The finding that Montessori students missed fewer trials and had more incorrect trials could reflect the emphasis on exploratory learning in Montessori classrooms (Livstrom, Szostkowski, & Roehrig, 2019; Marshall, 2017). The extent of exploratory learning through trial-and-error is known to depend on the structure of the environment, the task complexity and the instructions given; these features together have been shown to impact self-directed executive functions and curiosity among children (Barker et al., 2014; Baranes, Oudeyer, & Gottlieb, 2014; Gottlieb, Lopes, & Oudeyer, 2014; Oudeyer, 2017). This explanation would also be consistent with the fact that, in our study, Montessori students' showed stronger neural activation during math processing in bilateral occipital and parietal cortices, involved in multisensory integration (Corbetta & Shulman, 2002; Vossel, Geng, & Fink, 2014), in the right inferior parietal lobule, known to be recruited for math processing (Arsalidou, Pawliw-Levac, Sadeghi, & Pascual-Leone, 2018), and showed increased connectivity during incorrect trials with frontal areas. One testable hypothesis for future work is that these patterns of results in Montessori students reflect a more exploratory, self-corrective and multisensory approach to mathematical cognitive processes. Conversely, it may be that the traditionally-schooled children's increased functional connectivity between the ACC and the hippocampus reflects a strategic inclination to either memorize or recall correct answers, consistent with instrumental learning (Brovelli, Nazarian, Meunier, & Boussaoud, 2011; Vogel & Schwabe, 2016) and/or reinforcement learning (Ballard, Wagner, & McClure, 2019), with less reliance on self-direction and self-monitoring of errors. Together, these results suggest that daily pedagogical experience may have important implications for learning, behavior and related mindsets (i.e., being more oriented towards *processes* versus *outcomes;* Dweck, 1999) that should be further explored.

Limitations. First, our study compared groups that were not randomly assigned to either Montessori or Traditional type of education, making it possible that, despite our efforts to control for relevant variables, the observed effects were also driven by other factors (e.g., family-related or motivation-related) than the pedagogical ones. Second, our study has a modest sample size and a cross-sectional design. While our results suggest that error-monitoring is modulated by pedagogical experience, further longitudinal and larger-scale studies will need to investigate the extent to which pedagogy contributes to the emergence of robust psychological and neural error processing. Such studies would help to further probe the role of context in the development of error-monitoring for both Montessori and traditionally-schooled students, and could inform pedagogical and policy-related decisions that aim to foster process-oriented learning behaviors (Slavin, 2019; Darling-Hammond, 2016; Seghier, Fahim, & Habak, 2019). Another limitation of our study is that it did not include adult participants so we could not test whether brain patterns to correct and incorrect responses in our specific task are development-related or persist in adulthood. To confirm brain activation differences are development-related, it would be of interest to have adults perform the same experiment (with track of their pedagogical history).

Conclusion. Our findings suggest that 8-12 year old students may process correct and incorrect responses differently than do adults, and that they may attend especially to correct responses. Our findings also suggest that pedagogical experience in school modulates error-monitoring behavior, and its underlying brain activity and connectivity. Together, these findings call for further research testing whether error-monitoring competencies and corresponding brain networks indeed undergo a shift with age that is modulated by pedagogical experience.

Data availability

Archives of behavioral data used in this study, including the e-Prime stimulus files can be accessed in Zenodo 10.5281/zenodo.3773305.

Competing interests

The authors declare no competing interests.

Contributions

Designed research: S.D., E.F., and D.S.; Performed research: S.D., E.F., and P.H.; Analyzed data: S.D., E.F., X-F. Y., and MH. I-Y. All authors contributed to the interpretation of the results and writing the paper.

Acknowledgement

We are grateful to the children and parents who enrolled in this study. We thank Remy Furrer for insights on the task design, Virginie Sterpenich and Michael Notter for insights on data analyses, and Claire Batilotti for helping with behavioral data collection.

This research was financially supported by the National Center of Competence in Research in (NCCR) Affective sciences; by the Swiss National Science Foundation (n°51NF40-104897); and by Prof. Micah Murray, who supported the ethics review. SD was funded by the Department of Radiology of Lausanne University Hospital (CHUV), and by the Boninchi and the Fondation Eden in Geneva. Imaging was supported in part by the Centre d'Imagerie Biomédicale (CIBM) of the Université de Lausanne (UNIL), Université de Genève (UNIGE), Hôpitaux Universitaires de Genève (HUG), Centre Hospitalier Universitaire Vaudois (CHUV), Ecole Polytechnique Fédérale de Lausanne (EPFL), and the Leenaards and Jeantet Foundations. MHIY and XFY were supported by a grant from the Templeton Foundation to MHIY (n°53-4400-0464). XFY was also funded by an Army Research Laboratory-West Interdisciplinary Collaboration Grant (n°53-4400-1158).

References

Aarts, K., De Houwer, J., & Pourtois, G. (2012). Evidence for the automatic evaluation of self-generated actions. *Cognition, 124*(2), 117-127. doi:10.1016/j.cognition.2012.05.009

Aarts, K., De Houwer, J., & Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: evidence from ERPs. *Emotion, 13*(5), 960-973. doi:10.1037/a0032808

Amalric, M., & Dehaene, S. (2017). Cortical circuits for mathematical knowledge: evidence for a major subdivision within the brain's semantic networks. *Philos Trans R Soc Lond B Biol Sci, 373*(1740). doi:10.1098/rstb.2016.0515

Arsalidou, M., Pawliw-Levac, M., Sadeghi, M., & Pascual-Leone, J. (2018). Brain areas associated with numbers and calculations in children: Meta-analyses of fMRI studies. *Dev Cogn Neurosci, 30*, 239-250. doi:10.1016/j.dcn.2017.08.002

Ballard, I. C., Wagner, A. D., & McClure, S. M. (2019). Hippocampal pattern separation supports reinforcement learning. *Nat Commun, 10*(1), 1073. doi:10.1038/s41467-019-08998-1

Baranes, A. F., Oudeyer, P. Y., & Gottlieb, J. (2014). The effects of task difficulty, novelty and the size of the search space on intrinsically motivated exploration. *Front Neurosci, 8*, 317. doi:10.3389/fnins.2014.00317

Barker, J. E., Semenov, A. D., Michaelson, L., Provan, L. S., Snyder, H. R., & Munakata, Y. (2014). Less-structured time in children's daily lives predicts self-directed executive functioning. *Front Psychol, 5*, 593. doi:10.3389/fpsyg.2014.00593

Bediou, B., Koban, L., Rosset, S., Pourtois, G., & Sander, D. (2012). Delayed monitoring of accuracy errors compared to commission errors in ACC. NeuroImage, 60, 1925-1936.

Benedek, M., Jauk, E., Fink, A., Koschutnig, K., Reishofer, G., Ebner, F., & Neubauer, A. C. (2014). To create or to recall? Neural mechanisms underlying the generation of creative new ideas. *Neuroimage, 88*, 125-133. doi:10.1016/j.neuroimage.2013.11.021

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

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences, 8*(12), 539-546. doi:10.1016/j.tics.2004.10.003

Brovelli, A., Nazarian, B., Meunier, M., & Boussaoud, D. (2011). Differential roles of caudate nucleus and putamen during instrumental learning. *Neuroimage, 57*(4), 1580-1590. doi:10.1016/j.neuroimage.2011.05.059

Bruyer, R., & Brysbaert, M. (2011). Combining Speed and Accuracy in Cognitive Psychology: Is the Inverse Efficiency Score (Ies) a Better Dependent Variable Than the Mean Reaction Time (Rt) and the Percentage of Errors (Pe)? *Psychologica Belgica, 51*(1), 5-13. doi:DOI 10.5334/pb-51-1-5

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4*(6), 215-222.

Buzzell, G. A., Richards, J. E., White, L. K., Barker, T. V., Pine, D. S., & Fox, N. A. (2017). Development of the error-monitoring system from ages 9-35: Unique insight provided by MRI-constrained source localization of EEG. *Neuroimage, 157*, 13-26. doi:10.1016/j.neuroimage.2017.05.045

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*(5364), 747-749. doi:DOI 10.1126/science.280.5364.747

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain, 129*(Pt 3), 564-583. doi:10.1093/brain/awl004
Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci, 3*(3), 201-215. doi:10.1038/nrn755
Danielmeier, C., Eichele, T., Forstmann, B. U., Tittgemeyer, M., & Ullsperger, M. (2011). Posterior medial frontal cortex activity predicts post-error adaptations in task-related visual and motor areas. *J Neurosci, 31*(5), 1780-1789. doi:10.1523/JNEUROSCI.4299-10.2011
Darling-Hammond, L. (2016). Research on Teaching and Teacher Education and Its Influences on Policy and Practice. *Educational Researcher, 45*(2), 83-91. doi:DOI: 10.3102/0013189X16639597
de Bruijn, E. R., de Lange, F. P., von Cramon, D. Y., & Ullsperger, M. (2009). When errors are rewarding. *J Neurosci, 29*(39), 12183-12186. doi:10.1523/JNEUROSCI.1751-09.2009
Denervaud, S., Hess, A., Sander, D., & Pourtois, G. (*in revision*). Children's automatic evaluation of self-generated actions is different from adults.
Denervaud, S., Knebel, J. F., Hagmann, P., & Gentaz, E. (2019). Beyond executive functions, creativity skills benefit academic outcomes: Insights from Montessori education. *PLoS One, 14*(11), e0225319. doi:10.1371/journal.pone.0225319
Denervaud, S., Knebel, J. F., Mullier, E., Hagmann, P., & Murray, M.M. (*in prep*). Error-monitoring is influenced by school pedagogy: ERN and Pe study.
Dihoff, R. E., Brosvic, G. M., & Epstein, M. L. (2003). The Role of Feedback During Academic Testing: The Delay Retention Effect Revisited. . *The Psychological Record, 53*, 533-548. doi:https://doi.org/10.1007/BF03395451
Dweck, C. S. (1999). *Self-theories : their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.
Epstein, M. L., Epstein, B. B., & Brosvic, G. M. (2001). Immediate feedback during academic testing. *Psychol Rep, 88*(3 Pt 1), 889-894. doi:10.2466/pr0.2001.88.3.889
Gottlieb, J., Lopes, M., & Oudeyer, P. Y. (2014). Active learning and decision making: an introduction to the collection. *F1000Res, 3*, 276. doi:10.12688/f1000research.5757.2
Grammer, J. K., Carrasco, M., Gehring, W. J., & Morrison, F. J. (2014). Age-related changes in error processing in young children: a school-based investigation. *Dev Cogn Neurosci, 9*, 93-105. doi:10.1016/j.dcn.2014.02.001
Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biol, 6*(7), e159. doi:10.1371/journal.pbio.0060159
Hammerer, D., Muller, V., & Li, S. C. (2014). Performance monitoring across the lifespan: still maturing post-conflict regulation in children and declining task-set monitoring in older adults. *Neurosci Biobehav Rev, 46 Pt 1*, 105-123. doi:10.1016/j.neubiorev.2014.06.008
Heitzeg, M. M., Nigg, J. T., Hardee, J. E., Soules, M., Steinberg, D., Zubieta, J. K., & Zucker, R. A. (2014). Left middle frontal gyrus response to inhibitory errors in children prospectively predicts early problem substance use. *Drug Alcohol Depend, 141*, 51-57. doi:10.1016/j.drugalcdep.2014.05.002
Iannaccone, R., Hauser, T. U., Staempfli, P., Walitza, S., Brandeis, D., & Brem, S. (2015). Conflict monitoring and error processing: new insights from simultaneous EEG-fMRI. *Neuroimage, 105*, 395-407. doi:10.1016/j.neuroimage.2014.10.028

Immordino-Yang, M. H., Christodoulou, J. A., & Singh, V. (2012). Rest Is Not Idleness: Implications of the Brain's Default Mode for Human Development and Education. *Perspect Psychol Sci, 7*(4), 352-364. doi:10.1177/1745691612447308

Izuma, K., Akula, S., Murayama, K., Wu, D. A., Iacoboni, M., & Adolphs, R. (2015). A causal role for posterior medial frontal cortex in choice-induced preference change. *J Neurosci, 35*(8), 3598-3606. doi:10.1523/JNEUROSCI.4591-14.2015

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science, 6*(5), 498-504. doi:Doi 10.1111/1467-7687.00307

Kang, S. K., Hirsh, J. B., & Chasteen, A. L. (2010). Your mistakes are mine: Self-other overlap predicts neural response to observed errors. *Journal of Experimental Social Psychology, 46*(1), 229-232. doi:10.1016/j.jesp.2009.09.012

Karpicke, J. D., & Roediger, H. L., 3rd. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *J Exp Psychol Learn Mem Cogn, 33*(4), 704-719. doi:10.1037/0278-7393.33.4.704

Kelly, A. M., Di Martino, A., Uddin, L. Q., Shehzad, Z., Gee, D. G., Reiss, P. T., . . . Milham, M. P. (2009). Development of anterior cingulate functional connectivity from late childhood to early adulthood. *Cereb Cortex, 19*(3), 640-657. doi:10.1093/cercor/bhn117

Keren-Happuch, E., Shen-Hsing, A., Moon-Ho, R. H., & Desmond, J. E. (2014). A meta-analysis of cerebellar contributions to higher cognition from PET and fMRI studies. *Hum Brain Mapp, 35*(2), 593-615. doi:10.1002/hbm.22194

Klucharev, V., Munneke, M. A., Smidts, A., & Fernandez, G. (2011). Downregulation of the posterior medial frontal cortex prevents social conformity. *J Neurosci, 31*(33), 11934-11940. doi:10.1523/JNEUROSCI.1869-11.2011

Koban, L., Pourtois, G., Vocat, R., & Vuilleumier, P. (2010). When your errors make me lose or win: event-related potentials to observed errors of cooperators and competitors. *Soc Neurosci, 5*(4), 360-374. doi:10.1080/17470911003651547

Laubach, M., Caetano, M. S., & Narayanan, N. S. (2015). Mistakes were made: neural mechanisms for the adaptive control of action initiation by the medial prefrontal cortex. *J Physiol Paris, 109*(1-3), 104-117. doi:10.1016/j.jphysparis.2014.12.001

Lavin, C., Melis, C., Mikulan, E., Gelormini, C., Huepe, D., & Ibanez, A. (2013). The anterior cingulate cortex: an integrative hub for human socially-driven interactions. *Front Neurosci, 7*, 64. doi:10.3389/fnins.2013.00064

Leech, R., & Sharp, D. J. (2014). The role of the posterior cingulate cortex in cognition and disease. *Brain, 137*, 12-32. doi:10.1093/brain/awt162

Lillard, A. S. (2012). Preschool children's development in classic Montessori, supplemented Montessori, and conventional programs. *J Sch Psychol, 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*(5795), 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. *Front Psychol, 8*, 1783. doi:10.3389/fpsyg.2017.01783

Livstrom, I. C., Szostkowski, A. H., & Roehrig, G. H. (2019). Integrated STEM in practice: Learning from Montessori philosophies and practices. *School Science and Mathematics, 119*(4), 190-202. doi:10.1111/ssm.12331

Magno, E., Foxe, J. J., Molholm, S., Robertson, I. H., & Garavan, H. (2006). The anterior cingulate and error avoidance. *Journal of Neuroscience, 26*(18), 4769-4773. doi:10.1523/Jneurosci.0369-06.2006

Maloney, E. A., Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2015). Intergenerational Effects of Parents' Math Anxiety on Children's Math Achievement and Anxiety. *Psychol Sci, 26*(9), 1480-1488. doi:10.1177/0956797615592630

Marsh, E. J., Roediger, H. L., 3rd, Bjork, R. A., & Bjork, E. L. (2007). The memorial consequences of multiple-choice testing. *Psychon Bull Rev, 14*(2), 194-199. doi:10.3758/bf03194051

Marshall, C. (2017). Montessori education: a review of the evidence base. *npj Science of Learning, 2*(1). doi:10.1038/s41539-017-0012-7

Marvel, C. L., & Desmond, J. E. (2010). Functional topography of the cerebellum in verbal working memory. *Neuropsychology Review, 20*(3), 271-279. doi:10.1007/s11065-010-9137-7

Melby-Lervag, M., Redick, T. S., & Hulme, C. (2016). Working Memory Training Does Not Improve Performance on Measures of Intelligence or Other Measures of "Far Transfer": Evidence From a Meta-Analytic Review. *Perspect Psychol Sci, 11*(4), 512-534. doi:10.1177/1745691616635612

Montessori, M. (1936). *The Secret of Childhood* (1981 ed.). New York: Ballantine.

New Vision for Education; Unlocking the Potential of Technology. (2015). Retrieved from www.weforum.org

Oudeyer, P. Y. (2017). Autonomous development and learning in artificial intelligence and robotics: Scaling up deep learning to human-like learning. *Behav Brain Sci, 40*, e275. doi:10.1017/S0140525X17000243

Oudeyer, P. Y., Gottlieb, J., & Lopes, M. (2016). Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. *Changing Brains Applying Brain Plasticity to Advance and Recover Human Ability, 229*, 257-284. doi:10.1016/bs.pbr.2016.05.005

Posner, M. I., Rothbart, M. K., Sheese, B. E., & Tang, Y. (2007). The anterior cingulate gyrus and the mechanism of self-regulation. *Cogn Affect Behav Neurosci, 7*(4), 391-395.

Powers Iii, A. R., Hillock-Dunn, A., & Wallace, M. T. (2016). Generalization of multisensory perceptual learning. *Sci Rep, 6*, 23374. doi:10.1038/srep23374

Rathunde, K. (2001). Montessori Education and optimal experieunce: a framework for new research. *The NAMTA Journal, 26*(1), 11-43.

Rathunde, K., & Csikszentmihalyi, M. (2005). Middle School Students' Motivation and Quality of Experience: A Comparison of Montessori and Traditional School Environments. *American Journal of Education, 111*(3).

Redick, T. S., Shipstead, Z., Wiemers, E. A., Melby-Lervag, M., & Hulme, C. (2015). What's working in working memory training? An educational perspective. *Educ Psychol Rev, 27*(4), 617-633. doi:10.1007/s10648-015-9314-6

Ryali, S., Supekar, K., Chen, T., Kochalka, J., Cai, W., Nicholas, J., . . . Menon, V. (2016). Temporal Dynamics and Developmental Maturation of Salience, Default and Central-Executive Network Interactions Revealed by Variational Bayes Hidden Markov Modeling. *PLoS Comput Biol, 12*(12), e1005138. doi:10.1371/journal.pcbi.1005138

Schmahmann, J. D. (2010). The role of the cerebellum in cognition and emotion: personal reflections since 1982 on the dysmetria of thought hypothesis, and its historical evolution from theory to therapy. *Neuropsychology Review, 20*(3), 236-260. doi:10.1007/s11065-010-9142-x

Seghier, M. L., Fahim, M. A., & Habak, C. (2019). Educational fMRI: From the Lab to the Classroom. *Front Psychol, 10*, 2769. doi:10.3389/fpsyg.2019.02769

Simonart, G. (2008). *ECHAS; échelle d'apprentissages scolaires primaires* (E. éditions Ed.).

Slavin, R. E. (2019). How evidence-based reform will transform research and practice in education. *Educational Psychologist*. doi:https://doi.org/10.1080/00461520.2019.1611432

Smulders, S. F., Soetens, E., & van der Molen, M. W. (2016). What happens when children encounter an error? *Brain Cogn, 104*, 34-47. doi:10.1016/j.bandc.2016.02.004

Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: an integrative review. *Perspect Psychol Sci, 10*(2), 176-199. doi:10.1177/1745691615569000

Spielberger, C. D., & Vagg, P. R. (1984). Psychometric properties of the STAI: a reply to Ramanaiah, Franzen, and Schill. *J Pers Assess, 48*(1), 95-97. doi:10.1207/s15327752jpa4801_16

Tamnes, C. K., Walhovd, K. B., Torstveit, M., Sells, V. T., & Fjell, A. M. (2013). Performance monitoring in children and adolescents: a review of developmental changes in the error-related negativity and brain maturation. *Dev Cogn Neurosci, 6*, 1-13. doi:10.1016/j.dcn.2013.05.001

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiol Rev, 94*(1), 35-79. doi:10.1152/physrev.00041.2012

van den Bos, W., Guroglu, B., van den Bulk, B. G., Rombouts, S. A. R. B., & Crone, E. A. (2009). Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. *Frontiers in Human Neuroscience, 3*. doi:ARTN 52 10.3389/neuro.09.052.2009

van Duijvenvoorde, A. C., Zanolie, K., Rombouts, S. A., Raijmakers, M. E., & Crone, E. A. (2008). Evaluating the negative or valuing the positive? Neural mechanisms supporting feedback-based learning across development. *J Neurosci, 28*(38), 9495-9503. doi:10.1523/JNEUROSCI.1485-08.2008

Velanova, K., Wheeler, M. E., & Luna, B. (2008). Maturational Changes in Anterior Cingulate and Frontoparietal Recruitment Support the Development of Error Processing and Inhibitory Control. *Cerebral Cortex, 18*(11), 2505-2522. doi:10.1093/cercor/bhn012

Vogel, S., & Schwabe, L. (2016). Learning and memory under stress: implications for the classroom. *NPJ Sci Learn, 1*, 16011. doi:10.1038/npjscilearn.2016.11

von Aster, M. (2005). *ZAREKI-R Batterie pour l'évaluation du traitement des nombres et du calcul chez l'enfant.* (G. Dellatolas, Trans. Pearson Ed.): ECPA.

Vossel, S., Geng, J. J., & Fink, G. R. (2014). Dorsal and ventral attention systems: distinct neural circuits but collaborative roles. *Neuroscientist, 20*(2), 150-159. doi:10.1177/1073858413494269

Wechsler, D. (2005). *WISC-IV Echelle d'intelligence de Wechsler pour enfants et adolescents* (Pearson Ed. 4th ed.): ECPA.

Yang, X. F., Pavarini, G., Schnall, S., & Immordino-Yang, M. H. (2018). Looking up to virtue: averting gaze facilitates moral construals via posteromedial activations. *Soc Cogn Affect Neurosci, 13*(11), 1131-1139. doi:10.1093/scan/nsy081

1

Children's automatic evaluation of self-generated actions is different from adults

Solange Denervaud[1,2], Adrien Hess[3], David Sander[1,3] & Gilles Pourtois[4]

[1]The Center for Affective Sciences, Campus Biotech, University of Geneva, 1202 Geneva, Switzerland

[2]Connectomics Lab, Department of Radiology, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), 1011 Lausanne, Switzerland

[3]Faculty of Psychology and Educational Sciences (FAPSE), University of Geneva, 1204 Geneva, Switzerland

[4] Cap Lab, Department of Experimental Clinical and Health Psychology, Ghent University, 9000 Gent, Belgium

E-mails: gilles.pourtois@ugent.be, david.sander@unige.ch

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ACKNOWLEDGEMENT

We are grateful to Marie Vetterli and Claire Batilotti for help with data collection and Lien De Saedeleer for early help with coding. This research was supported by the National Center of Competence in Research (NCCR) Affective sciences financed by the Swiss National Science Foundation (n°51NF40-104897) and hosted by the University of Geneva. GP is supported by a research grant (n°3G024716) from the FWO (Research Foundation - Flanders).

Children's automatic evaluation of self-generated actions is different from adults

Running head: CHILDREN'S AUTOMATIC EVALUATION OF ACTIONS

Research highlights

- Response errors led to post error slowing in both children and adults.
- Response errors were associated with negative affect in adults only.
- In children, correct responses were related to positive affect.
- Children experiencing the traditional or the Montessori pedagogy showed different action monitoring effects.

Abstract

Performance monitoring (PM) is central to learning and decision making. It allows to swiftly detect deviations between actions and intentions, such as response errors, and adapt behavior accordingly. Previous research in adult participants showed that error monitoring is associated with two distinct and robust behavioral effects. First, a systematic slowing down of reaction time speed is usually observed following error commission, and known as post error slowing (PES). Second, response errors are automatically evaluated as negative events "on the fly". However, it remains unclear whether children (i) process response errors as adult do, (ii) also evaluate them as negative events, and (iii) the specific school environment they encounter could influence error monitoring. To address these questions, we adapted a simple decision making task validated previously in adults, and allowing to measure PES as well as the affective processing of response errors. We recruited 8-12 years old schoolchildren enrolled either in traditional (N=56) or Montessori (N=45) schools, and compared them to adults (N=46) on the exact same task. Results showed that children processed correct actions as positive events whereas adults processed errors as negative events. Moreover, this former effect was observed in traditional schoolchildren only, but not in Montessori schoolchildren. In comparison, PES was not modulated by age, nor pedagogy. These findings suggest that unlike PES that likely reflects an automatic attention orienting towards response errors, their affective processing depends on both age and school environment in 8-12 years children.

Keywords: performance monitoring development – action monitoring – evaluative priming – post-error slowing – school environment – Montessori pedagogy

Central to learning and decision making stands the remarkable ability to rapidly evaluate the outcome of our actions as good or bad, and adapt our behavior accordingly. In adults, response errors provide a unique window into performance monitoring (PM), which is tightly related to self-regulation (Inzlicht, Bartholow, & Hirsh, 2015), as well as value-based decision making (Ullsperger, Danielmeier, & Jocham, 2014). In comparison, research on PM in children is limited. More specifically, whether or not children can process response errors as adults, remains an open question. This gap in the literature is surprising given that PM is a building block of self-regulation and decision making. The main goal of our study was to address this question.

Post-Error Slowing

In adults, the cognitive architecture underlying PM has been conceived as a feedback loop that monitors possible deviations between action and goal, and this way assigns value to it. Based on this evaluation, remedial processes can subsequently take place (Ullsperger, Danielmeier, & Jocham, 2014; Ullsperger, Fischer, Nigbur, & Endrass, 2014). At the behavioral level, these remedial processes can be explored using Post-Error Slowing (PES) (Rabbitt, 1966). PES translates the systematic slowing-down in reaction time (RT) speed for trials following response errors versus correct responses. Although the PES has long been conceived as adaptive (i.e. increasing the likelihood of post-error accuracy and/or reflecting enhanced cognitive control; see also Botvinick, Braver, Barch, Carter, & Cohen, 2001), recent models and data (see Ullsperger & Danielmeier, 2016 for a review) have challenged this view, suggesting that the PES could also probably reflect unspecific attention processes to some degree, including an automatic orienting response to deviant events, and bearing in mind that response errors are usually "oddball" in the trial series (Notebaert et al., 2009). According to this view (see also Wessel, 2018), the PES reflects a blend of both adaptive and unspecific adjustment effects following error commission.

In comparison, research on PES in children is scant. Accordingly, it remains unclear whether they automatically orient to them (like adults do too). Earlier work already showed that the PES could be found in children as young as 3 years old (Jones, Rothbart, & Posner, 2003), suggesting an early onset in life for it, in agreement with the view that it is likely subtended by an exogenous attention control system that can operate and maturate rapidly after birth (Colombo, 2001). Given this evidence, it is likely that older children (e.g. 8-12 years old), very much like adults, could exhibit PES (see also Smulders, Soetens, & van der Molen, 2016). The first goal of our study was to address this question.

Errors are negative

Besides the behavioral adaptation following errors (i.e., PES), these worse-than-expected events are also usually associated with a distinctive affective processing. More specifically, accumulating evidence shows that response errors are perceived as negative events compared to correct responses in adults (Pourtois et al., 2010; Koban & Pourtois, 2014; see also Dignath, Eder, Steinhauser, & Kiesel, 2019). Moreover, this evaluation is deemed rapid and automatic (in young adults - undergraduate students; see Aarts, De Houwer, & Pourtois, 2012). Using a simple priming methodology, it has been shown that young adults categorize negative words faster and better after response errors than positive words, suggesting that this evaluation of response errors as negative events could be traced at the behavioral level too (see also Aarts, De Houwer, & Pourtois, 2013 and De Saedeleer & Pourtois, 2016 for replications). Interestingly, the reverse effect (i.e. assigning a positive value to correct responses) was much weaker in these earlier studies, suggesting an asymmetry in the affective processing of self-generated actions in young adult participants. Further, this evaluative effect did not correlate with PES, suggesting that the processing of response errors as aversive is unrelated to the automatic orienting towards these deviant events in young adults. Presumably, in analogy with PES, the affective processing of response errors as negative events could also be deemed adaptive since it might serve to quickly identify them, and in turn foster error-based learning, with the goal to protect the organism from their possible bad or deleterious consequences. While it can be adaptive to perceive error as negative events, inhibition or mistakes retention, on the other hand, undermine the learning rate (Thompson-Schill, Ramscar, & Chrysikou, 2009).

However, whether or not young children automatically assign a negative value to their response errors, like adult do (Aarts et al., 2012), remains an open question. Previous research showed that toddlers express complex emotions such as shame or anger when failing to reach a goal, suggesting that they can assign negative value to breakdowns in self-efficacy. More generally, they usually show a negative bias whereby “bad” is stronger than “good” when it comes to stimulus or outcome evaluation (Vaish, Grossmann, & Woodward, 2008). Accordingly, one could conjecture that response errors are probably processed as negative events in toddlers. However, toddlers’ behavior is usually characterized by active exploration, and guided by trials and errors, which suggests indirectly that response errors are not necessarily associated with negative value during childhood, and unlike what has been found in young adult participants (Aarts et al., 2012). For children, response errors, conflicts or challenges usually correspond to valuable learning opportunities that allow them to acquire and transform knowledge (see e.g., Gopnik & Wellman, 2012). These distinctive events allow them to adjust and update the mental representations that form and structure newly acquired information (Fischer & Rose, 1996; Montessori, 1936; Piaget, 1952; Vygotsky, 1978). Interestingly, the minimization of error probability is thought to underlie and drive cognitive development (Oudeyer & Smith, 2016). Moreover, children actually allocate attention preferentially towards surprising events, such as novel stimuli, and they exhibit an intrinsic motivation, or curiosity, to learn from them.This tendency to lower error prediction is observed in the self-organization of language acquisition (Moulin-Frier, Nguyen, & Oudeyer, 2013), or when they engage in traditional games for example (Pellegrini, Dupuis, & Smith, 2007; Lillard, 2017). Hence, whereas children undoubtedly can detect and react to events that violate or challenge their expectations, it remains unclear whether response errors are automatically processed as negative events by them or not. The second goal of our study was to assess the affective processing of response errors in children, and compare it to adults.

Influence of School Pedagogy

As a matter of fact, during childhood, exploration and learning are strongly influenced by the environment in which they take place. Therein, the specific pedagogy encountered by the children in their school is an important determinant of how exploration and learning develop and manifest (Kaplan & Patrick, 2016; Kang et al., 2009; Oudeyer, Gottlieb, & Lopes, 2016), and it might therefore also influence their "natural" processing of response errors as more or less negative, or even as positive events. In many Western developed countries, a traditional pedagogy is often used (PISA; Grisay, de Jong, Gebhardt, Berezner, & Halleux-Monseur, 2007). This pedagogy entails that learning progresses are evaluated through formal assessments, typically with the use of grades or other forms of evaluative feedback, such as rewards or punishments. In contrast, the Montessori pedagogy, which is less frequently used and encountered in these countries, offers an alternative approach, where learning and development are promoted without the use of these incentives and reinforcers (Lillard & Else-Quest, 2006; Lillard, 2012; Marshall, 2017; Montessori, 1936; Rathunde, 2001). More specifically, through independent or peer-to-peer exploration in the absence of evaluative feedback (emanating from the teacher), learning is facilitated, and self-efficacy is eventually stimulated (Denervaud, Knebel, Hagmann, & Gentaz, 2019; Denervaud, Knebel, Immordino-Yang, & Hagmann, 2020; Lillard et al., 2017). Accordingly, it is conceivable that the specific pedagogy encountered by the children may have a modulatory effect on the way they process response errors as distinctive affective events, and orient to them (as expressed by PES). Presumably, the Montessori pedagogy might have a different impact on the affective processing of response errors than the traditional pedagogy, even though in both cases, a PES could be found. The last goal of our study was to put to the test this hypothesis.

To this end, in this study, we used and adapted the experimental procedure previously devised and validated by Aarts et al. (2012) and De Saedeleer & Pourtois (2016) in young adults. More specifically, we asked 8-12 years old schoolchildren (experiencing either the traditional or Montessori pedagogy) and young adults to perform the same and simple speeded Go/noGo task. Given the strict response deadline imposed, participants committed now and then response errors. Importantly, after each trial of the Go/noGo task, participants had to categorize as fast as possible an emotional word shown on the screen as being either positive or negative (second task). We borrowed the logic of evaluative priming (Jones, Olson, & Fazio, 2010), and used this second task to probe at the behavioral level the affective processing of response errors (first task) by the participants. More specifically, we assessed if emotional word categorization was globally delayed following responses errors compared to hits (suggesting PES), as well as whether negative words were processed faster than positive words following response errors, selectively (suggesting evaluative priming of response errors as negative events). In light of the literature reviewed here above, we hypothesized that PES should be observed in young adults as well as children. Moreover, we surmised that in young adults, response errors would be processed as negative events, thereby replicating Aarts et al., (2012). In children, we explored if a similar evaluative processing of errors could be found (see also Vaish, Grossmann, & Woodward, 2008), and whether it could be influenced by the pedagogy encountered by the children at school, here with a focus on the direct comparison between Montessori and traditional pedagogy."

Material and Method

Ethics

The experiment was conducted in accordance with the Declaration of Helsinki. Written parental consent was obtained for each child, and informed consent was provided by each adult participant.

Participants

Hundred-and-ten schoolchildren participated in the experiment. Selection criteria were age (8 to 13 y.o.) and schooling system. Children with missing data (n=2, Montessori schoolchildren), or outside target age (n=7) were excluded from the study (N=9). In total, the data from 101 children (M_{age}=10.4, SD=1.1) were included in the analysis. Forty-five of them were enrolled in the Montessori schooling system (M_{age}=10.3, SD=1.2, 17 girls), and 56 in the traditional one (M_{age}=10.5, SD=1.1, 29 girls). In addition, 55 adult participants took part in the study either in exchange of course credits (28 undergraduate psychology students), or were compensated 15CHF (27 recruited outside University). Adults who did not commit errors in all conditions and hence had missing data, were removed (N=9). Following this criterion, 46 subjects were included in the analysis (M_{age}=28.0, SD=9.4, 30 women).

Demographic and socio-economic variables

For children, we collected age, gender, fluid intelligence (Raven, Raven, & Court, 2003) and information about the socio-economic background (Genoud, 2002) to assess whether the two groups were comparable on these variables or not. For adults, we only collected age and gender.

Evaluative Priming Task

Participants performed an adapted version of the speeded Go/noGo task from Vocat, Pourtois & Vuilleumier (2008) intertwined with an affective word categorization task (see Figure 1; see Aarts et al., 2012). The main interest being to evaluate how responses are affectively evaluated, responses from the first task (Go/noGo) served as primes for the word categorization task.

Go/noGo task. We adapted the stimuli of the Go/noGo task to make it child-friendly. Instead of arrows, we used rich and colourful stimuli (i.e. diamonds) that the children were asked to chase in a game-like environment. The diamonds (diameter of ~7.14 cm) had different colors: green (average relative luminance 32.8%), red (average relative luminance 23.0%) or pink (average relative luminance 35.1%). These stimuli were retrieved from an online open-source data base (www.pexels.com). On each trial, the first diamond that appeared on screen was always green. It was followed by a second diamond that could be similar (green) or change in color (red or pink; see Figure 1). The former corresponded to the imperative stimulus (i.e. Go trial), while the latter required response inhibition (i.e. noGo trials).

Evaluative Categorization Task. Stimuli were 15 positive and 15 negative words selected from the Affective norms for French words rated by children and adolescents (Monnier & Syssau, 2017). These words were either nouns or adjectives (see Table S1). Based on this database (Monnier & Syssau, 2017, 2014), we could ascertain that children and adults had similar valence ratings for the selected words ($F(1,56)=0.016$, $p=.90$).

Procedure

The task was performed on a computer. The stimuli were presented in the center of the screen, on a white background. Given the limited and fluctuating attention capacity of children, we shortened the experiment compared to Aarts et al. (2012). We used 100 trials in total, whereas Aarts et al. (2012) used 540 trials. The experiment was composed of a training block (24 trials, corresponding to 16 Go and 8 noGo trials), followed by 4 test blocks, amounting 100 trials (68 Go and 32 noGo, randomly presented). Each trial started with a fixation cross (500 ms), followed by a green diamond shown for a duration varying randomly between 1000 and 2000 ms. This jittering was introduced to reduce anticipatory effects for the second diamond. After its presentation, a blank screen (250 ms) was presented before the second diamond appeared. Its actual duration was determined based on reaction times recorded during the first test block, ensuring personalized calibration of speeded reaction time speed. Similar to Aarts et al. (2012), we used a conservative cutoff and adjusted the stimulus duration of the second diamond in the three subsequent test blocks to be 70% of the mean RT on Go trials (first test block).

In analogy with Aarts et al. (2012), RTs on go trials were labelled online as either fast or slow hits. Fast hits corresponded to RTs falling below this arbitrary RT cutoff, and were associated with a positive performance feedback. In comparison, slow hits were RTs falling above it and were associated with a negative performance feedback. This procedure was used to promote the use of rapid/speedy decisions, and hence used to increase the likelihood of error making on the noGo trials. After the Go/noGo decision, a blank screen was presented for 300 ms, followed by the written affective word (with either a positive or negative valence, see Figure 1), presented until a response was recorded. Participants were asked to perform a two alternative forced choice (2AFC) task based on the valence of the word. Across trials and participants, the presentation was random, such that both the Go and noGo trials were followed by a similar amount of positive and negative words on average. Moreover, this procedure ensured that on average, the 30 words were sampled a similar amount of times. At the end of each trial, a general performance feedback was presented for 1000 ms and informed participants about the accuracy and speed of the Go/noGo decision, as well as the accuracy of the emotion word categorization task.

Participants were asked to use their non-dominant hand for the Go/noGo task and their dominant hand for the 2AFC categorization task. This way, we could rule out that evaluative priming was explained by a motor effector shared between the two tasks

Data Analyses

Demographic and socio-economic variables

For each control variable separately, a t-test (independent or Welch's according to the preliminary data check with Q-Q plots and Levene's test) with a 95% confidence interval (CI) for the mean difference was used, with a false-rate discovery (FDR) p-value correction set at p=.05. A chi-squared test was used for gender. None of them was significant (p>.05), revealing that the two groups of children did not differ significantly for age, gender, socio-economic status, as well as fluid intelligence (Table 1).

Go/noGo Task

Accuracy. We extracted False Alarms -FAs-, Hits, Correct Rejections and Misses for each group (adults, traditional, and Montessori schoolchildren) separately (see Table S2). Next, a mixed-model ANOVA was used to assess possible group differences in accuracy. We also assessed whether the ratio of fast vs. slow hits significantly differed between them.

Reaction time. We computed the mean reaction time (RT) for Hits and compared the three GROUPs on this metric using an ANOVA.

Affective Word Categorization Task

Reaction time. Given the large RT differences between adults and children that precluded a direct comparison between them, RTs for correct responses were first z-transformed using the following formula $(RT\text{-}RT_{group\ mean})/SD_{group}$. To test our a priori hypotheses, we first performed a mixed-model ANOVA on these z-scored RTs with VALENCE (positive vs. negative) and ACTION (hits vs. FAs) as within-subject factors, and GROUP (adults, traditional or Montessori schoolchildren) as between-subjects factor. Fast and slow Hits were combined for this analysis, as the procedure was kept short to remain child friendly, and hence the Go/noGo task eventually generated a limited number of Hits in total. As the three-way interaction was significant (see Results), we next performed three ANOVAs on the non-transformed RTs, for each group separately (adult, traditional and Montessori schoolchildren), with ACTION and VALENCE as within-subject factors (with $\alpha < .05$). Post-hoc Tukey tests were computed when appropriate.

Accuracy. We analyzed the percentage of correct responses using the same statistical model as for the RTs.

Results

Go/noGo Task

Accuracy. Across the three groups, participants' mean accuracy was higher in the Go (M = 75.4%, SE = 23.8%) than in the noGo condition (M = 33.5%, SE = 23.8%), $F(1, 144) = 201.8$, $p < .001$, $\eta^2_p = 0.58$. Furthermore, schoolchildren's mean accuracy for Go and noGo trials collapsed together (M= 41.3%, SE = 3.2%) was lower than adults' mean accuracy (M = 80.8%, SE = 3.2%), $F(2, 144) = 48.5$, $p < .001$, $\eta^2_p = .40$. However, the two groups of children did not differ from each other, $F(2, 144) = .63$, $p = .537$, $\eta^2_p = .01$. Moreover, the ratio of Fast vs. Slow Hits did not differ between the three groups, $F(2, 144) = 1.93$, $p = .148$, $\eta^2_p = .03$.

Reaction time. Mean RTs (in ms) for hits were significantly faster for adults than traditional schoolchildren ($p_{\text{bonferroni}}$<.001), and Montessori schoolchildren ($p_{\text{bonferroni}}$=.019), $F(2, 144) = 7.67$, $p < .001$, $\eta^2_p = .10$ (see Table S2). However, the two groups of children did not differ on this metric ($t(144) = .88$, $p_{\text{tukey}} = .654$).

Affective Word Categorization Task

The number of trials per condition (Hit-positive, Hit-negative, FA-positive and FA-negative) did not differ between groups and conditions, $F(2, 144) = .982$, $p = .377$, $\eta^2_p = .01$ (see Table S3).

Reaction Time. The ANOVA showed a significant three-way interaction, $F(2, 144) = 5.32$, $p = .006$, $\eta^2_p = .07$, suggesting that ACTIONS were differently processed at the affective level (VALENCE) depending on the GROUP (Table S4, Figure 2A and 2B). Since accuracy was different (lower) for children and adults, we selected a subset of errors in children (using a down-sampling technique) to match error frequency with the adults. After down-sampling, we recomputed PES and priming, hence based on a smaller number of data (for the errors of children). The results of this control analysis confirmed those of the main analysis: the omnibus ANOVA (run on the z-transformed RTs) showed a significant three-way interaction between GROUP, ACTION and VALENCE, $F(2, 144) = 4.58$, $p = .012$, $\eta^2_p = .06$. This analysis confirmed that whereas PES was evidenced in the three groups alike, $F(1, 144) = 77.95$, $p < .001$, $\eta^2_p = .35$, priming for errors was only found in adults ($p_{Tukey} < .001$). In comparison, priming for hits was found in children ($p_{Tukey} = .020$), and was different between schoolchildren experiencing the Montessori ($p_{Tukey} = 1.0$) or the traditional pedagogy ($p_{Tukey} = .014$). Accordingly, the imbalance in accuracy between adults and children cannot easily account for the differential affective processing of actions found between them."

Adults Participants. The main effect of ACTION was significant, $F(1, 45) = 35.4$, $p < .001$, $\eta^2_p = .44$, indicating slower RTs after FAs ($M = 755$, $SE = 27.6$) than following Hits ($M = 579$, $SE = 27.6$), and hence the presence of a substantial PES effect (Ullsperger, Danielmeier, et al., 2014). The main effect of VALENCE was trend significant, $F(1 ,45) = 3.86$, $p = .056$, $\eta^2_p = .08$, translating slightly faster RTs for negative ($M = 639$, $SE = 27.4$) than positive words ($M = 695$, $SE = 27.4$). Importantly, the two-way interaction was also significant, $F(1, 45) = 6.39$, $p = .015$, $\eta^2_p = .12$. Post-hoc t-tests revealed that mean RT for negative words was faster than for positive ones after FAs ($p_{\text{tukey}} < .011$), whereas RTs for negative and positive words after hits did not differ ($p_{\text{tukey}} = .973$) (see Figure 2A).

Schoolchildren Experiencing the Traditional Pedagogy. The effect of ACTION was significant, $F(1, 55) = 25.54$, $p < .001$, $\eta^2_p = .32$, showing that RTs following FAs were slower ($M = 1394$, $SE = 69.1$) than following hits ($M = 1173$, $SE = 69.1$), and hence that a PES effect was present in this group as well. VALENCE was also significant, $F(1, 55) = 11.88$, $p = .001$, $\eta^2_p = .18$, with faster RTs for positive ($M = 1208$, $SE = 69.2$) than negative words ($M = 1360$, $SE = 69.2$). Importantly, the two-way interaction was also significant, $F(1, 55) = 4.57$, $p = .037$, $\eta^2_p = .08$. A post-hoc t-test revealed that RTs for positive words were faster than for negative ones after hits ($p_{tukey} < .001$), whereas RTs did not differ between negative and positive words after FAs ($p_{tukey} = .877$) (see Figure 2C). Hence, this result suggests opposite patterns for children and adults: the traditional schoolchildren showed affective priming for correct actions only, whereas adults showed affective priming for errors only.

Schoolchildren Experiencing the Montessori Pedagogy. The effect of ACTION was significant, $F(1, 44) = 27.41$, $p < .001$, $\eta^2_p = .38$, with slower RTs for words following FAs (M = 1634, SE = 87) than following hits (M = 1297, SE = 87), suggesting that PES was also observed in Montessori schoolchildren. VALENCE was significant as well, $F(1, 44) = 5.591$, $p = .023$, $\eta^2_p = .11$. Unlike the traditional schoolchildren, the two-way interaction was not significant however in this group, $F(1, 44) = .802$, $p = .375$, $\eta^2_p = .02$ (see Figure 3D). Accordingly, Montessori schoolchildren did not show a differential affective priming depending on the value of the preceding action.

Schoolchildren Experiencing the Traditional Versus the Montessori Pedagogy. Based on the fact that traditionally-schooled and Montessori schoolchildren did not process the affective valence of words after correct actions in a similar fashion, we additionally ran a mixed model ANOVA comparing directly the two groups of children for evaluative word categorization following hits. This analysis confirmed that the two groups of children differed from each other, $F(1, 99) = 3.99$, $p = .049$, $\eta^2_p = .04$. Whereas the affective priming was significant after Hits for traditional schoolchildren ($t(99) = -4.04$, $p_{tukey} < .001$), it was not for Montessori schoolchildren ($t(99) = -0.60$, $p_{tukey} = .933$). When controlling for gender and SES in an ANCOVA, this effect was trend significant, $F(1, 94) = 3.95$, $p = .050$, $\eta^2_p = .04$. We did not add age and fluid intelligence as covariates in this ANCOVA, as they correlated with one another, and moreover, they both correlated strongly with the mean RT speed, making the interpretation of these results difficult.

Accuracy. The ANOVA showed a significant three-way interaction, $F(2, 144) = 6.82$, $p = .001$, $\eta^2_p = .09$ (Table S5).

Adults Participants. The main effect of ACTION was trend significant only, $F(1, 45) = 3.54$, $p = .067$, $\eta^2_p = .07$, translating a slightly higher accuracy following hits ($M = 87.9$, $SE = 2.0$) than FAs ($M = 83.6$, $SE = 2.0$). VALENCE was significant, $F(1, 45) = 14.46$, $p < .001$, $\eta^2_p = .24$, with a higher accuracy for negative ($M = 90.3$, $SE = 2.0$) than positive words ($M = 81.2$, $SE = 2.0$). The two-way interaction was significant, $F(1, 45) = 21.49$, $p < .001$, $\eta^2_p = .32$, with a higher accuracy for negative than positive words after FAs ($p_{\text{tukey}} < .001$), but no significant difference between them after hits ($p_{\text{tukey}} = .956$), in agreement with a previous study performed in adults (De Saedeleer & Pourtois, 2016).

Schoolchildren Experiencing the Traditional Pedagogy. There was a significant main effect of ACTION, $F(1, 55) = 8.06$, $p = .006$, $\eta^2_p = .13$, with a higher accuracy after hits ($M = 89.7$, $SE = 1.7$) than FAs ($M = 84.9$, $SE = 1.7$). VALENCE was not significant, $F(1, 55) = 0.19$, $p = .664$, $\eta^2_p = .003$, nor the interaction between VALENCE and ACTION, $F(1, 55) = 2.18$, $p = .145$, $\eta^2_p = .04$.

Schoolchildren Experiencing the Montessori Pedagogy. The effect of ACTION was significant, $F(1, 44) = 4.94$, $p = .031$, $\eta^2_p = .10$, with a higher accuracy after hits ($M = 89.6$, $SE = 1.9$) than FAs ($M = 86.6$, $SE = 1.9$). The main effect of VALENCE was significant, $F(1, 44) = 10.32$, $p = .002$, $\eta^2_p = .19$, with a higher accuracy for positive ($M = 91.7$, $SE = 2.1$) than negative words ($M = 84.5$, $SE = 2.1$. The two-way interaction was not significant, $F(1, 44) = .05$, $p = .831$, $\eta^2_p = .001$.

Discussion

In this study, we compared PM in 8-12 years old schoolchildren and adults. We also tested whether the pedagogy experienced at school could modulate PM in children. Based on earlier studies performed in adults only (Aarts et al., 2012; De Saedeleer & Pourtois, 2016), we used a dual task procedure in order to derive two dissociable correlates of PM at the behavioral level: the affective processing of actions (suggesting that response errors are processed as negative events at the adult age) and the PES (suggesting an automatic attention orienting to response errors). Our results showed that even though response errors led to a PES in the three groups, the affective processing of actions substantially differed between them. More specifically, although the adults evaluated their response errors as negative events, the children did not. Moreover, and contrary to the adults, these children, if experiencing the traditional pedagogy, actually evaluated correct responses as positive events. Further, the children experiencing the Montessori pedagogy did not show this latter priming effect. Here after, we discuss the possible implications of these new results, which suggest that PM is qualitatively different for children when compared to adults. More generally, they also lend support to the notion that the automatic attention orienting towards response errors, as captured by the PES, and their affective processing as negative events (as visible in the priming effect), are two distinct components of PM (e.g. Koban & Pourtois, 2014).

Our results are consistent with previous studies showing that young children, very much like adults, systematically slow down following response errors (Smulders et al., 2016). Hence, it appears unlikely that PES would follow a U-shaped developmental trajectory. Given that the PES could reflect an automatic orienting response to deviant events (i.e., “oddball” response errors in the trial series, Danielmeier & Ullsperger, 2011; Notebaert et al., 2009), our results suggest that this attention-based PM effect is probably mature in 8-12 years old schoolchildren. This interpretation is compatible with a vast literature in developmental psychology showing that the stimulus-driven attentional system (i.e. exogenous attention) is functional and active early in life, and before the one involved in the top-down control of attention (Johnson, Posner, & Rothbart, 1991), with an asymmetric development observed between them (Farrant & Uddin, 2015). This dissociation has been confirmed across many modalities or tasks, including language processing (de Diego-Balaguer, Martinez-Alvarez, & Pons, 2016). In fact, young children's attention is easily attracted toward salient stimuli or events in their environment (such as response errors in the present case), but disengaging and switching their attention require more years of development to be fully functional (Farrant & Uddin, 2015; Wainwright & Bryson, 2002). Interestingly, we found that this behavioral adaptation following errors was not smaller or larger in magnitude for children compared to adults in our study, suggesting indirectly that the PES seen at the adult age likely reflects the operations of a core PM component that is already active early in life (e.g. Basirat, Dehaene, & Dehaene-Lambertz, 2014), and does not undergo major change between childhood and adulthood. Moreover, since we failed to observe a difference for the PES between the traditional and Montessori schoolchildren, it is likely that this PM component is not easily liable to contextual effects, including the affective meaning of response errors (or the lack thereof), and its reinforcement by external factors or agents (such as the schooling system encountered).

This age-invariance of the PES sharply contrasts with the significant modulation of the affective processing of self-generated actions by age found in our study. Replicating previous results for adult participants (Aarts et al., 2012, 2013; De Saedeleer & Pourtois, 2016), we found here that response errors were aversive for them (Hajcak & Foti, 2008), even though a child-friendly version of the Go/noGo task was used, and these response errors only threatened the self indirectly, or "self-efficacy" broadly speaking (e.g. they did not entail monetary losses for instance). Moreover, we could rule out a speed accuracy tradeoff underlying this evaluative priming effect because the adult participants were not only faster for negative than positive words after errors, they were also more accurate in the former case. However, and strikingly, this effect was not found in 8-12 years old schoolchildren, who showed instead a RT facilitation for positive compared to negative words following hits, selectively, suggesting that, unlike adults, they processed correct actions as positive events. Hence, our results suggest that the affective processing of actions is asymmetrical, but this imbalance takes different forms and expressions depending on age. Importantly, because the positive and negative words used as targets in our study were rated in a similar way for the children and the adults, it is unlikely that this asymmetry arose because negative or positive words were perceived as less or more negative/positive by the children compared to the adults. Instead, our results suggest that the way the correct or incorrect action preceding this word was evaluated substantially differed between the two groups.

The affective processing of correct actions as positive events seen in these children aligns with earlier work showing a stronger impact of positive than negative feedback on learning in 8-9 years old children, with a reversal of this effect occurring later during the development, at around 11-13 years old (van Duijvenvoorde, Zanolie, Rombouts, Raijmakers, & Crone, 2008). Moreover, this shift seems to reflect a change in what children perceive as salient during learning, as opposed to be driven by valence only (van den Bos, Guroglu, van den Bulk, Rombouts, & Crone, 2009). Accordingly, it is likely that the opposite priming effects found for children and adults in this study occurred as a result of a change through development and maturation in the saliency of the action value, where children mostly assigned a positive value to correct decisions whereas adults assigned more weight to incorrect ones. However, future studies will be needed to unveil the cognitive and emotional factors that enable this profound shift in the way self-generated actions are evaluated by children vs. adults.

Tentatively, the lack of distinctive evaluative processing of errors in these children could be explained by the fact that these events are often instrumental to learning at that age and/or these events do not pose a main threat or challenge to the self (Chrysikou et al., 2013; Chrysikou, Novick, Trueswell, & Thompson-Schill, 2011; Thompson-Schill et al., 2009). In line with this idea, it was found previously that children are actually better than adults to learn abstract causal relationships as they could more easily update prior knowledge, and this way more flexibly solve problems (Lucas, Bridgers, Griffiths, & Gopnik, 2014). Hence, a greater flexibility and lower error-avoidance could explain why children do not automatically assign negative value to response errors, even though they are generating them now and them during decision making, and automatically orient to them after their occurrences (as reflected by the PES). Likewise, this specific processing style of children could also explain why they actually assign a positive value to correct actions, which usually manifest that the task goal has been met (i.e. an overt response in face of an imperative go stimulus has been made in the present case), and hence that learning was successful. Further and more generally, this specific processing style could stem from the fact that the prefrontal cortex is not fully matured yet in these children (Crone & van der Molen, 2007). As a result, evaluative processes, including those involved in action and outcome, are already functional, but probably recruiting a network of (subcortical) brain areas that are different than those used by adult participants, and characterized by reward sensitivity (van Duijvenvoorde, Peters, Braams, & Crone, 2016).

Remarkably and unlike the PES, this priming effect in children was restricted to those enrolled in the traditional schooling system. In comparison, Montessori schoolchildren were slower following errors, but they did not process correct actions as positive events. This difference suggests that the affective processing of actions, unlike the PES, is shaped by both age and pedagogy. At that age, the way self-generated actions are assessed by peers and evaluators (e.g. school teachers) is likely to influence profoundly how they are eventually processed along an affective dimension by the children who execute them. Because children experiencing Montessori pedagogy are usually much less confronted with evaluative feedback and reinforcers for their actions than those experiencing traditional pedagogy (Lillard, 2013; Lillard, 2012; Rathunde, 2001; Rathunde & Csikszentmihalyi, 2005), it is possible that their actions acquire less specific affective values, as our new results suggest indirectly. Hence, we contend that the difference in affective priming found between Montessori versus traditional schoolchildren could stem from a differential reinforcement learning (RL) effect. Although speculative at this stage, it is feasible that the schooling environment actually shapes PM by influencing specific RL parameters. In this perspective, it appears relevant to consider the difference between model-free and model-based RL (Dolan & Dayan, 2013; Glascher, Daw, Dayan, & O'Doherty, 2010; Neftci & Averbeck, 2019). Whereas the former refers to how value/weight for a given outcome or action is computed and expected, the latter relates to a state transition for it (model-based RL), thereby enabling an optimal processing of sensory cues that is weighted according to prior beliefs. In adults, it has been shown that this framework is extremely valuable as it can account for a wide range of phenomena during RL, including modulatory effects of feedback type or reward on it (Mattar, Thompson-Schill, & Bassett, 2018). Accordingly, it would be extremely informative in future studies to link more directly changes in PM with possible alterations of specific RL parameters (using computational modelling methods for example), in order to obtain a more mechanistic understanding of how development and prefrontal cortex maturation can influence it. Finally, we found that accuracy was higher for positive than negative words in Montessori schoolchildren, but not in traditionally schoolchildren. This result is compatible with previous findings showing that the former children can exhibit a bias for positive emotional stimuli (Denervaud *et al.*, *in revision*). Moreover, this finding is interesting when considering the fact that they did not show evaluative priming of actions compared to traditionally schoolchildren. Accordingly, this bias for positive stimuli does not contribute to influence the evaluative processing of actions as good or bad, suggesting indirectly that this bias could be restricted to external stimuli only.

A few limitations warrant comment. First, we used a child friendly version of the dual task previously devised for adult participants (Aarts et al., 2012), and as a result, we only had a limited trial number per condition. A way to overcome this limitation in future studies would be to increase trial number, although this might be detrimental to the selective attention or task's involvement of these children. Second and more importantly, here we performed a cross-sectional study comparing children to adults, but it appears important to assess how the PES and the evaluative processing of actions could change as a function of development and prefrontal cortex maturation, which would require the use of longitudinal studies instead. Although these studies are more expensive and time-consuming than cross-sectional ones, they are the ones who could provide a unique and unprecedented insight into the developmental trajectory of PM until the adult age. Third, there also might be a selection bias in our sample as we drew, for practical and ethical reasons, schoolchildren experiencing the Montessori pedagogy from private schools exclusively, whereas the schoolchildren experiencing the traditional pedagogy attended regular schools instead (where practices regarding grades and formal assessments are quite homogenized due to local policies). Accordingly, it remains to establish whether Montessori pedagogy as such, or alternatively any private schooling system, eventually yields a differential affective processing of (correct) actions in children. A way to address this limitation would be to compare, using the same experimental design as used here, Montessori children to children enrolled in a private school as well, but where a different pedagogy than Montessori is used. Also, it might be interesting to consider parental attitude and education doctrine in future studies, as these variables might influence the way actions, and more specifically response errors, are appraised by children and in turn influence their behavior. Finally, it should be noted that while information about socio-economic status and fluid intelligence were collected in children, adults were not tested and compared on these variables.

To conclude, our findings shed new lights on PM in children, and more specifically two fundamental components that underlie this utmost important cognitive ability. Results suggest that children of that age, very much like adults, orient automatically to response errors, as reflected by the PES. However, unlike what we observed for the adults, we found that response errors were not evaluated as negative events by the children. Instead, they evaluated correct actions as positive events, with this reward-related effect being only found for traditional schoolchildren. All in all, these results therefore suggest that PM is composed of an age-invariant component that allows individuals to orient attention to (deviant) errors, while the affective evaluation of actions is shaped by both development and school environment. This in turn may allow children as well as adults to assign value to action in a flexible and context-dependent fashion, and ultimately foster goal-adaptive behavior in an ever-changing environment.

REFERENCES

Aarts, K., De Houwer, J., & Pourtois, G. (2012). Evidence for the automatic evaluation of self-generated actions. *Cognition, 124*(2), 117-127. doi:10.1016/j.cognition.2012.05.009

Aarts, K., De Houwer, J., & Pourtois, G. (2013). Erroneous and correct actions have a different affective valence: evidence from ERPs. *Emotion, 13*(5), 960-973. doi:10.1037/a0032808

Basirat, A., Dehaene, S., & Dehaene-Lambertz, G. (2014). A hierarchy of cortical responses to sequence violations in three-month-old infants. *Cognition, 132*(2), 137-150. doi:10.1016/j.cognition.2014.03.013

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624-652. doi:10.1037//0033-295x.108.3.624

Chrysikou, E. G., Hamilton, R. H., Coslett, H. B., Datta, A., Bikson, M., & Thompson-Schill, S. L. (2013). Noninvasive transcranial direct current stimulation over the left prefrontal cortex facilitates cognitive flexibility in tool use. *Cogn Neurosci, 4*(2), 81-89. doi:10.1080/17588928.2013.768221

Chrysikou, E. G., Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2011). The other side of cognitive control: can a lack of cognitive control benefit language and cognition? *Top Cogn Sci, 3*(2), 253-256. doi:10.1111/j.1756-8765.2011.01137.x

Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology, Vol 64, 52*, 337-367. doi:10.1146/annurev.psych.52.1.337

Crone, E. A., & van der Molen, M. W. (2007). Development of decision making in school-aged children and adolescents: evidence from heart rate and skin conductance analysis. *Child Dev, 78*(4), 1288-1301. doi:10.1111/j.1467-8624.2007.01066.x

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Front Psychol, 2*, 233. doi:10.3389/fpsyg.2011.00233

de Diego-Balaguer, R., Martinez-Alvarez, A., & Pons, F. (2016). Temporal Attention as a Scaffold for Language Development. *Front Psychol, 7*, 44. doi:10.3389/fpsyg.2016.00044

Denervaud, S., Knebel, J. F., Hagmann, P., & Gentaz, E. (2019). Beyond executive functions, creativity skills benefit academic outcomes: Insights from Montessori education. *PLoS One, 14*(11), e0225319. doi:10.1371/journal.pone.0225319

Denervaud, S., Knebel, J. F., Immordino-Yang, M. H., & Hagmann, P. (2020). Effects of Traditional Versus Montessori Schooling on 4- to 15-Year Old children's Performance Monitoring. *Mind Brain and Education*. doi:https://doi.org/10.1111/mbe.12233

Denervaud, S., Mumenthaler, C., Gentaz, E., & Sander, D. (*in revision*). Emotion Recognition Development: Preliminary Evidence for an Effect of School Pedagogy.

De Saedeleer, L., & Pourtois, G. (2016). Evaluative priming reveals dissociable effects of cognitive versus physiological anxiety on action monitoring. *Emotion, 16*(4), 498-514. doi:10.1037/emo0000149

Dignath, D., Eder, A. B., Steinhauser, M., & Kiesel, A. (2019). Conflict Monitoring and the Affective Signaling Hypothesis – an Integrative Review. *Psychonomic Bulletin & Review*.

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron, 80*(2), 312-325. doi:10.1016/j.neuron.2013.09.007

Farrant, K., & Uddin, L. Q. (2015). Asymmetric development of dorsal and ventral attention networks in the human brain. *Dev Cogn Neurosci, 12*, 165-174. doi:10.1016/j.dcn.2015.02.001

Fischer, K. W., & Rose, S. P. (1996). Dynamic growth cycles of brain and cognitive development. In R. Thatcher, G. R. Lyon, J. Rumsey, & N. Krasnegor (Eds.), *Developmental neuroimaging: Mapping the development of brain and behavior* (pp. 263-279). New York: Academic Press.

Glascher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron, 66*(4), 585-595. doi:10.1016/j.neuron.2010.04.016

Gopnik, A., & Wellman, H. M. (2012). Reconstructing Constructivism: Causal Models, Bayesian Learning Mechanisms, and the Theory Theory. *Psychological Bulletin, 138*(6), 1085-1108. doi:10.1037/a0028044

Grisay, A., de Jong, J. H., Gebhardt, E., Berezner, A., & Halleux-Monseur, B. (2007). Translation equivalence across PISA countries. *J Appl Meas, 8*(3), 249-266.

Hajcak, G., & Foti, D. (2008). Errors are aversive: defensive motivation and the error-related negativity. *Psychol Sci, 19*(2), 103-108. doi:10.1111/j.1467-9280.2008.02053.x

Inzlicht, M., Bartholow, B. D., & Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends Cogn Sci, 19*(3), 126-132. doi:10.1016/j.tics.2015.01.004

Johnson, M. H., Posner, M. I., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: contingency learning, anticipatory looking, and disengaging. *J Cogn Neurosci, 3*(4), 335-344. doi:10.1162/jocn.1991.3.4.335

Jones, C. R., Olson, M. A., & Fazio, R. H. (2010). Evaluative Conditioning: The "How" Question. *Adv Exp Soc Psychol, 43*, 205-255. doi:10.1016/S0065-2601(10)43005-1

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science, 6*(5), 498-504. doi:Doi 10.1111/1467-7687.00307

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). The wick in the candle of learning: epistemic curiosity activates reward circuitry and enhances memory. *Psychol Sci, 20*(8), 963-973. doi:10.1111/j.1467-9280.2009.02402.x

Kaplan, A., & Patrick, H. (2016). Learning environments and motivation. . In K. Wentzel & D. Miele (Eds.), *Handbook of motivation at school* (2nd ed., pp. 251-274). New York: Routlege.

Koban, L., & Pourtois, G. (2014). Brain systems underlying the affective and social monitoring of actions: an integrative review. *Neurosci Biobehav Rev, 46 Pt 1*, 71-84. doi:10.1016/j.neubiorev.2014.02.014

Lillard, A. (2013). Playful Learning and Montessori Education. *American Journal of Play, 5*(2), 157-186.

Lillard, A., & Else-Quest, N. (2006). The early years. Evaluating Montessori education. *Science, 313*(5795), 1893-1894. doi:10.1126/science.1132362

Lillard, A. S. (2012). Preschool children's development in classic Montessori, supplemented Montessori, and conventional programs. *J Sch Psychol, 50*(3), 379-401. doi:10.1016/j.jsp.2012.01.001

Lillard, A. S. (2017). Why Do the Children (Pretend) Play? *Trends in Cognitive Sciences, 21*(11), 826-834. doi:10.1016/j.tics.2017.08.001

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. *Front Psychol, 8*, 1783. doi:10.3389/fpsyg.2017.01783

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition, 131*(2), 284-299. doi:10.1016/j.cognition.2013.12.010

Marshall, C. (2017). Montessori education: a review of the evidence base. *npj Science of Learning, 2*(1). doi:10.1038/s41539-017-0012-7

Mattar, M. G., Thompson-Schill, S. L., & Bassett, D. S. (2018). The network architecture of value learning. *Netw Neurosci, 2*(2), 128-149. doi:10.1162/netn_a_00021

Monnier, C., & Syssau, A. (2017). Affective norms for 720 French words rated by children and adolescents (FANchild). *Behav Res Methods, 49*(5), 1882-1893. doi:10.3758/s13428-016-0831-0

Moulin-Frier, C., Nguyen, S. M., & Oudeyer, P. Y. (2013). Self-organization of early vocal development in infants and machines: the role of intrinsic motivation. *Front Psychol, 4*, 1006. doi:10.3389/fpsyg.2013.01006

Montessori, M. (1936). *The Secret of Childhood* (1981 ed.). New York: Ballantine.

Neftci, E. O., & Averbeck, B. B. (2019). Reinforcement learning in artificial and biological systems. *Nature Machine Intelligence, 1*(3), 133-143. doi:10.1038/s42256-019-0025-4

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition, 111*(2), 275-279. doi:10.1016/j.cognition.2009.02.002

Oudeyer, P. Y., Gottlieb, J., & Lopes, M. (2016). Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. *Changing Brains Applying Brain Plasticity to Advance and Recover Human Ability, 229*, 257-284. doi:10.1016/bs.pbr.2016.05.005

Oudeyer, P. Y., & Smith, L. B. (2016). How Evolution May Work Through Curiosity-Driven Developmental Process. *Top Cogn Sci, 8*(2), 492-502. doi:10.1111/tops.12196

Pellegrini, A. D., Dupuis, D., & Smith, P. K. (2007). Play in evolution and development. *Developmental Review, 27*(2), 261-276. doi:10.1016/j.dr.2006.09.001

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Pourtois, G., Vocat, R., N'Diaye, K., Spinelli, L., Seeck, M., & Vuilleumier, P. (2010). Errors recruit both cognitive and emotional monitoring systems: simultaneous intracranial recordings in the dorsal anterior cingulate gyrus and amygdala combined with fMRI. *Neuropsychologia, 48*(4), 1144-1159. doi:10.1016/j.neuropsychologia.2009.12.020

Rabbitt, P. M. (1966). Errors and error correction in choice-response tasks. *J Exp Psychol, 71*(2), 264-272.

Rathunde, K. (2001). Montessori Education and optimal experieunce: a framework for new research. *The NAMTA Journal, 26*(1), 11-43.

Rathunde, K., & Csikszentmihalyi, M. (2005). Middle School Students' Motivation and Quality of Experience: A Comparison of Montessori and Traditional School Environments. *American Journal of Education, 111*(3).

Raven, J., Raven, J. C., & Court, J. H. (2003). Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview. . San Antonio, TX: Harcourt Assessment.

Smulders, S. F., Soetens, E., & van der Molen, M. W. (2016). What happens when children encounter an error? *Brain Cogn, 104*, 34-47. doi:10.1016/j.bandc.2016.02.004

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Curr Dir Psychol Sci, 18*(5), 259-263. doi:10.1111/j.1467-8721.2009.01648.x

Ullsperger, M., & Danielmeier, C. (2016). Reducing Speed and Sight: How Adaptive Is Post-Error Slowing? *Neuron, 89*(3), 430-432. doi:10.1016/j.neuron.2016.01.035

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiol Rev, 94*(1), 35-79. doi:10.1152/physrev.00041.2012

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends Cogn Sci, 18*(5), 259-267. doi:10.1016/j.tics.2014.02.009

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: the negativity bias in social-emotional development. *Psychol Bull, 134*(3), 383-403. doi:10.1037/0033-2909.134.3.383

van den Bos, W., Guroglu, B., van den Bulk, B. G., Rombouts, S. A. R. B., & Crone, E. A. (2009). Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. *Frontiers in Human Neuroscience, 3*. doi:ARTN 52
10.3389/neuro.09.052.2009

van Duijvenvoorde, A. C., Peters, S., Braams, B. R., & Crone, E. A. (2016). What motivates adolescents? Neural responses to rewards and their influence on adolescents' risk taking, learning, and cognitive control. *Neurosci Biobehav Rev, 70*, 135-147. doi:10.1016/j.neubiorev.2016.06.037

van Duijvenvoorde, A. C., Zanolie, K., Rombouts, S. A., Raijmakers, M. E., & Crone, E. A. (2008). Evaluating the negative or valuing the positive? Neural mechanisms supporting feedback-based learning across development. *J Neurosci, 28*(38), 9495-9503. doi:10.1523/JNEUROSCI.1485-08.2008

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: MA: Harvard University Press.

Wainwright, A., & Bryson, S. E. (2002). The development of exogenous orienting: mechanisms of control. *J Exp Child Psychol, 82*(2), 141-155.

Wessel, J. R. (2018). An adaptive orienting theory of error processing. *Psychophysiology, 55*(3). doi:10.1111/psyp.13041

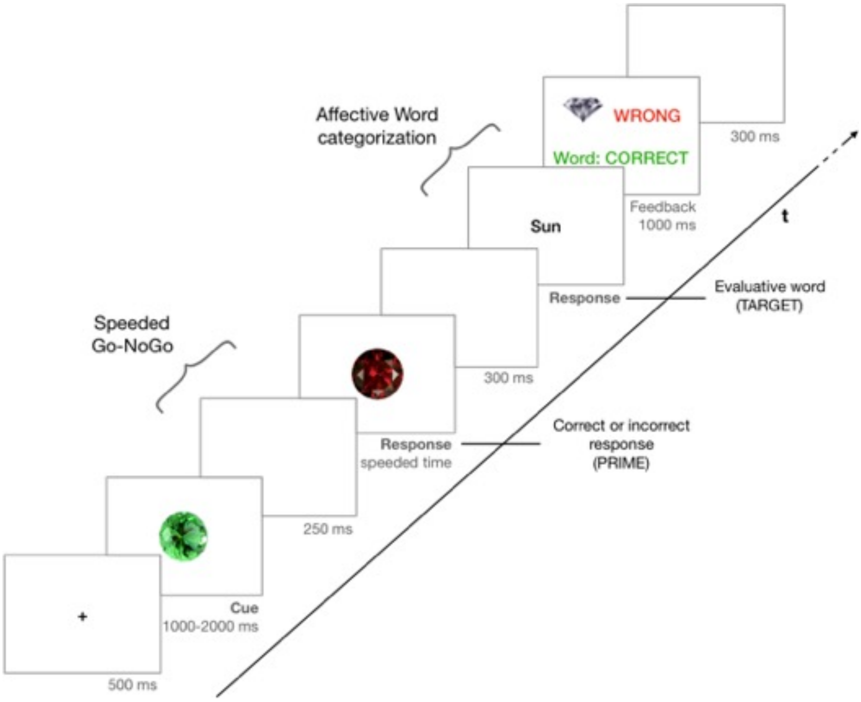


Figure 1 ***Evaluative priming task.*** On each trial, participants performed two tasks: first a speeded Go/noGo task (that led either to correct or incorrect responses), followed by an affective word categorization task (based on positive and negative words), serving respectively as primes and targets in an evaluative priming procedure.

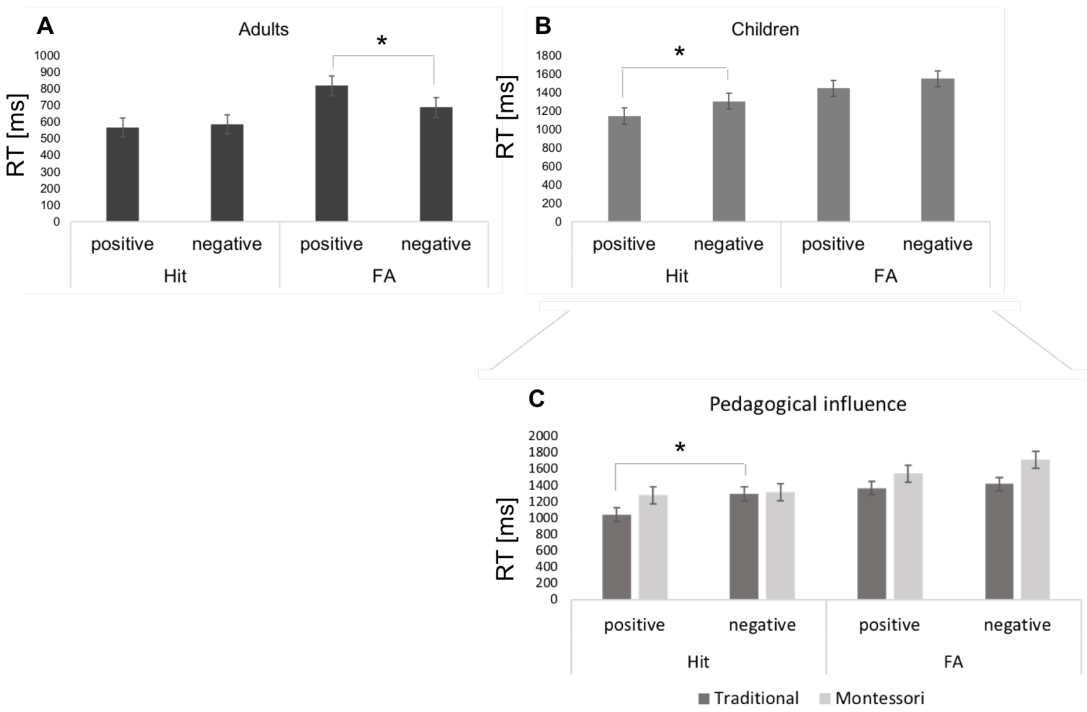


Figure 2 ***Affective word categorization task.*** Mean RTs for (A) adults and (B) children. (C) Children were split into two groups, according to the pedagogy experienced, either traditional or Montessori. RT stands for Reaction Time expressed in milliseconds (ms) and the error bar corresponds to the standard error of the mean.

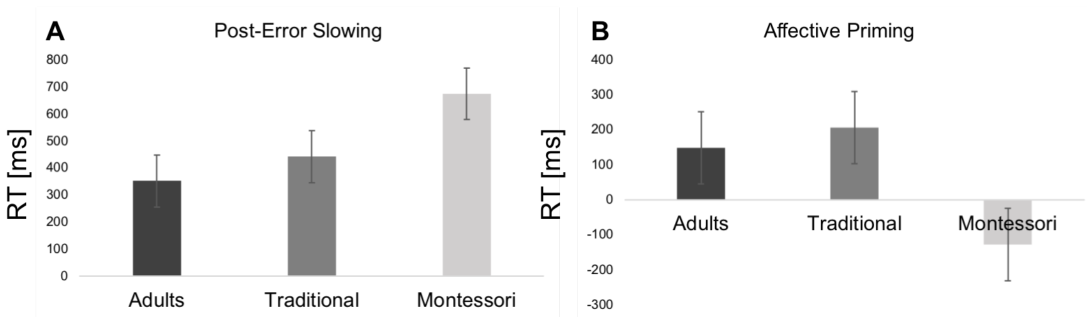


Figure 3 ***Summary of main results.*** (A) PES and (B) Affective Priming, separately for each group. PES was computed as (RT_{FA}-RT_{Hit}) whereas Affective Priming was computed as ($RT_{Hit\ Neg}$ + $RT_{FA\ Pos}$)-($RT_{Hit\ Pos}$ + $RT_{FA\ Neg}$). RT stands for Reaction Time in milliseconds, error bar corresponds to the standard error of the mean.

	Schoolchildren Group				
	M	T	t or X^2	p-values FDR corrected	Cohen's d
N (girls)	45 (17)	56 (29)	3.40	0.13	
Age [years]	10.3 (1.2)	10.5 (1.1)	0.82	0.42	0.16
min, max	8.31-12.8	8.5-12.8			
SES [au]	7.10 (0.8)	6.77 (1.1)	1.69	0.13	0.34
Fluid intelligence [score]	34.1 (1.6)	33.4 (2.3)	1.78	0.13	0.35

	Adult Group
N (women)	46 (30)
Age [years]	28.0 (9.4)
min, max	20-40

Note. Mean and SD. Au = arbitrary unit, M=Montessori schooling background T=traditional schooling background.

Table 1 Descriptive statistics of control variables, and group comparisons.

SUPPLEMENTARY MATERIAL

Positive targets		**Negative targets**	
Ami (friend)	Liberté (freedom)	Cauchemar (nightmare)	Maladie (disease)
Blague (joke)	Paix (peace)	Chagrin (grief)	Malheur (misfortune)
Bonheur (happiness)	Paradis (paradise)	Diable (devil)	Méchanceté (wickedness)
Cadeau (gift)	Plaisir (pleasure)	Douleur (pain)	Peur (fear)
Chance (luck)	Rêve (dream)	Enfer (hell)	Regret (regret)
Fête (party)	Rire (laugh)	Fatigue (tiredness)	Souffrance (misery)
Humour (humor)	Soleil (sun)	Guerre (war)	Tristesse (sadness)
Joie (joy)		Larme (tear)	

Table S1 Target words selected from the Affective norms for French words rated by children and adolescents (FANchild) (Monnier & Syssau, 2017)

Descriptives

	Group	**Hit**	**Correct rejection**	**Miss**	**FA**	**Mean Accuracy**
Mean responses (SD) [%]	Adults	100	81.4 (12.4)	0.00	18.6 (12.4)	90.7 (6.2)
	Traditional	84.8 (19.9)	61.3 (20.3)	18.2 (9.6)	38.7 (20.4)	73.0 (13.5)
	Montessori	81.0 (20.1)	58.4 (20.8)	19.6 (9.8)	41.6 (20.8)	69.7 (12.9)
RT (SD) [ms]	Adults	282 (53)				
	Traditional	370 (135)				
	Montessori	350 (135)				

Table S2 Descriptive statistics of the Go/noGo task.

Descriptives

	Group	NB_HitPos	NB_HitNeg	NB_FAPos	NB_FANeg
Mean (SD)	Adults	34.6 (3.51)	33.4 (3.51)	3.13 (1.34)	4.04 (2.71)
	Traditional	35.7 (4.79)	32.3 (4.79)	6.23 (3.29)	6.79 (3.61)
	Montessori	36.6 (3.91)	31.4 (3.93)	7.40 (4.01)	6.38 (3.20)

Table S3 Number of trials per condition included in the analyses.

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
ACTION	96.371	1	96.371	78.418	< .001	0.353
ACTION ✻ Group	40.671	2	20.336	16.547	< .001	0.187
Residual	176.967	144	1.229			
VALENCE	0.122	1	0.122	0.112	0.739	0.001
VALENCE ✻ Group	15.892	2	7.946	7.252	< .001	0.092
Residual	157.775	144	1.096			
ACTION ✻ VALENCE	8.254	1	8.254	6.540	0.012	0.043
ACTION ✻ VALENCE ✻ Group	**13.420**	**2**	**6.710**	**5.317**	**0.006**	**0.069**
Residual	181.727	144	1.262			

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
Group	48.3	2	24.13	5.65	0.004	0.073
Residual	615.3	144	4.27			

Note. Type 3 Sums of Squares

Table S4 Results of the omnibus ANOVA performed on the mean z-RTs (affective word categorization task).

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
ACTION	0.23416	1	0.23416	14.6871	< .001	0.093
ACTION ✻ Group	0.00851	2	0.00426	0.2670	0.766	0.004
Residual	2.29585	144	0.01594			
VALENCE	9.35e-4	1	9.35e-4	0.0314	0.860	0.000
VALENCE ✻ Group	0.62825	2	0.31413	10.5406	< .001	0.128
Residual	4.29145	144	0.02980			
ACTION ✻ VALENCE	0.33385	1	0.33385	16.4932	< .001	0.103
ACTION ✻ VALENCE ✻ Group	**0.27602**	**2**	**0.13801**	**6.8182**	**0.001**	**0.087**
Residual	2.91479	144	0.02024			

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
Group	0.0531	2	0.0265	0.519	0.596	0.007
Residual	7.3672	144	0.0512			

Note. Type 3 Sums of Squares

Table S5 Results of the omnibus ANOVA performed on the accuracy score (% correct affective word categorization).

SUPPLEMENTARY MATERIAL

Positive targets		Negative targets	
Ami (friend)	Liberté (freedom)	Cauchemar (nightmare)	Maladie (disease)
Blague (joke)	Paix (peace)	Chagrin (grief)	Malheur (misfortune)
Bonheur (happiness)	Paradis (paradise)	Diable (devil)	Méchanceté (wickedness)
Cadeau (gift)	Plaisir (pleasure)	Douleur (pain)	Peur (fear)
Chance (luck)	Rêve (dream)	Enfer (hell)	Regret (regret)
Fête (party)	Rire (laugh)	Fatigue (tiredness)	Souffrance (misery)
Humour (humor)	Soleil (sun)	Guerre (war)	Tristesse (sadness)
Joie (joy)		Larme (tear)	

Table S1 Target words selected from the Affective norms for French words rated by children and adolescents (FANchild) (Monnier & Syssau, 2017)

Descriptives

	Group	Hit	Correct rejection	Miss	FA	Mean Accuracy
Mean responses (SD) [%]	Adults	100	81.4 (12.4)	0.00	18.6 (12.4)	90.7 (6.2)
	Traditional	84.8 (19.9)	61.3 (20.3)	18.2 (9.6)	38.7 (20.4)	73.0 (13.5)
	Montessori	81.0 (20.1)	58.4 (20.8)	19.6 (9.8)	41.6 (20.8)	69.7 (12.9)
RT (SD) [ms]	Adults	282 (53)				
	Traditional	370 (135)				
	Montessori	350 (135)				

Table S2 Descriptive statistics of the Go/noGo task.

Descriptives

	Group	NB_HitPos	NB_HitNeg	NB_FAPos	NB_FANeg
Mean (SD)	Adults	34.6 (3.51)	33.4 (3.51)	3.13 (1.34)	4.04 (2.71)
	Traditional	35.7 (4.79)	32.3 (4.79)	6.23 (3.29)	6.79 (3.61)
	Montessori	36.6 (3.91)	31.4 (3.93)	7.40 (4.01)	6.38 (3.20)

Table S3 Number of trials per condition included in the analyses.

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
ACTION	96.371	1	96.371	78.418	< .001	0.353
ACTION ✻ Group	40.671	2	20.336	16.547	< .001	0.187
Residual	176.967	144	1.229			
VALENCE	0.122	1	0.122	0.112	0.739	0.001
VALENCE ✻ Group	15.892	2	7.946	7.252	< .001	0.092
Residual	157.775	144	1.096			
ACTION ✻ VALENCE	8.254	1	8.254	6.540	0.012	0.043
ACTION ✻ VALENCE ✻ Group	**13.420**	**2**	**6.710**	**5.317**	**0.006**	**0.069**
Residual	181.727	144	1.262			

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
Group	48.3	2	24.13	5.65	0.004	0.073
Residual	615.3	144	4.27			

Note. Type 3 Sums of Squares

Table S4 Results of the omnibus ANOVA performed on the mean z-RTs (affective word categorization task).

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
ACTION	0.23416	1	0.23416	14.6871	< .001	0.093
ACTION ✻ Group	0.00851	2	0.00426	0.2670	0.766	0.004
Residual	2.29585	144	0.01594			
VALENCE	9.35e-4	1	9.35e-4	0.0314	0.860	0.000
VALENCE ✻ Group	0.62825	2	0.31413	10.5406	< .001	0.128
Residual	4.29145	144	0.02980			
ACTION ✻ VALENCE	0.33385	1	0.33385	16.4932	< .001	0.103
ACTION ✻ VALENCE ✻ Group	**0.27602**	**2**	**0.13801**	**6.8182**	**0.001**	**0.087**
Residual	2.91479	144	0.02024			

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
Group	0.0531	2	0.0265	0.519	0.596	0.007
Residual	7.3672	144	0.0512			

Note. Type 3 Sums of Squares

Table S5 Results of the omnibus ANOVA performed on the accuracy score (% correct affective word categorization).

EMOTION RECOGNITION DEVELOPMENT:

Preliminary evidence for an effect of school pedagogical practices.

Solange Denervaud[1,2], Christian Mumenthaler[1,3], Edouard Gentaz[1,3] & David Sander[1,3*]

[1] *Swiss Center for Affective Sciences, Campus Biotech, University of Geneva, 1202 Geneva, Switzerland*

[2] *Connectomics Lab, Department of Radiology, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), 1011 Lausanne, Switzerland*

[3] *Faculté de Psychologie et des Sciences de l'Education (FAPSE), University of Geneva, 1204 Geneva, Switzerland*

*corresponding authors: solange.denervaud@unil.ch, david.sander@unige.ch

Acknowledgement We are grateful to all the children, parents, teachers and school directors who made this study possible. We thank Jennifer Malsert, PhD, for sharing the offset task, Camila Camacho and Morgane Budry, for helping with data collection, Andres Posada, PhD, Nicolas Denervaud, PhD, and Lucas Tamarit, PhD for technical support, and Vanessa Sennwald for editing and revising support. This research was supported by the National Center of Competence in Research (NCCR) Affective sciences financed by the Swiss National Science Foundation (n° 51NF40-104897) and hosted by the University of Geneva, SD was also funded by the Swiss Center for Affective Sciences, and by the Department of Radiology, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), as well as the Boninchi Foundation in Geneva.

- Emotion recognition differ according to different pedagogical practices
- Traditionally-schooled children show higher fear recognition sensitivity
- Montessori schoolchildren integrate more social context and are positively biased

EMOTION RECOGNITION DEVELOPMENT:

Preliminary evidence for an effect of school pedagogical practices.

While emotion recognition is shaped through social interactions from a child's early years through at least late adolescence, no emphasis has thus far been given to the effects of daily experiences at school. We posited that enriched, more diverse, and less competitive social interactions fostered by some pedagogical practices may contribute to emotion recognition processes in children. Here, we investigated differences in emotion recognition among schoolchildren experiencing the Montessori versus traditional practices. Children performed two tasks; one measuring the impact of social context on fear-surprise perception, and one measuring their bias toward happiness or anger. Results suggest that children experiencing traditional practices show a higher sensitivity to fear-recognition, while children attending Montessori schools show a higher integration of social cues and perceive expressions of happiness for longer durations. Such preliminary findings call for replication and further research to determine which pedagogical features from the Montessori method may explain these effects.

Keywords: Emotion recognition development; Social Context; School pedagogical practices; Montessori Education

Research concerning the role of emotions in school-related social behavior, well-being, and academic performance, as well as the role of school interventions on the development of social and emotional competencies has substantially grown in the last two decades (see e.g., Nathanson et al., 2016; Pekrun & Linnenbrink-Garcia, 2014). Emotion recognition abilities, broadly defined here as the way individuals perceive, identify, categorize, or interpret others' emotional expressions, are typically considered to be a key process involved in socio-emotional competencies (e.g., Ohl, Fox, & Mitchell, 2013). Culture and early life experiences have been shown to influence such emotion recognition abilities (e.g., Yik, Widen, & Russell, 2013; Gendron, Roberson, van der Vyver, & Barrett, 2014), and evidence indicates that targeted school-based interventions also have the potential to improve them (Garner & Waajid, 2008; Nathanson, Rivers, Flynn, & Brackett, 2016, for example), thereby benefiting scholastic outcomes, personal well-being, as well as long-term positive and cooperative social interactions (Immordino-Yang, Darling-Hammond, & Krone, 2019). For instance, MacCann et al. (2020) recently showed that emotional intelligence can predict academic performance. To the best of our knowledge, no study to date has examined the extent to which unsupervised socio-emotional learning induced by school pedagogical practices contributes to schoolchildren's emotion recognition abilities. Given some divergent characteristics of the Montessori versus traditional pedagogical principles, we aimed at obtaining preliminary evidence that emotion recognition performances differ in children attending Montessori versus traditional schools.

Emotion Recognition Development

The ability to adequately identify and categorize emotions emerges early in life, (see Widen & Russell, 2013) with a protracted maturation through adolescence (e.g., Thomas, De Bellis, Graham, & LaBar, 2007) and life-long modifications (Ruffman, Henry, Livingstone, & Phillips, 2008). While happiness is the first emotion that infants easily recognize (Herba & Phillips, 2004; Palama, Malsert, & Gentaz, 2018), the ability to recognize fear, surprise, and disgust improves from 3 to 10 years of age (Coenen, Aarnoudse, Huitema, Braams, & Veenstra, 2013; Widen & Russell, 2003, 2013), suggesting that some emotions require more cognitive development and/or more complex socio-emotional experiences to be learned. Emotion recognition not only requires us to track emotional cues from the face, voice, and body, but to also integrate contextual information for the attribution of an emotion to be made (e.g., Mumenthaler & Sander, 2012; 2019).

Contextual Social Cues

It takes more than perceiving an isolated facial, vocal, or postural expression to recognize an emotion: There is growing evidence indicating that the very process of emotion recognition is integrating contextual information (such as co-occurring body cues or social information) in order to identify an emotion (Aviezer et al., 2017; Mumenthaler & Sander, 2019). Evidence suggests that some contextual effects in emotion recognition may even automatically take place in adults (e.g., Aviezer et al., 2011; Leitzke & Pollak, 2016; Mumenthaler & Sander, 2015). The contextual effects are particularly observed when the to-be-recognized expressions are ambiguous such as when they are perceived as expressing both fear and surprise (see Mumenthaler & Sander, 2015; 2019). Studies in children even suggest that children rely more on social cues than facial cues to efficiently recognize an emotion, and that they look longer at co-occurring contextual cues to identify emotions such as fear, surprise, or disgust than adults (Durand, Gallay, Seigneuric, Robichon, & Baudouin, 2007; Widen & Russell, 2010). The efficiency of such processing increases with age (e.g., Theurel et al., 2016) and is modulated by targeted training in children aged 3-12 years (for a review, see Theurel & Gentaz, 2015).

Positivity/Negativity Bias

Studies propose a process of probabilistic learning across development that improves emotion recognition: Associations experienced over time cumulate to guide selective attention (Plate, Fulvio, Shutts, Green, & Pollak, 2018). While babies and infants look longer at happy faces than angry ones (Farroni, Menon, Rigato, & Johnson, 2007; Grossmann, Striano, & Friederici, 2007), adults seem to exhibit the opposite pattern, even cross-culturally (Marinetti, Mesquita, Yik, Cragwall, & Gallagher, 2012). This is sometimes interpreted as a “threat advantage”, an effect where dangerous cues (e.g., angry faces) are more salient and thus processed longer than safer cues (e.g., happy faces). Interestingly, a recent study reported developmental changes (8 to 23 years of age) in the brain networks subserving salience detection and cognitive control of emotion recognition, fearful and angry faces being subject to more considerable functional reorganization (Zhang, Padmanabhan, Gross, & Menon, 2019). The developmental shift in emotional valence perception (Kauschke, Bahn, Vesker, & Schwarzer, 2019) parallels the calibration of threat perception across development, with a potential cascading effect from early childhood to late adolescence. In fact, the miscalibration of threatening signals can lead to an overcautious attentional bias toward negative emotional stimuli, as is the case with social anxiety (Maoz et al., 2016) or depression (Gollan et al., 2016), which typically emerges during adolescence (Siegel & Dickstein, 2012). In addition, studies on early experiences, such as exposure to family violence or acute adversity, align with this idea. Indeed, they report a link between early-life adversity, an attentional bias toward negative stimuli (e.g., Dannlowski et al., 2013) or threat, and a higher level of anxiety at an older age (Briggs-Gowan et al., 2015). Importantly, the capacity to assign valence to a stimuli can be biased toward positive or negative emotional stimuli (positive/negative bias) and has a direct impact on the person’s interpretation and handling of a situation (Moser, Hajcak, Huppert, Foa, & Simons, 2008). In fact, a bias toward positive stimuli (positivity bias) is related to an increase in positive emotion and better regulation of negative emotions, as well as predicts an individual's resilience to stress (Thoern, Grueschow, Ehlert, Ruff, & Kleim, 2016; Van Bockstaele et al., 2018).

Effective School Practices

If early social experiences modulate an individual's emotion recognition abilities, one can then hypothesize that social interactions experienced at school also contribute to the development of emotion recognition competencies. Children from 4 to 12 years of age spend at least six hours a day in school environments, which are essentially social settings. It is therefore crucial to understand the impact that school pedagogical practices can have not only on children's academic outcomes but also on their socio-emotional competencies, such as emotion recognition abilities, that can, in turn, also predict academic performance (MacCann et al., 2020).

So far, the few studies that have linked education with emotion recognition skills have focused on the education level. These studies reported that students with a higher education level perform better on emotion recognition tasks (Wolfgang & Cohen, 1988; Mill, Allik, Realo, & Valk, 2009; Trauffer, Widen, & Russell, 2013) and show differential brain activation in emotion-related neural substrates (e.g., the amygdala; Demenescu et al., 2014). It could be that growing up in dense social contexts, such as those inherent to educational environments where social interactions are intense and diverse, offer unsupervised learning of emotion recognition (Huelle, Sack, Broer, Komlewa, & Anders, 2014). If true, schoolchildren experiencing enriched social environments and fostered peer-to-peer interactions throughout their daily school practices would show different emotion recognition capacities than schoolchildren of the same age experiencing less diverse social practices. As a "case study", we compared two school pedagogical practices: the Montessori method and traditional practices. Both can be of high quality, but vary in how unsupervised learning and social interactions take place both quantitatively (e.g., amount of time allocated for social interactions), and qualitatively (e.g., with respect to the form and diversity of the social interactions within the environment).

The so-called traditional schools have quite homogeneous pedagogical practices as described by the local policies about school curricula. Schoolchildren typically (i) interact out-class with peers, mainly during recess (twice per day, for 20-30 minutes), and are otherwise asked to work individually at their desk for most of their time; (ii) are in class environments with peers of a similar age lead by one teacher at a time; (iii) receive formal assessments with grades and receive punishments for their behavior (e.g., class exclusion, extra hours after school). Within a typical competitive class climate (Hayek, Toma, Guidotti, Oberle, & Butera, 2017; Hayek, Toma, Oberle, & Butera, 2015), children may undermine their emotion recognition capacities, or even bias them. In adults, competitive climates have been shown to bias intra- and inter-group emotion recognition (Lazerus, Ingbretsen, Stolier, Freeman, & Cikara, 2016). Furthermore, school-related anxiety is also often reported in students experiencing traditional pedagogical practices (Steinmayr, Crede, McElvany, & Wirthwein, 2015; Briggs-Gowan et al., 2015). In adults, anxiety increases threat sensitivity (i.e., more fear or anger perceived; Meyer & Gawlowska, 2017; Notebaert et al., 2018; Proudfit, Inzlicht, & Mennin, 2013).

On the other hand, the Montessori practices, when implemented with high fidelity, provide schoolchildren (i) with in-class peer interactions by keeping the teacher-to-children ratio low so they communicate during on-going work all day long, or share learning moments in small groups; (ii) with multi-grade classrooms (i.e., 3-6 years old, 6-9 years old, 9-12 years old stay together); (iii) with no grades or punishments

(Denervaud, Knebel, Hagmann, & Gentaz, 2019; Lillard, 2011, 2012; Lillard, 2019; Marshall, 2017; Montessori, 1936; Rathunde, 2001). Montessori-schooled children experience a higher percentage of social interactions with peers, as well as more individualized exchanges with their teacher (Baines & Snortum, 1973; Hojnoski et al., 2008). They may learn more from their peers, not only at a cognitive level, but also at a social-emotional one, potentially increasing their emotion recognition capacity. However, there is scarce and indirect evidence suggesting socio-emotional advantages for schoolchildren experiencing Montessori practices compared with their peers exposed to traditional practices. More precisely, Montessori students were reported to be better at self-regulation, to have more positive social interactions, better conflict-monitoring skills, and a higher well-being at school (Ervin, Wash, & Mecca, 2010; Alves et al., 2015; Denervaud et al., 2019; Lillard & Else-Quest, 2006; Lillard et al., 2017a; Rathunde & Csikszentmihalyi, 2005). This form of self-monitoring and resilient behavior is related to a bias toward positive information in adults (Thoern, Grueschow, Ehlert, Ruff, & Kleim, 2016; Van Bockstaele et al., 2018).

Taken together, these elements suggest that school pedagogical practices have an impact on the development of socio-emotional processes such as emotion recognition.

Hypotheses

The aim of this study was to investigate how the differences in social interactions and environments experienced over the last 6 years would impact 8-12 years old's emotion recognition abilities. In particular, we were interested in investigating (i) the processing of social emotional cues displayed in context; and (ii) the bias toward positive emotion. Accordingly, we measured these effects by adapting two tasks. First, we adapted an existing social appraisal task for 8- to 12-year-old schoolchildren to test the influence of social contextual cues on the categorization of an ambiguous facial expression (50% surprise-50% fear morphed face). It may be that children who learn more from peers on a daily basis could be particularly sensitive to children's expressions. Therefore, both child and adult faces were used as stimuli for the social context. Second, we used an offset reaction time task of dynamic emotional changes (morphing video clips) to test for the presence of any positivity bias. We hypothesized that, based on the indirect preliminary evidence previously cited, schoolchildren experiencing Montessori practices compared with traditional practices would (i) be more efficient at integrating social cues in the emotion recognition process; (ii) be more biased toward positive emotional faces; and (iii) show a lower fear recognition sensitivity.

Methods

Study Sites and Participants

Selection Criteria for the Schools

Montessori private schools (4 classes from 3 different schools) were selected following the criteria set by the International Montessori Association (https://montessori-ami.org), and ensured to have a high fidelity in the implementation of their curriculum (Lillard, 2012):

(i) self-directed activities, through the use of Montessori educational materials;
(ii) self-correction and no formal assessments;
(iii) children have the opportunity to work for 3 uninterrupted hours;
(iv) classes with at least 3 different age-levels.

The Swiss public schools were selected based on their rigorous application of the local policies for traditional pedagogical practices, which was observed and discussed with the school Directors. Traditional public schools (3 classes from 2 different schools) were selected in a specific area based on the city's official statistical data on mean salary to only include an upper-class population (as a way to control some family related variables when comparing this group with the group of children attending Montessori schools), and they were controlled for their application of the official local study plan:

(i) lecture-style, adult-driven interactions;
(ii) feedback in the form of grades and summative assessments;
(iii) children are given a break every hour, and mainly interact with their during these breaks;
(iv) one age-level per class.

The teachers' and students' participation was voluntary.

Participants

In total, a subset of 57 children were recruited in the framework of a larger study, which includes neuroimaging and other behavioral measures aimed at evaluating the school environment's impact on a series of psychological processes. The present study was conducted in accordance with the Declaration of Helsinki and with ethics approval from the department of Psychology from the University of XXX. Written parental consent was obtained for each child and informed consent was provided by each adult participant. For this specific study, inclusion criteria were the age of the participants (8-12 years of age) and belonging to one of the two schooling systems for the last 6 years (according to parental report). The Montessori group was composed of

28 children (M_{age} = 10.07, SD = 1.35), 16 boys and 12 girls recruited from 3 different schools. The group of children attending traditional public schools was composed of 29 children (M_{age} = 10.64, SD = 1.02), 15 boys and 14 girls recruited from 3 different schools.

Group Variables. ***Socio-Economic Status (SES).*** Due to local policies in Switzerland, no public Montessori schools exist. In order to control for the fact that the Montessori schools included in this study were all private schools, the selected traditional public schools were located in specific areas to include a disproportionately upper-class population. Parents were also asked to complete a socio-economic questionnaire to assess their education level and professional level. More precisely, in the questionnaire, the parent(s) had to select which of the four options best described their education level (e.g., less than a high school diploma to university level) and professional level (e.g., unemployed to senior executive employee) (Genoud, 2011).

Fluid Intelligence. To account for the effect of intelligence in emotion recognition abilities (Schlegel et al., 2020), fluid intelligence was determined using the black and white short version of the Raven matrices (Raven, Raven, & Court, 2003). The child is presented with a pattern that is missing a piece and is asked to select a piece from several options to complete the given pattern. Children from both school systems showed a similar level of fluid intelligence.

Measures

Each participant completed two separate experimental tasks. Given the exploratory nature of this study and its sample size, we computed sensitivity analyses to determine a priori the critical F and t for the expected effects to ensure a statistical power of 80% (G*Power 3; Faul, Erdfelder, Lang, & Buchner, 2007).

Emotion Categorization with Social Context

The social context task, adapted from Mumenthaler & Sander (2012), was programmed using Matlab with the Psychophysics Toolbox extensions (Brainard, 1997). During the task, the participants were asked to look at a fixation cross on one side of the screen while a social context was displayed on the other side. The social context consisted of a Caucasian front-facing adult or child (male or female) face from the Radboud database (Langner et al., 2010), with either an angry or neutral facial expression (referred to as the "contextual face"). A congruent non-verbal emotional prosody (Banse & Scherer, 1996) was simultaneously played bilaterally to foster the contextual emotion. For the so-called contextual faces of children, the pitch was adapted by an upward transposition (shift in tone) of the adult voice to mimic a child's prosody. After a 50 ms presentation of the fixation cross, it was replaced by a target face. All target faces were of front-facing static Caucasian children with either a neutral or a morphed emotional expression. The morphed expression was always a mixed facial expression composed of 50% surprise and 50% fear (see Figure 1A). In order to create these faces, surprise and fear faces from the CAFE database (LoBue, Baker, & Thrasher, 2018) were morphed at 50% with FantaMorph-Abrosoft. The neutral faces served as a control condition, while the morphed faces were used for the experimental condition. The presentation side was counter balanced across the block, and the total presentation time of the target faces was 2 seconds. After the presentation of the stimuli, the child was asked to select the emotional label corresponding to the target face from 5 different options (fear, surprise, shame, sadness, or "something else"), which were presented in a randomized order between participants. No time limit was set for the response. At the beginning of the task, the child was instructed to "look at the fixation cross and rate the following target picture." Each child performed 2 familiarization trials, followed by 4 blocks of 32 trials, each of which was pseudo-randomized with a controlled number of stimuli for gender, age, and left or right gaze per block (see Fig. SI 1).

Positivity/Negativity Bias

Short morphing video-clips with either an adult or a baby face displaying a continuum of a 100% (pure) happy or anger expression that was gradually changing into a 100% of the other expression (from Korb et al., 2015) were used. Each 5-second clip consisted of 60 frames (see Fig 1B). The child was asked to press the space bar as soon as s/he could no longer perceive the first emotion. Every child performed one familiarization trial of the baby and adult trials, followed by two blocks (randomly starting by adult or baby) of 24 video-clips of faces (see Fig. SI 2), with a balanced gender presentation.

Procedure

The experiments were conducted *in situ* at the end of the school year, where children performed the tasks in a random order on a laptop with headphones, in a separate quiet room.

Results

Statistical analyses were performed with R (R Core Team, 2014) and with the Jamovi open-source software (Version 0.9).

Group Variables

Participants with missing data for one of the two tasks (n=9), outside of the target age (2 SD from the mean; n=4) or with low SES (n=1) or low fluid intelligence (n=1) were excluded from the analysis. In total, data from 57 children were analyzed. Regarding the fluid intelligence measure, correct answers for the Raven matrices task (PM-47) were summed (maximum 36 points) and reported as a single score for each participant (missing data from one participant). The socio-economic status (SES) was rated from the parental questionnaires; answers were summed (maximum score is 8) and scores were averaged when both parents responded (missing data from one participant). All scores were normalized. Three independent-sample t-tests were conducted comparing the age, fluid intelligence, and SES scores of the Montessori versus the traditional schoolchildren. There were no significant differences in the scores (all p's > 0.05; see table SI 1), suggesting a between-group homogeneity in these variables.

Emotion Categorization with Social Context

The responses to the adapted social appraisal task were summed per conditions over all trials for each participant. Two repeated measures analyses of variance (ANOVA) were conducted on the means of each condition to examine the effects of the contextual face (child versus adult), the contextual emotion (anger versus neutral), and the type of schooling system (Montessori versus traditional) on emotion recognition (fear, surprise, shame, sadness, "something else") of the target face; one analysis was performed on the control condition (neutral target face), and the other, on the morphed fear-surprise target face (experimental condition) (see Fig. SI 3 for the mean number of responses per condition). For the experimental condition, we established the critical F accounting for an interaction between responses and context as 1.60 based on a sensitivity analysis with a power of at least 80% (Faul et al., 2007).

For the control condition, in which the target was a neutral face and the contextual faces could be either a neutral or an angry face, no significant between-group effect was observed (see Fig. SI 4, table SI 2). For both the control and the experimental conditions, there was no significant difference observed between the adult and children faces (see Fig. SI 5).

For the experimental condition, where the target face was morphed to express 50% fear and 50% surprise, there was a significant interaction between the school pedagogical practices experienced by the children and the responses given, $F(4,220) = 2.93$, $p = 0.022$, $\eta p^2 = 0.05$[1]. The actual F-value was higher than the critical F computed, confirming the reliability of our measure (Faul et al., 2007). This was further confirmed by a reliable post-hoc power of 86.6% (Fig. 2). The results indicated that schoolchildren exposed to Montessori pedagogical practices perceived more "surprise" in the ambiguous faces than traditionally-schooled students did when the social context was neutral ($t(220) = 4.00$, $p_{Tukey} = 0.013$, $d = 0.31$). However, when the social context displayed an angry face, no significant difference was observed for the "surprise" responses between the two groups. In addition, for both angry and neutral contexts, schoolchildren exposed to traditional practices significantly reported more "fear" than "surprise" in the ambiguous faces ($t(295) = 3.77$, $p_{Tukey} = 0.026$, $d = 0.39$; $t(295) = 4.61$, $p_{Tukey} < 0.001$, $d = 0.37$, respectively), a pattern that was not observed in schoolchildren exposed to Montessori practices. Furthermore, independently of the context,

[1] When including all data collected (i.e., adding the data from the children who did not participate in the Positivity/negativity Bias task, N=62), the triple interaction was robustly present, $F(4,220) = 3.27$, $p = 0.012$, $\eta p^2 = 0.05$ (see Table SI 4).

schoolchildren exposed to Montessori pedagogical practices gave significantly less fear responses for the ambiguous faces than the schoolchildren exposed to traditional practices ($t(55) = -2.47$, $p = 0.017$, $d = 0.65$) (see table SI 3 for details).

Positivity/Negativity Bias

We computed a score for the positivity bias using the offset reaction times (RT) of the pooled data from both the baby and adult conditions (i.e., starting with an anger expression *versus* starting with a happy expression). By computing the reaction time for each condition, we derived individual differences in positive emotion perception (*positivity bias [ms]* = $RT_{happy} - RT_{angry}$). An independent t-test was conducted to compare the positivity bias between the Montessori and traditional schoolchildren. We established a critical t of 1.67 based on a sensitivity analysis with a power of 80%. The results indicated a significant positivity bias in favor of the Montessori schoolchildren (M = 80.4 [ms], SD = 320 [ms]) compared to the traditional schoolchildren (M = -83.3 [ms], SD = 276 [ms]), $t(55) = 2.07$, $p = 0.043$. The actual t-value was higher than the critical t computed a priori (2.07 > 1.67), suggesting a reliable measured effect. As an additional measure of positive versus negative sensitivity, children with either a negative (< 0 [ms]) or a positive bias (> 0 [ms]) were classified into two groups, “positively” or “negatively” biased (Fig. 3). An independent Chi-square test was computed to compare the frequency of positively *versus* negatively biased schoolchildren in the Montessori and traditional systems. A significant interaction was found, X^2 (1, N = 57) = 9.27, p = 0.002[2]. The schoolchildren enrolled in the Montessori system were more likely to present a positivity bias (67.9%) than the traditional schoolchildren (27.6%) (Fig. SI 6).

Discussion

This study was a first attempt at investigating whether the social environment, as operationalized by school pedagogical practices, has the potential to modulate emotion recognition in children. We tested 57 children (8-12 years old) experiencing either Montessori or traditional practices for at least the past six years (as reported on the parental questionnaire). We first measured individual sensitivity to contextual social cues in fear perception. Second, we investigated the bias toward positive or negative emotional facial expressions.

Results from the social context task suggest that the contextual cues were integrated differently depending on the pedagogical practices the schoolchildren experienced. Emotion attribution for the ambiguous target face (morphed to contain 50% fear and 50% surprise) was found to be different in the angry *versus* neutral social contexts for schoolchildren experiencing Montessori practices: these children attributed less "surprise" when exposed to the angry versus neutral social context. No evidence was found for such an integration of social cues (i.e., differential recognition pattern according to the context) in children experiencing traditional practices. Interestingly, these children attributed more "fear" than "surprise" to the ambiguous faces when the context displayed an angry face and a neutral face, suggesting a higher fear-recognition

[2] The results are significantly different even when introducing more data (adding participants that had not participated in the the emotion categorization with social context task); X^2 (1, N = 61) = 8.70, p = 0.003.

sensitivity in these children compared with those attending Montessori schools. Taken together, these results suggest that the integration of social cues in the process of emotion recognition is modulated by the early social environment. Although further research is needed to reach such a conclusion, it is possible that daily enriching social interactions, which are promoted in the Montessori method, may hasten the maturation of the contextual cue integration (training-effect, such as is found in unsupervised learning, see Huelle et al., 2014). Furthermore, the higher fear sensitivity and related threat calibration in schoolchildren exposed to traditional practices may tentatively reflect their experience of a more competitive environment (through grading for example, Hayek et al., 2017), or less peer interactions during learning. These effects may have an impact on the underlying flexibility in cognitive processes through daily cumulative social experiences, thereby potentially causing long-term effects on social behaviors (van Duijvenvoorde, Peters, Braams, & Crone, 2016) and undermining the integration of social contexts that could lead to suboptimal interpersonal relationships (Maoz et al., 2016). It would be interesting to test, in a future study, whether systematically introducing more peer-to-peer working interactions on a daily basis in traditional practices would engender an increase in social cue integration. In fact, when working together, schoolchildren need to coordinate their goals, overcome conflicts, develop their theory of mind mechanisms, and regulate their emotions (Ainsworth & Baumeister, 2013; Domberg, Koymen, & Tomasello, 2018). All these aspects may benefit socio-emotional skills. Finally, to confirm our preliminary findings, studying the response patterns of younger children (i.e., with less experience in each pedagogical method) and tracking their development, within a longitudinal and randomized framework, seems necessary.

In addition, results from the second experiment also suggest a different pattern of emotion processing for children who attended Montessori *versus* traditional schools. More specifically, in the task where children had to notice when the dynamical facial expression changed, schoolchildren attending the Montessori schools perceived happy expressions for a longer duration than they perceived angry expressions, an effect that we can refer to as a positivity bias. In contrast, anger expressions were perceived for a longer duration than happy expressions in schoolchildren exposed to traditional practices. These results suggest that the school pedagogical practices experienced by schoolchildren contribute to the emergence of such a positive/negative bias.

The pattern observed in Montessori-schooled children is consistent with previous studies reporting a relatively high positive affect at - and toward - school in children experiencing Montessori practices (Lillard & Else-Quest, 2006; Rathunde & Csikszentmihalyi, 2005; Denervaud et al., 2019; Lillard et al., 2017b). A bias toward positive stimuli, which has been shown in many paradigms (Pool et al., 2016), can be influenced by a current positive mood (e.g., Wadlinger & Isaacowitz, 2006), and impact emotion regulation that can lead to more positive social interactions (Thoern, Grueschow, Ehlert, Ruff, & Kleim, 2016; Van Bockstaele et al., 2018). Conversely, the pattern observed in children attending traditional schools parallels adults' longer looking times at angry faces compared to happy faces (Marinetti et al., 2012), and suggests a precocious "threat advantage" bias (Marinetti et al., 2012; Martinez, Falvello, Aviezer, & Todorov, 2016). However, a too large bias toward negative stimuli may have deleterious implications on emotion regulation or affective disorders such as anxiety or depression (Bar-Haim et al., 2007; Bone et al., 2019).

Future studies could test whether the school climate, and its direct impact on students' well-being (Steinmayr, Heyder, Naumburg, Michels, & Wirthwein, 2018), may be an underlying feature shaping attentional bias. More research is clearly needed to replicate these effects and understand their origins, using multiple tasks, larger populations, randomized designs, and manipulating variables that relate to school-induced mood and anxiety.

Consistent with the proposal that emotion recognition ability depends on education, our research provides preliminary evidence suggesting that this ability not only depends on the level of education (e.g., Trauffer et al., 2013) but also on the pedagogical practices. Crucially, some specific pedagogical features from the Montessori education may explain such differences, but cannot be inferred from our study. The measured effects are certainly not specific to the Montessori pedagogy, but rather to some variables found in this pedagogy (e.g., a focus on collaborative learning with peers, multi-grade classrooms). From a different perspective, recent results showing that the more a student talked in class, the better they performed in a reading literacy test are inspiring in this respect (Sedova et al., 2019). An interesting approach would be to compare specific pedagogical practices by systematically as well as empirically testing them using designs manipulating specific variables such as the diversity of social interactions during learning hours (e.g., age and diversity of the children who interact and the social contexts in which the interactions take place), moods induced, feedback given, or active collaborative learning. Although new research should test whether these results can be replicated, and if they are directly caused by the school environment and/or by other factors, such as family-related variables (Castro, Halberstadt, Lozada, & Craig, 2015), our findings suggest that the early social environment influences emotion recognition mechanisms. With respect to theories of emotion, our results are particularly compatible with appraisal theory's account of emotion recognition (e.g., Sander et al., 2007) as well as with theories focusing on emotion attribution (see Widen, 2013). Both account for the contextual effects (see Aviezer, Ensenberg, & Hassi, 2017) and for the existence of environmental modulators of emotion recognition (see Trauffer, Widen, & Russell, 2013) in children.

Disclosure Statement

No potential conflict of interest was reported by the authors.

References

Aviezer, H., Bentin, S., Dudareva, V., & Hassin, R. (2011). Automaticity in contextualized emotion perception. Emotion, 11, 1406-1414.

Aviezer, H., Ensenberg, N., & Hassin, R. R. (2017). The inherently contextualized nature of facial emotion perception. Current opinion in psychology, 17, 47-54.

Ainsworth, S. E., & Baumeister, R. F. (2013). Cooperation and fairness depend on self-regulation. *Behav Brain Sci, 36*(1), 79-80. doi:10.1017/S0140525X12000696

Alves, F. L., Ribeiro, M. A., Hahn, R. C., de Melo Teixeira, M., de Camargo, Z. P., Cisalpino, P. S., & Marini, M. M. (2015). Transposable elements and two other molecular markers as typing tools for the genus Paracoccidioides. *Med Mycol, 53*(2), 165-170. doi:10.1093/mmy/myu074

Baines, M. R., & Snortum, J. R. (1973). A time-sampling analysis of Montessori versus traditional classroom interaction. . *Journal of Educational Research, 66*, 313-316.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J Pers Soc Psychol, 70*(3), 614-636.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. Psychological Bulletin, 133, 1–24. http://dx.doi.org/10.1037/0033-2909.133.1.1

Bone, J. K., Lewis, G., Button, K. S., Duffy, L., Harmer, C. J., Munafo, M. R., . . . Lewis, G. (2019). Variation in recognition of happy and sad facial expressions and self-reported depressive symptom severity: A prospective cohort study. *J Affect Disord, 257*, 461-469. doi:10.1016/j.jad.2019.06.025

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis, 10*(4), 433-436.

Briggs-Gowan, M. J., Pollak, S. D., Grasso, D., Voss, J., Mian, N. D., Zobel, E., . . . Pine, D. S. (2015). Attention bias and anxiety in young children exposed to family violence. *J Child Psychol Psychiatry, 56*(11), 1194-1201.

Castro, V. L., Halberstadt, A. G., Lozada, F. T., & Craig, A. B. (2015). Parents' Emotion-Related Beliefs, Behaviors, and Skills Predict Children's Recognition of Emotion. *Infant Child Dev, 24*(1), 1-22. doi:10.1002/icd.1868

Coenen, M., Aarnoudse, C., Huitema, R., Braams, O., & Veenstra, W. S. (2013). *Development of Facial Emotion Recognition in Childhood: Age-related*

Differences in a Shortened Version of the Facial Expressions of Emotion - Stimuli and Tests. Paper presented at the INS Midyear meeting.

Dannlowski, U., Kugel, H., Huber, F., Stuhrmann, A., Redlich, R., Grotegerd, D., . . . Suslow, T. (2013). Childhood maltreatment is associated with an automatic negative emotion processing bias in the amygdala. *Hum Brain Mapp, 34*(11), 2899-2909. doi:10.1002/hbm.22112

Demenescu, L. R., Stan, A., Kortekaas, R., van der Wee, N. J., Veltman, D. J., & Aleman, A. (2014). On the connection between level of education and the neural circuitry of emotion perception. *Front Hum Neurosci, 8*, 866. doi:10.3389/fnhum.2014.00866

Denervaud, S., Knebel, J. F., Hagmann, P., & Gentaz, E. (2019). Beyond executive functions, creativity skills benefit academic outcomes: Insights from Montessori education. *PLoS One, 14*(11), e0225319. doi:10.1371/journal.pone.0225319

Domberg, A., Koymen, B., & Tomasello, M. (2018). Children's reasoning with peers in cooperative and competitive contexts. *Br J Dev Psychol, 36*(1), 64-77. doi:10.1111/bjdp.12213

Durand, K., Gallay, M., Seigneuric, A., Robichon, F., & Baudouin, J. Y. (2007). The development of facial emotion recognition: the role of configural information. *J Exp Child Psychol, 97*(1), 14-27. doi:10.1016/j.jecp.2006.12.001

Ervin, B., Wash, P. D., & Mecca, M. E. (2010). A 3-Year Study of Self-Regulation in Montessori and Non-Montessori Classrooms. *Montessori Life*(2).

Farroni, T., Menon, E., Rigato, S., & Johnson, M. H. (2007). The perception of facial expressions in newborns. *Eur J Dev Psychol, 4*(1), 2-13. doi:10.1080/17405620601046832

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods, 39*(2), 175-191.

Flynn, T. M. (1991). Development of social, personal and cognitive skills of preschool children in Montessori and traditional preschool programs. . *Early Child Development & Care, 72*, 117-124.

Garner, P. W., & Waajid, B. (2008). The associations of emotion knowledge and teacher–child relationships to preschool children's school-related developmental competence. *Journal of Applied Developmental Psychology, 29*(2), 89-100. doi:10.1016/j.appdev.2007.12.001

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014). Cultural relativity in perceiving emotion from vocalizations. *Psychol Sci, 25*(4), 911-920. doi:10.1177/0956797613517239

Genoud, P. (2011). Indice de position socioéconomique (IPSE) : un calcul simplifié.

Gollan, J. K., Hoxha, D., Hunnicutt-Ferguson, K., Norris, C. J., Rosebrock, L., Sankin, L., & Cacioppo, J. (2016). Twice the negativity bias and half the positivity offset: Evaluative responses to emotional information in depression. *J Behav Ther Exp Psychiatry, 52*, 166-170. doi:10.1016/j.jbtep.2015.09.005

Grossmann, T., Striano, T., & Friederici, A. D. (2007). Developmental changes in infants' processing of happy and angry facial expressions: a neurobehavioral study. *Brain Cogn, 64*(1), 30-41. doi:10.1016/j.bandc.2006.10.002

Hayek, A. S., Toma, C., Guidotti, S., Oberle, D., & Butera, F. (2017). Grades degrade group coordination: deteriorated interactions and performance in a cooperative motor task. *European Journal of Psychology of Education, 32*(1), 97-112. doi:10.1007/s10212-016-0286-9

Hayek, A. S., Toma, C., Oberle, D., & Butera, F. (2015). Grading Hampers Cooperative Information Sharing in Group Problem Solving. *Social Psychology, 46*(3), 121-131. doi:10.1027/1864-9335/a000232

Herba, C., & Phillips, M. (2004). Annotation: Development of facial expression recognition from childhood to adolescence: behavioural and neurological perspectives. *J Child Psychol Psychiatry, 45*(7), 1185-1198. doi:10.1111/j.1469-7610.2004.00316.x

Hojnoski, R. L., Margulies, A. S., Barry, A., Bose-Deakins, J., Sumara, K. M., & Harman, J. L. (2008). Analysis of two early childhood education settings: Classroom variables and peer verbal interaction. *Journal of Research in Childhood Education, 23*, 193-209.

Huelle, J. O., Sack, B., Broer, K., Komlewa, I., & Anders, S. (2014). Unsupervised learning of facial emotion decoding skills. *Front Hum Neurosci, 8*, 77. doi:10.3389/fnhum.2014.00077

Immordino-Yang, M.-H., Darling-Hammond, L., & Krone, C. R. (2019). Nurturing Nature: How Brain Development Is Inherently Social and Emotional, and What This Means for Education. *Educational Psychologist*, 1-20. doi:10.1080/00461520.2019.1633924

Kauschke, C., Bahn, D., Vesker, M., & Schwarzer, G. (2019). The Role of Emotional Valence for the Processing of Facial and Verbal Stimuli-Positivity or Negativity Bias? *Front Psychol, 10*, 1654. doi:10.3389/fpsyg.2019.01654

Korb, S., Malsert, J., Rochas, V., Rihs, T. A., Rieger, S. W., Schwab, S., . . . Grandjean, D. (2015). Gender differences in the neural network of facial mimicry of smiles--An rTMS study. *Cortex, 70*, 101-114. doi:10.1016/j.cortex.2015.06.025

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion, 24*(8), 1377-1388. doi:10.1080/02699930903485076

Lazerus, T., Ingbretsen, Z. A., Stolier, R. M., Freeman, J. B., & Cikara, M. (2016). Positivity bias in judging ingroup members' emotional expressions. *Emotion, 16*(8), 1117-1125. doi:10.1037/emo0000227

Leitzke, B. T., & Pollak, S. D. (2016). Developmental changes in the primacy of facial cues for emotion recognition. *Dev Psychol, 52*(4), 572-581. doi:10.1037/a0040067

Lillard, A., & Else-Quest, N. (2006). The early years. Evaluating Montessori education. *Science, 313*(5795), 1893-1894. doi:10.1126/science.1132362

Lillard, A. S. (2011). Mindfulness Practices in Education: Montessori’s Approach. *Mindfulness, 2*(2), 78-85. doi:10.1007/s12671-011-0045-6

Lillard, A. S. (2012). Preschool children's development in classic Montessori, supplemented Montessori, and conventional programs. *J Sch Psychol, 50*(3), 379-401. doi:10.1016/j.jsp.2012.01.001

Lillard, A. S. (2019). Shunned and Admired: Montessori, Self-Determination, and a Case for Radical School Reform. *Educational Psychology Review*. doi:doi.org/10.1007/s10648-019-09483-3

Lillard, A. S., Heise, M. J., Richey, E. M., Tong, X., Hart, A., & Bray, P. M. (2017a). Montessori Preschool Elevates and Equalizes Child Outcomes: A Longitudinal Study. *Frontiers in Psychology, 8*. doi:ARTN 1783
10.3389/fpsyg.2017.01783

Lillard, A. S., Heise, M. J., Richey, E. M., Tong, X., Hart, A., & Bray, P. M. (2017b). Montessori Preschool Elevates and Equalizes Child Outcomes: A Longitudinal Study. *Front Psychol, 8*, 1783. doi:10.3389/fpsyg.2017.01783

LoBue, V., Baker, L., & Thrasher, C. (2018). Through the eyes of a child: preschoolers' identification of emotional expressions from the child affective facial expression (CAFE) set. *Cogn Emot, 32*(5), 1122-1130. doi:10.1080/02699931.2017.1365046

MacCann, C., Jiang, Y., Brown, L., Double, K., Bucich, M., Minbashian, A. (2020). Emotional Intelligence Predicts Academic Performance: A Meta-Analysis. 146(2), 150–186. *Psychological Bulletin*.

Maoz, K., Eldar, S., Stoddard, J., Pine, D. S., Leibenluft, E., & Bar-Haim, Y. (2016). Angry-happy interpretations of ambiguous faces in social anxiety disorder. *Psychiatry Res, 241*, 122-127. doi:10.1016/j.psychres.2016.04.100

Marinetti, C., Mesquita, B., Yik, M., Cragwall, C., & Gallagher, A. H. (2012). Threat advantage: perception of angry and happy dynamic faces across cultures. *Cogn Emot, 26*(7), 1326-1334. doi:10.1080/02699931.2011.644976

Marshall, C. (2017). Montessori education: a review of the evidence base. *npj Science of Learning, 2*(1). doi:10.1038/s41539-017-0012-7

Martinez, L., Falvello, V. B., Aviezer, H., & Todorov, A. (2016). Contributions of facial expressions and body language to the rapid perception of dynamic emotions. *Cogn Emot, 30*(5), 939-952. doi:10.1080/02699931.2015.1035229

Meyer, A., & Gawlowska, M. (2017). Evidence for specificity of the impact of punishment on error-related brain activity in high versus low trait anxious individuals. *Int J Psychophysiol, 120*, 157-163. doi:10.1016/j.ijpsycho.2017.08.001

Mill, A., Allik, J., Realo, A., & Valk, R. (2009). Age-related differences in emotion recognition ability: a cross-sectional study. *Emotion, 9*(5), 619-630. doi:10.1037/a0016562

Montessori, M. (1936). *The Secret of Childhood* (1981 ed.). New York: Ballantine.

Moser, J. S., Hajcak, G., Huppert, J. D., Foa, E. B., & Simons, R. F. (2008). Interpretation Bias in Social Anxiety as Detected by Event-Related Brain Potentials. *Emotion, 8*(5), 693-700. doi:10.1037/a0013173

Mumenthaler, C., & Sander, D. (2012). Social appraisal influences recognition of emotions. *J Pers Soc Psychol, 102*(6), 1118-1135. doi:10.1037/a0026885

Mumenthaler, C., & Sander, D. (2015). Automatic integration of social information in emotion recognition. *J Exp Psychol Gen, 144*(2), 392-399. doi:10.1037/xge0000059

Mumenthaler, C., & Sander, D. (2019). Socio-affective inferential mechanisms involved in emotion recognition. In D. Dukes & F. Clément (Eds.), Foundations of affective social learning: Conceptualizing the social transmission of value (pp. 142-164). Cambridge, UK: Cambridge University Press.

Nathanson, L., Rivers, S. E., Flynn, L. M., & Brackett, M. A. (2016). Creating Emotionally Intelligent Schools With RULER. *Emotion Review, 8*(4), 305-310. doi:10.1177/1754073916650495

Notebaert, L., Georgiades, J. V., Herbert, M., Grafton, B., Parsons, S., Fox, E., & MacLeod, C. (2018). Trait anxiety and the alignment of attentional bias with controllability of danger. *Psychol Res*. doi:10.1007/s00426-018-1081-9

Ohl, M., Fox, P., & Mitchell, K. (2013). Strengthening socio-emotional competencies in a school setting: data from the Pyramid project. *Br J Educ Psychol, 83*(Pt 3), 452-466. doi:10.1111/j.2044-8279.2012.02074.x

Pekrun, R., & Linnenbrink-Garcia, L. (2014). International Handbook of Emotions in Education. (2014). *Routledge*.

Palama, A., Malsert, J., & Gentaz, E. (2018). Are 6-month-old human infants able to transfer emotional information (happy or angry) from voices to faces? An eye-tracking study. *PLoS One, 13*(4), e0194579. doi:10.1371/journal.pone.0194579

Plate, R. C., Fulvio, J. M., Shutts, K., Green, C. S., & Pollak, S. D. (2018). Probability Learning: Changes in Behavior Across Time and Development. *Child Dev, 89*(1), 205-218. doi:10.1111/cdev.12718

Pool, E. R., Brosch, T., Delplanque, S., & Sander, D. (2016). Attentional bias for positive emotional stimuli: A meta-analytic investigation. Psychological Bulletin, 142(1), 79-106.

Proudfit, G. H., Inzlicht, M., & Mennin, D. S. (2013). Anxiety and error monitoring: the importance of motivation and emotion. *Front Hum Neurosci, 7*, 636. doi:10.3389/fnhum.2013.00636

Rathunde, K. (2001). Montessori Education and optimal experieunce: a framework for new research. *The NAMTA Journal, 26*(1), 11-43.

Rathunde, K., & Csikszentmihalyi, M. (2005). Middle School Students' Motivation and Quality of Experience: A Comparison of Montessori and Traditional School Environments. *American Journal of Education, 111*(3).

Raven, J., Raven, J. C., & Court, J. H. (2003). Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview. . San Antonio, TX: Harcourt Assessment.

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: implications for neuropsychological models of aging. *Neurosci Biobehav Rev, 32*(4), 863-881. doi:10.1016/j.neubiorev.2008.01.001

Schlegel, K., Palese, T., Mast, M. S., Rammsayer, T. H., Hall, J. A., & Murphy, N. A. (2020). A meta-analysis of the relationship between emotion recognition ability and intelligence. *Cognition & Emotion, 34*(2), 329-351. doi:10.1080/02699931.2019.1632801

Sedova, K., Sedlacek, M., Svaricek, R., Majcik, M., Navaratilova, J., Drexlerova, A., . . . Salamounova, Z. (2019). Do those who talk more learn more? The relationship between student classroom talk and student achievement. *Learning and Instruction, 63*.

Siegel, R. S., & Dickstein, D. P. (2012). Anxiety in adolescents: Update on its diagnosis and treatment for primary care providers. *Adolesc Health Med Ther, 3*, 1-16. doi:10.2147/AHMT.S7597

Steinmayr, R., Crede, J., McElvany, N., & Wirthwein, L. (2015). Subjective Well-Being, Test Anxiety, Academic Achievement: Testing for Reciprocal Effects. *Front Psychol, 6*, 1994. doi:10.3389/fpsyg.2015.01994

Steinmayr, R., Heyder, A., Naumburg, C., Michels, J., & Wirthwein, L. (2018). School-Related and Individual Predictors of Subjective Well-Being and Academic Achievement. *Front Psychol, 9*, 2631. doi:10.3389/fpsyg.2018.02631

Theurel, A., & Gentaz, E. (2015). Entraîner les compétences émotionnelles à l'école. *A.N.A.E, 139*.

Theurel, A., Witt, A., Malsert, J., Lejeune, F., Fiorentini, C., Barisnikov, K., & Gentaz, E. (2016). The integration of visual context information in facial emotion recognition in 5- to 15-year-olds. *J Exp Child Psychol, 150*, 252-271. doi:10.1016/j.jecp.2016.06.004

Thoern, H. A., Grueschow, M., Ehlert, U., Ruff, C. C., & Kleim, B. (2016). Attentional Bias towards Positive Emotion Predicts Stress Resilience. *PLoS One, 11*(3), e0148368. doi:10.1371/journal.pone.0148368

Thomas, L. A., De Bellis, M. D., Graham, R., & LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science, 10*(5), 547-558. doi:10.1111/j.1467-7687.2007.00614.x

Trauffer, N. M., Widen, S. C., & Russell, J. A. (2013). Education and the Attribution of Emotion to Facial Expressions. *Psychological Topics, 22*(2), 237-247.

Van Bockstaele, B., Notebaert, L., MacLeod, C., Salemink, E., Clarke, P. J. F., Verschuere, B., . . . Wiers, R. W. (2018). The effects of attentional bias modification on emotion regulation. *J Behav Ther Exp Psychiatry, 62*, 38-48. doi:10.1016/j.jbtep.2018.08.010

van Duijvenvoorde, A. C., Peters, S., Braams, B. R., & Crone, E. A. (2016). What motivates adolescents? Neural responses to rewards and their influence on adolescents' risk taking, learning, and cognitive control. *Neurosci Biobehav Rev, 70*, 135-147. doi:10.1016/j.neubiorev.2016.06.037

Wadlinger, H. A., & Isaacowitz, D. M. (2006). Positive mood broadens visual attention to positive stimuli. *Motivation and Emotion, 30*(1), 89-101. doi:10.1007/s11031-006-9021-1

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Dev Psychol, 39*(1), 114-128.

Widen, S. C., & Russell, J. A. (2010). Children's scripts for social emotions: causes and consequences are more central than are facial expressions. *Br J Dev Psychol, 28*(Pt 3), 565-581.

Widen, S. C., & Russell, J. A. (2013). Children's recognition of disgust in others. *Psychol Bull, 139*(2), 271-299. doi:10.1037/a0031640

Wolfgang, A., & Cohen, M. (1988). Sensitivity of Canadians, Latin Americans, Ethiopians, and Israelis to Interracial Facial Expressions of Emotions. *International Journal of Intercultural Relations, 12*, 139-150.

Yik, M., Widen, S. C., & Russell, J. A. (2013). The within-subjects design in the study of facial expressions. *Cogn Emot, 27*(6), 1062-1072. doi:10.1080/02699931.2013.763769

Zhang, Y., Padmanabhan, A., Gross, J. J., & Menon, V. (2019). Development of human emotion circuits investigated using a Big-Data analytic approach: Stability, reliability, and robustness. *J Neurosci*.

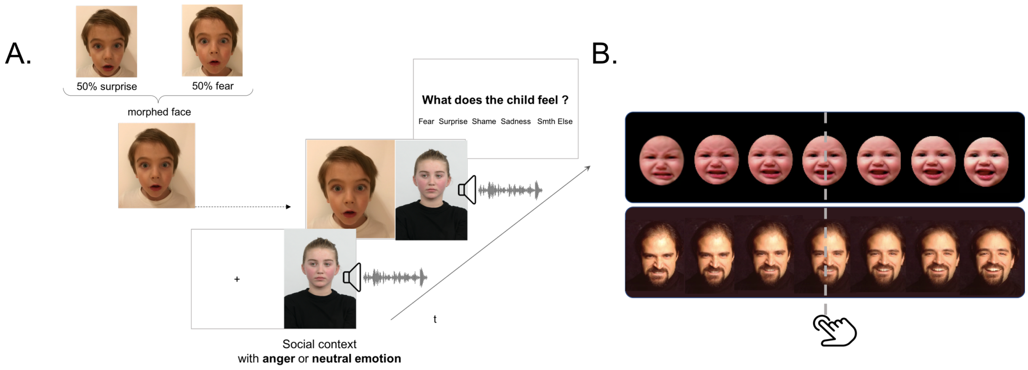


Figure 1. **A.** ***Emotion Categorization with Social Context.*** In this task, schoolchildren were first presented with a social context in the form of an adult or child face expressing either anger or no emotion (neutral) concomitant with a congruent emotional prosody; then, a target face appeared in the form of a neutral child expression or an ambiguous (50% fear-50% surprise morphed) child expression. After the stimuli presentation, schoolchildren were asked to label the emotion of the target face (forced choice between fear, surprise, shame, sadness, or “something else”). **B.** ***Positivity/Negativity Bias.*** In this task, schoolchildren were asked to track a dynamic change in emotions from positive to negative or the opposite (morphing video clips) in baby or adult faces. They were asked to detect when the first emotion (happiness or anger) had stopped.

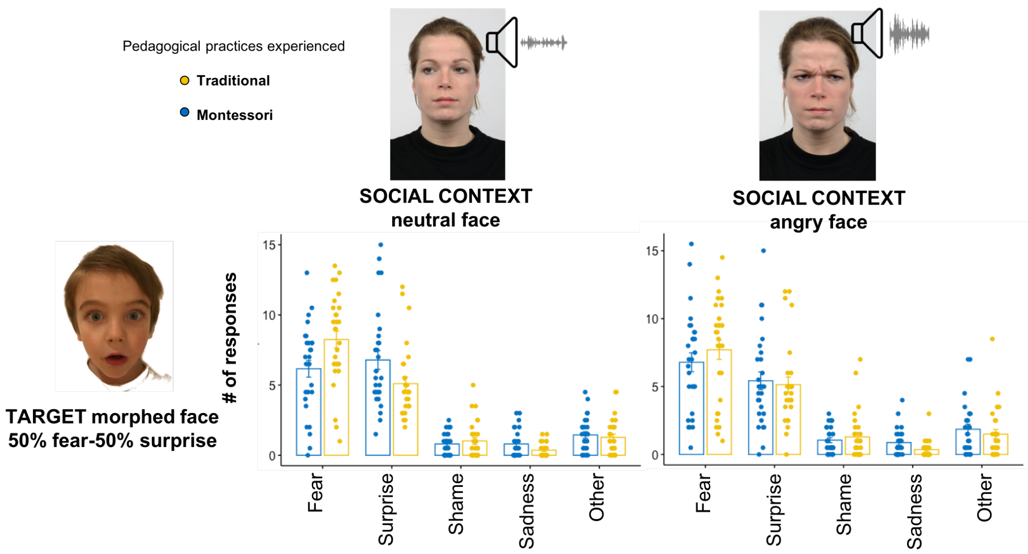


Figure 2. ***Emotion Categorization with Social Context***. Emotion recognition in the ambiguous condition (50% fear-50% surprise morphed face) with either a child or an adult in context, with an angry or neutral facial expression (from the Radboud database) combined with a congruent emotional prosody. The descriptive plot shows the results of the triple interaction (contextual face × responses given × school pedagogy). Error bars represent SE.

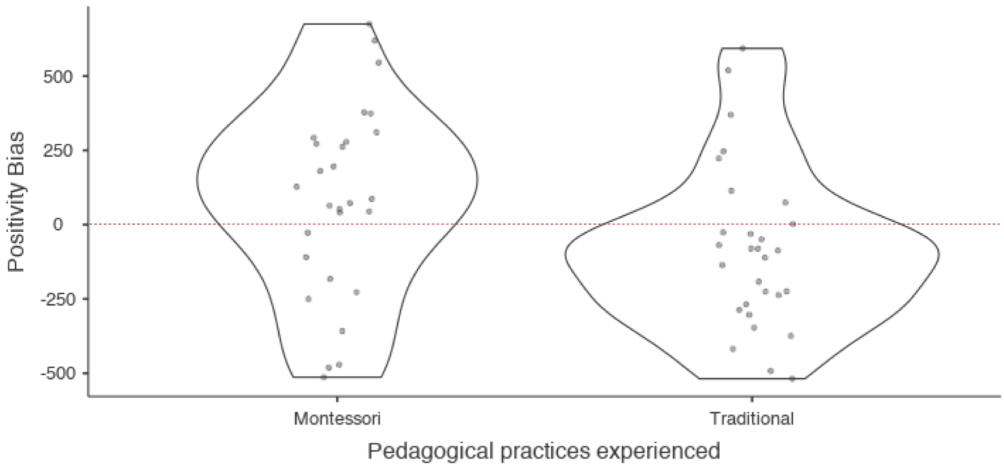


Figure 3. ***Positivity /Negativity Bias.*** The capacity to assign valence to a stimulus can be biased toward positive *versus* negative emotional stimuli, and is measured through the positivity bias: the difference between the amount of time (RT) spent perceiving positive stimuli (here, happy faces) and the amount of time (RT) spent perceiving negative stimuli (here, angry faces).

EMOTION RECOGNITION DEVELOPMENT:

Preliminary evidence for an effect of school pedagogical practices

Solange Denervaud[1,2], Christian Mumenthaler[1,2], Edouard Gentaz[1,2] & David Sander[1,2*]

[1] *Swiss Center for Affective Sciences, Campus Biotech, University of Geneva, 1202 Geneva, Switzerland*

[2] *Connectomics Lab, Department of Radiology, Lausanne University Hospital and University of Lausanne (CHUV-UNIL), 1011 Lausanne, Switzerland*

[3] *Faculté de Psychologie et des Sciences de l'Education (FAPSE), University of Geneva, 1204 Geneva, Switzerland*

*corresponding authors: solange.denervaud@unil.ch, david.sander@unige.ch

SUPPLEMENTARY MATERIAL

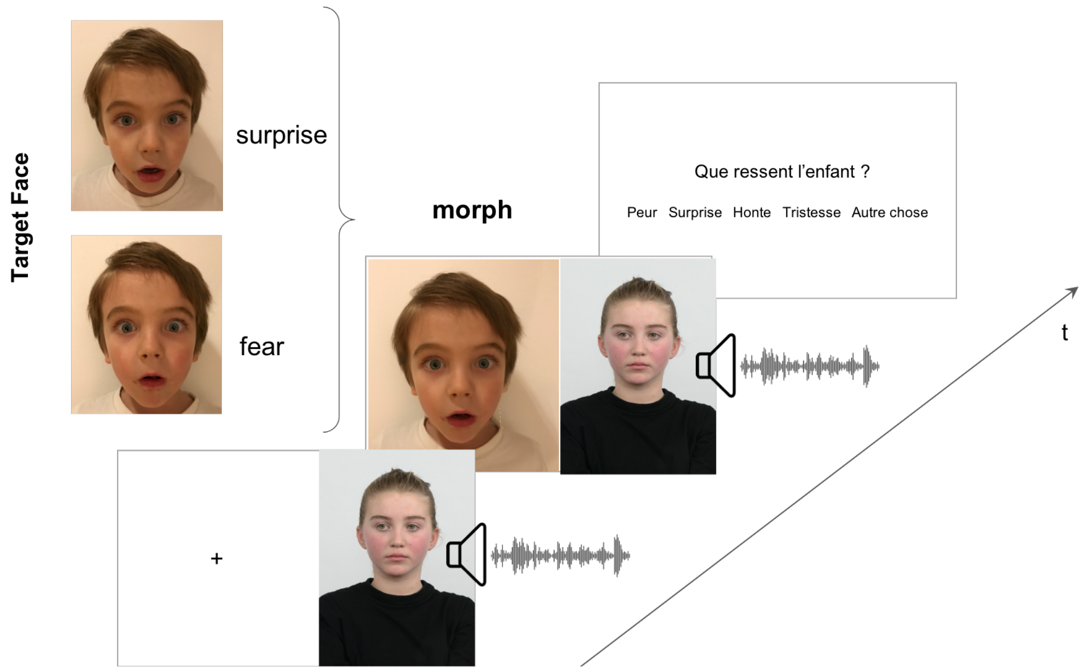


Fig. SI 1 Emotion categorization with Social Context Task. In each trial, a contextual face (from the Radboud database) with a congruent emotional prosody (from Banse & Scherer, 1996) was displayed for 50ms. Subsequently, a target face that was either a morph or a neutral face appeared alongside the contextual face. After the presentation of the stimuli, the child was asked to indicate which emotion the target face expressed using 5 different labels: fear, surprise, shame, sadness, or something else. Due to copyright protection, the pictures featured in this figure were made by the authors, while the pictures we used in the study were taken from the CAFE database. The target stimuli were made of surprise-fear morphed faces and the control ones were of neutral faces.

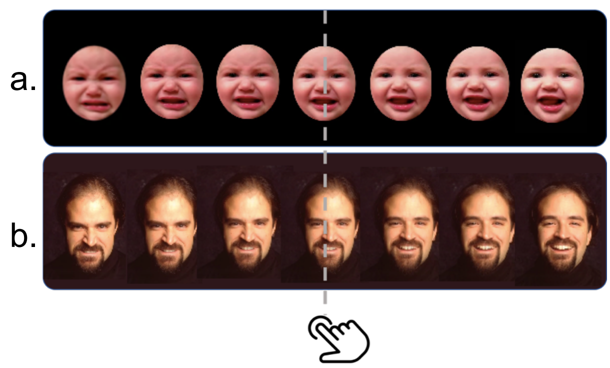


Fig. SI 2 Positivity/Negativity Bias Task. During this task, participants were shown 5-second morphing video clips showing either an angry (displayed on the left) or happy (displayed on the right) facial expression gradually changing to the other expression (happy or angry, respectively). The child was asked to press a key as fast as possible when s/he perceived the initial emotion to be over. The two conditions consisted of either baby faces (panel a) or adult faces (panel b).

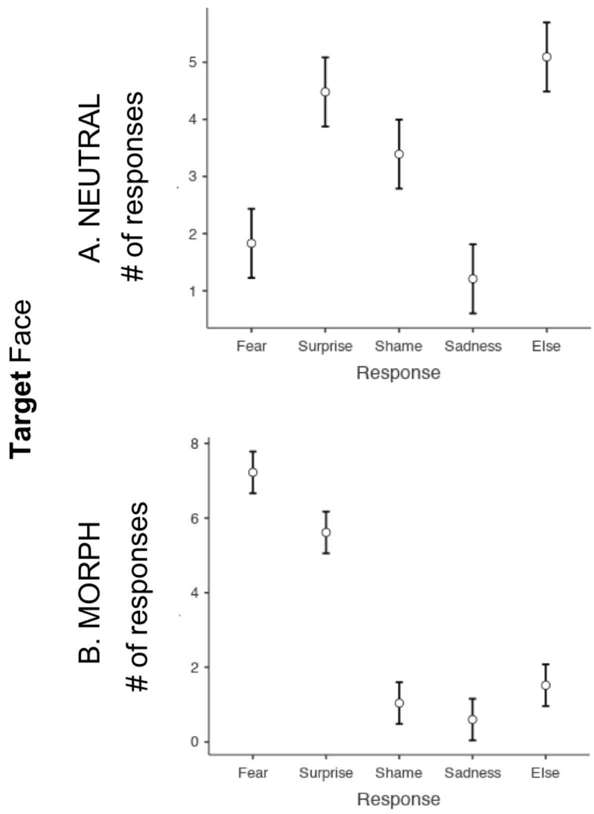


Fig. SI 3 Emotion categorization with Social Context; main effect of responses. Average number of responses given in (A) the control condition, when the target face was neutral, and (B) the experimental condition (B), when the target face was a fear-surprise morphed face. The error bars represent the CI.

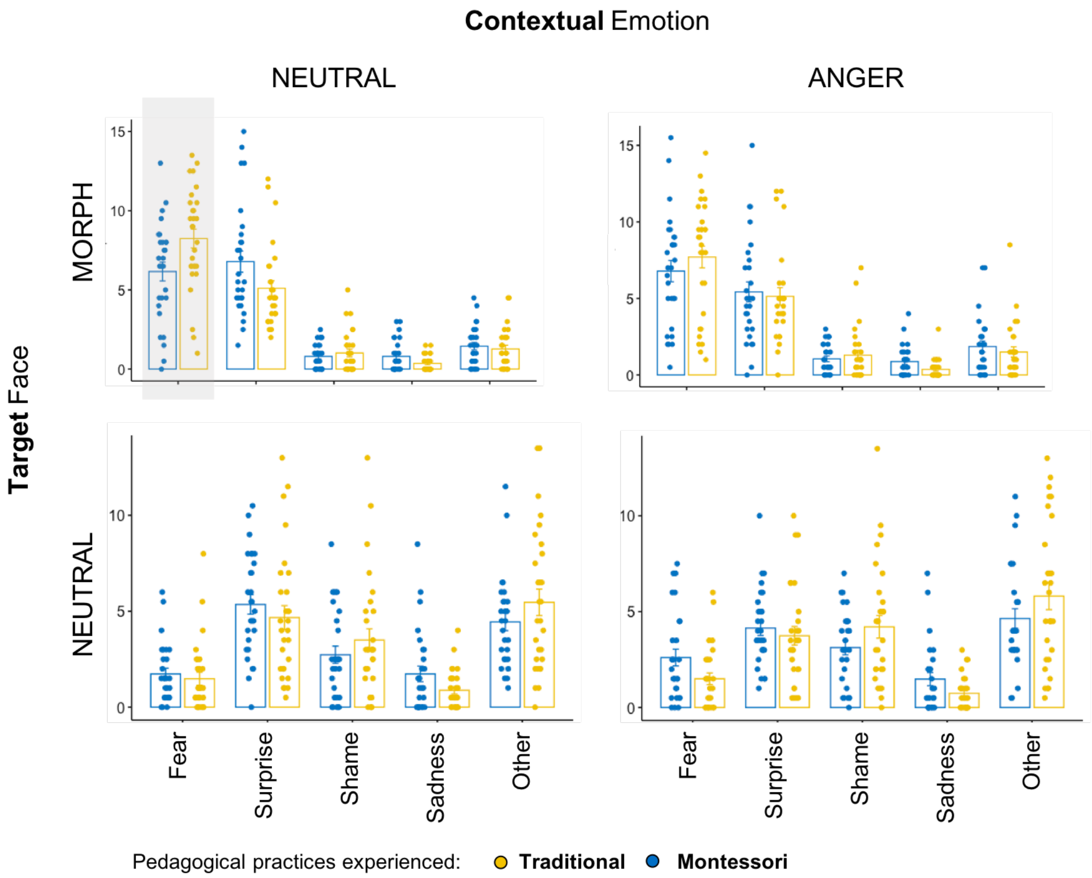


Fig. SI 4 Emotion categorization with Social Context. The figure illustrates the three-way ANOVA interactions in the control and testing conditions on the average number of responses. The top plots depict the testing condition of the fear-surprise morphed face: Children enrolled in traditional systems (in orange) keep on reporting higher fear compared to children enrolled in Montessori schools (blue line) when the contextual face was neutral, while Montessori schoolers report more surprise (highlighted in grey). The bottom plots depict the control condition, when the target face was a neutral child face. M stands for Montessori, T for traditional; the error-bars represent the SE.

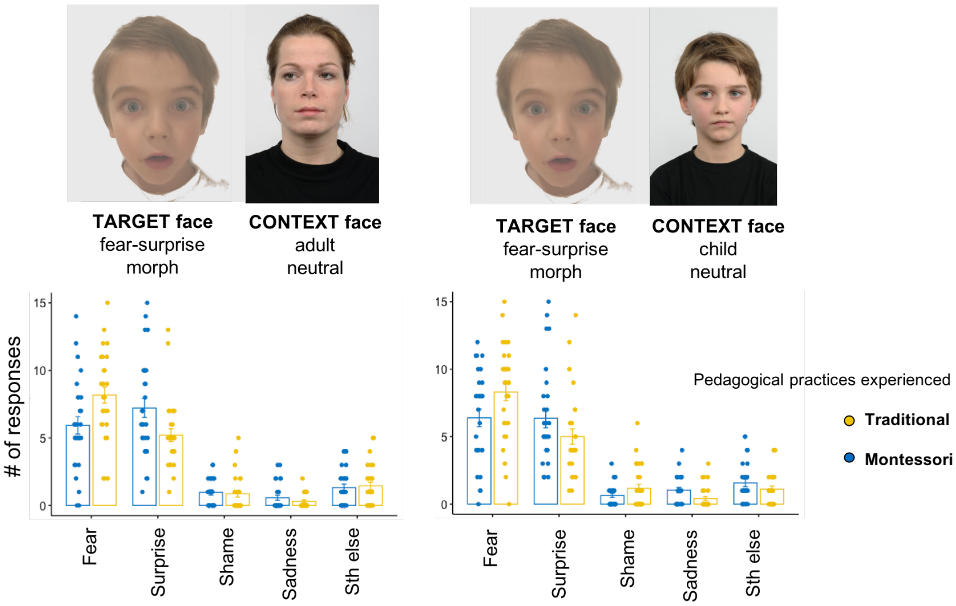


Fig. SI 5 Contextual faces (child or adult) in the morph target experimental condition. The graphs show the average number of responses given when the contextual neutral face was an adult (left) or a child (right). There was no significant effect between conditions (adult *versus* child) for traditional (T, in orange) or Montessori (M, in blue) schoolers. The original pictures for the target faces were taken from the CAFE database, but due to copy-right protection, the current pictures are used to depict at best the experimental design. The contextual faces are taken from the Radboud database. The bars represent the SE.

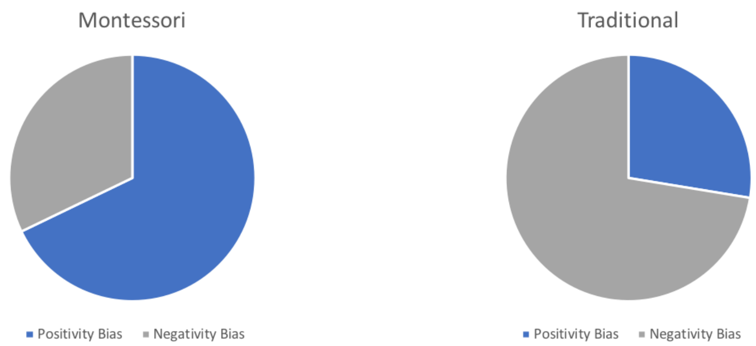


Fig. SI 6 Frequency of positively or negatively biased students.

	Montessori	Traditional	Statistics (t-test or χ^2)	df	p	Cohen's d
# schoolers (girls)	28 (12)	29 (14)	0.17	1	0.68	
Age (SD)	10.07 (1.35)	10.64 (1.02)	-1.82	55	0.08	-0.52
Fluid Intelligence (SD)	33.82 (1.70)	33.00 (2.37)	1.37	54	0.18	0.37
SES (SD)	0.72 (0.09)	0.69 (0.10)	1.11	54	0.27	0.30

Table SI 1 Independent t-tests for the Control variables.

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
Contextual face	2.25e-30	1	2.25e-30	-6.21e−15	1.000	-0.000
Contextual face ✻ System	6.88e-30	1	6.88e-30	-1.90e−14	1.000	-0.000
Residual	-1.99e−14	55	-3.62e−16			
Contextual emotion	2.02e-29	1	2.02e-29	-2.65e−14	1.000	-0.000
Contextual emotion ✻ System	1.26e-30	1	1.26e-30	-1.66e−15	1.000	-0.000
Residual	-4.19e−14	55	-7.62e−16			
Response	**2527.396**	**4**	**631.849**	**23.5870**	**< .001**	**0.300**
Response ✻ System	196.052	4	49.013	1.8297	0.124	0.032
Residual	5893.352	220	26.788			
Contextual face ✻ Contextual emotion	2.25e-30	1	2.25e-30	-1.06e−15	1.000	-0.000
Contextual face ✻ Contextual emotion ✻ System	0.000	1	0.000	0.0000	1.000	0.000
Residual	-1.17e−13	55	-2.12e−15			
Contextual face ✻ Response	13.678	4	3.420	1.5643	0.185	0.028
Contextual face ✻ Response ✻ System	1.977	4	0.494	0.2261	0.924	0.004
Residual	480.918	220	2.186			
Contextual emotion ✻ Response	**100.452**	**4**	**25.113**	**5.9280**	**< .001**	**0.097**
Contextual emotion ✻ Response ✻ System	13.522	4	3.380	0.7980	0.528	0.014
Residual	931.987	220	4.236			
Contextual face ✻ Contextual emotion ✻ Response	0.644	4	0.161	0.0840	0.987	0.002
Contextual face ✻ Contextual emotion ✻ Response ✻ System	6.100	4	1.525	0.7954	0.529	0.014
Residual	421.812	220	1.917			

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
System	5.62e-31	1	5.62e-31	-6.69e−16	1.000	-0.000
Residual	-4.62e−14	55	-8.40e−16			

Note. Type 3 Sums of Squares

Table SI 2 Repeated measures ANOVA for the control condition (neutral face) with contextual face (child, adult), contextual emotion (angry, neutral), and response (fear, surprise, shame, sadness, something else) as within-subject factors, and system (Montessori, traditional) as a between-subject factor.

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
Contextual face	3.51e-28	1	3.51e-28	7.91e-13	1.000	0.000
Contextual face ✻ System	4.00e-29	1	4.00e-29	9.01e-14	1.000	0.000
Residual	2.44e-14	55	4.44e-16			
Contextual emotion	2.75e-29	1	2.75e-29	-3.14e−14	1.000	-0.000
Contextual emotion ✻ System	9.27e-30	1	9.27e-30	-1.06e−14	1.000	-0.000
Residual	-4.82e−14	55	-8.77e−16			
Response	**8263.33**	**4**	**2065.832**	**90.312**	**< .001**	**0.622**
Response ✻ System	203.69	4	50.922	2.226	0.067	0.039
Residual	5032.37	220	22.874			
Contextual face ✻ Contextual emotion	2.75e-29	1	2.75e-29	9.52e-14	1.000	0.000
Contextual face ✻ Contextual emotion ✻ System	1.78e-31	1	1.78e-31	6.15e-16	1.000	0.000
Residual	1.59e-14	55	2.89e-16			
Contextual face ✻ Response	2.40	4	0.601	0.268	0.899	0.005
Contextual face ✻ Response ✻ System	5.74	4	1.434	0.639	0.635	0.011
Residual	493.88	220	2.245			
Contextual emotion ✻ Response	34.79	4	8.698	2.158	0.075	0.038
Contextual emotion ✻ Response ✻ System	**47.32**	**4**	**11.830**	**2.934**	**0.022**	**0.051**
Residual	886.89	220	4.031			
Contextual face ✻ Contextual emotion ✻ Response	14.49	4	3.623	2.107	0.081	0.037
Contextual face ✻ Contextual emotion ✻ Response ✻ System	10.46	4	2.614	1.521	0.197	0.027
Residual	378.26	220	1.719			

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η²
System	1.37e-29	1	1.37e-29	3.77e-14	1.000	0.000
Residual	2.00e-14	55	3.63e-16			

Note. Type 3 Sums of Squares

Table SI 3 Repeated measures ANOVA for the testing condition (morph face) with contextual face (child, adult), contextual emotion (angry, neutral), and response (fear, surprise, shame, sadness, something else) as within-subject factors, and system (Montessori, traditional) as a between-subject factor.

Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η^2
Response	8747.61	4	2186.904	97.386	< .001	0.619
Response ✻ System	144.24	4	36.061	1.606	0.173	0.026
Residual	5389.46	240	22.456			
Contextual face	1.85e-29	1	1.85e-29	6.81e-15	1.000	0.000
Contextual face ✻ System	1.87e-30	1	1.87e-30	6.90e-16	1.000	0.000
Residual	1.63e-13	60	2.71e-15			
Contextual emotion	4.41e-29	1	4.41e-29	2.76e-14	1.000	0.000
Contextual emotion ✻ System	1.37e-30	1	1.37e-30	8.60e-16	1.000	0.000
Residual	9.59e-14	60	1.60e-15			
Response ✻ Contextual face	2.22	4	0.556	0.259	0.904	0.004
Response ✻ Contextual face ✻ System	5.85	4	1.463	0.681	0.606	0.011
Residual	515.43	240	2.148			
Response ✻ Contextual emotion	26.38	4	6.595	1.702	0.150	0.028
Response ✻ Contextual emotion ✻ System	**50.69**	**4**	**12.672**	**3.270**	**0.012**	**0.052**
Residual	930.11	240	3.875			
Contextual face ✻ Contextual emotion	2.20e-29	1	2.20e-29	3.61e-14	1.000	0.000
Contextual face ✻ Contextual emotion ✻ System	9.54e-33	1	9.54e-33	1.57e-17	1.000	0.000
Residual	3.65e-14	60	6.09e-16			
Response ✻ Contextual face ✻ Contextual emotion	15.49	4	3.873	2.346	0.055	0.038
Response ✻ Contextual face ✻ Contextual emotion ✻ System	10.35	4	2.587	1.567	0.184	0.025
Residual	396.18	240	1.651			

Note. Type 3 Sums of Squares

Between Subjects Effects

	Sum of Squares	df	Mean Square	F	p	partial η^2
System	3.10	1	3.10	0.937	0.337	0.015
Residual	198.40	60	3.31			

Note. Type 3 Sums of Squares

Table SI 4 Repeated measure ANOVA *for the whole sample of children* in the testing condition (morph face) with contextual face (child, adult), contextual emotion (angry, neutral), and response (fear, surprise, shame, sadness, something else) as within-subject factors, and system (Montessori, traditional) as a between-subject factor.