Uncovering the mechanisms of real-world attentional control over the course of primary education

Nora Turoman[1-3], Ruxandra I. Tivadar[1,2,4], Chrysa Retsa[1], Anne M. Maillard[5], Gaia Scerif[6], Pawel J. Matusz[1,2,7*]

[1] The LINE (Laboratory for Investigative Neurophysiology), Department of Radiology and Clinical Neurosciences, University Hospital Center and University of Lausanne, Lausanne, Switzerland
[2] Information Systems Institute at the University of Applied Sciences Western Switzerland (HES-SO Valais), 3960, Sierre, Switzerland
[3] Working Memory, Cognition and Development Lab, Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland
[4] Cognitive Computational Neuroscience Group, Institute of Computer Science, University of Bern, Bern, Switzerland
[5] Service des Troubles du Spectre de l'Autisme, Department of Psychiatry, University Hospital Center and University of Lausanne, Lausanne, Switzerland
[6] Department of Experimental Psychology, University of Oxford, Oxfordshire, UK
[7] Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

* - Corresponding author:
Pawel Matusz
Information Systems Institute
University of Applied Sciences Western Switzerland (HES-SO) Valais
Rue Technopole 3
3960 Sierre, Switzerland
Tel: +41 27 606 9060

Abstract

Schooling may shape children's abilities to control their attention, but it is unclear if this impact extends from control over visual objects to encompass multisensory objects, which are more typical of everyday environments. We compared children across three primary school grades (Swiss 1st, 3rd, and 5th grade) on their performance on a computer game-like audio-visual attentional control task, while recording their EEG. Behavioural markers of *visual* attentional control were present from 3rd grade (after 2 years of schooling), whereas multisensory attentional control was not detected in any group. However, multivariate whole-brain EEG analyses ('electrical neuroimaging') revealed stable patterns of brain activity that indexed both types of attentional control – visual control in all groups, and multisensory attentional control – from 3rd grade onwards. Our findings suggest that using multivariate EEG approaches can uncover otherwise undetectable mechanisms of attentional control over visual and multisensory objects and characterise how they differ at different educational stages.

Keywords: attentional control, multisensory, visual attention, schooling, EEG, N2pc, Electrical Neuroimaging

Lay Abstract

We measured how visual and audiovisual distractors differ in capturing attention of 1st- to 5th-graders while recording the children's brain activity. Brain activity results showed that all children were sensitive to visual distraction, and from 3rd grade onwards, children were also sensitive to audiovisual distraction. These results deepen our understanding of how school children control their attention in everyday environments, which are made up of information that stimulates multiple senses at a time.

Acknowledgments

We thank (name to be added after peer review) for assistance with child IQ data collection and (name to be added after peer review) for assistance with child EEG data analysis. We would also like to thank Micah Murray for helpful comments during the conceptualisation of the study. Finally, we would like to thank all of the families who participated in the current study. This work was supported by the Pierre Mercier Foundation (to PJM), the Swiss National Science Foundation (PZ00P1_174150 to PJM), and the Fondation Asile des Aveugles.

The start of school marks a transition from a less regulated, play-oriented environment to one with increasing demands for focused attention and ignoring distractors. The development of such attentional control mechanisms remains poorly understood, especially in real-world environments, and in the context of schooling experience.

Most current knowledge about children's attentional control has come from research on executive functions (EF; Miyake et al., 2000; Fiske & Holmboe, 2018), which are closely linked with attentional control (Bavelier & Green, 2019). Both EF's and attentional control gradually improve over childhood (e.g., Cepeda et al., 2001; Donnelly et al., 2007; Trick & Enns, 1998). Such developmental progression has been linked to protracted structural changes in the prefrontal and parietal cortex (e.g., Casey et al., 2005; Gogtay et al., 2004; Tamnes et al., 2013) and changes in connectivity between these regions (e.g., Baum et al., 2017; Hwang et al., 2011; Rohr et al., 2016). Arguably, such findings have helped us understand when and how children's control skills compare to those of adults, rather than to map and understand how control skills function during various stages of childhood, when children are challenged by their educational environment in different ways. Though both approaches are important, we focus on the latter for two reasons.

First, it has been suggested that the development of control skills need not be uniformly linear when the multisensory nature of the environment and the child's schooling experience are taken into account. For example, Matusz and colleagues (2015) found that while 11-year-olds and young adults were distracted from a visual search task by audiovisual shape-sound stimuli, 6-year-olds were immune to such distraction. This suggested that young children's limited attentional control skills could shield them from distraction in real-world contexts, rather than making them perform worse than adults, as visual attentional research would typically suggest. Moreover, in some cases, 11-year-olds were more distracted by auditory-only distractors than either 6-year olds or adults. This suggested that distractibility, when not tested only in vision, can be qualitatively different at different levels of experience. In a recent educationally-relevant version of this study (Matusz et al, 2019a), similar developmental differences were observed for distraction by digits, which could be accounted for by differences in experience. That is, children with less schooling experience were less familiar with numerals, which in turn protected them from distraction by conjunctions of numerals presented as visual symbols and their familiar auditorily-presented names.

Second, there may be links between the way in which children control their attention and school learning. That schooling is a catalyst for developing cognitive control is supported by reports of increases in IQ with education (e.g., Ceci, 1991; Brinch & Galloway, 2012) that were too rapid to be accounted by maturation alone (Cliffordson & Gustafsson, 2008). Another line of support has come from evidence on strong improvements in EF skills during the period when most children enter formal schooling ('5-to-7-year shift'; e.g., Burrage et al., 2008; Roebers et al., 2011). Recently, Brod et al. (2017) showed that children with one year of schooling experience had better behavioural EF skills, and greater recruitment of right posterior parietal cortex than kindergarteners. However, the influence of schooling on attentional control in multisensory contexts has yet to be directly investigated.

Systematically investigating the development of multisensory attentional control skills during primary school is important insomuch as classroom settings in which learning occurs are multisensory in nature. Multisensory processes, like visual attentional processes, show developmental progression (Murray et al., 2016). Starting in infancy, young brains are sensitive to congruency across the senses (Lewkowicz & Turkewitz, 1980; Bahrick & Lickliter, 2000), but processes that rely on integrating different sensory inputs into a single product (multisensory integration) seem to mature much later (8 and 10-11 years; Gori et al. 2008, 2012; Barutchu et al. 2009, respectively). Studying attentional control and *distraction* engaged by multisensory as well as visual objects could have important implications for education. Distraction by unisensory content has been demonstrated to hinder learning (visual: Fisher et al., 2014; Godwin & Fisher, 2011; auditory: Massonnié et al., 2019). Meanwhile, proficiency in visually-assessed EF/attentional control skills has been linked with academic achievement (e.g., Gawrilow et al., 2014; Bull et al., 2008; Isbell et al., 2019). Further, learning multisensory letter-sound mappings has been shown to be crucial for the development of specific reading networks (Brem et al., 2010), and can predict reading outcomes years later (Bach et al., 2013).

The present study investigated attentional control processes in primary school children who took part in a previous study of ours (reference post-review). Here, our focus was elucidating attentional control mechanisms at different stages of schooling. We investigated the following research questions in a largely exploratory fashion: 1) How do children process task-irrelevant multisensory distractors vis-à-vis visual distractors, and what are the brain mechanisms governing such processes? 2) How do these mechanisms differ across different levels of school experience? We investigated the behavioural and brain mechanisms of attentional control using a child-friendly version of a multisensory spatial cueing paradigm, during which EEG was recorded. We analysed the traditional EEG correlate of attentional selection ("N2pc component") and enriched the understanding of the underlying brain mechanisms using multivariate analyses. We expected older children to show visual attentional control behaviourally; we had no strong hypotheses for multisensory attentional control or the underlying EEG mechanisms.

Methods

Participants

We tested 92 children from local primary schools. In the local school system, children enter formal education at age 4, the first two years of which are considered kindergarten. By 3rd grade (ages 6-7) children sit at desks and receive more structured classroom instruction. In total, we tested 26 children from 5th grade (10 male, M_{age}: 8y 10mo, *SD*: 5mo, range: 8y 1mo – 10y 1mo), 38 from 3rd grade (18 female, M_{age}: 6y 10mo, *SD*: 4mo, range: 6y 1mo, 7y 9mo), and 28 from 1st grade (13 female, M_{age}: 5y, *SD*: 4mo, range: 4y – 5y 7mo; for full details see Supplementary Online Materials, SOMs). All research procedures were approved by the local ethics committee.

Stimuli and procedure

Participants were tested in the local hospital in an experimental session lasting 1h–1h30mins, where their performance was assessed along with a simultaneous EEG measurement. The paradigm was a multisensory variant of Folk et al.'s (1992) spatial cueing paradigm (Matusz & Eimer, 2011, Exp.2) adapted to children by presenting it as a diamond-search game (see Fig.1).

---FIG.1 HERE---

Participants searched for a target diamond of a predefined colour, to respond as quickly and accurately as possible to the target's orientation (horizontal or vertical; randomly determined on each trial) by pressing one of two horizontally aligned round buttons (Lib Switch, Liberator Ltd.) fixed onto a tray bag on their lap. Each diamond was preceded by a spatially uninformative cue, which could match the target colour (e.g., blue cue for blue target), or have a target-nonmatching colour that was not present before (e.g., red cue for blue target). On half of all trials, the colour change coincided with the onset of a pure sine-wave tone (2000Hz), resulting in four cue conditions: TCCV (target colour-cue, Visual), NCCV (nontarget colour-cue, Visual), TCCAV (target colour-cue, AudioVisual), NCCAV (nontarget colour-cue, AudioVisual). Thus, visual attentional control was assessed through the presence of task-set contingent attentional capture (TAC), i.e., attentional capture by target-matching (TCC) but not nonmatching cues (NCC) (Folk et al., 1992; Yantis & Jonides, 1990). Attentional control by multisensory stimuli was assessed through the occurrence of multisensory enhancements of attentional capture (MSE), i.e. increases in attentional capture for audiovisual cues (AV) regardless of their target-matching (Matusz & Eimer, 2011). Full details of the paradigm and procedure are reported in SOMs.

EEG acquisition and preprocessing

Continuous EEG data sampled at 1000Hz was recorded using a 129-channel HydroCel Geodesic Sensor Net connected to a NetStation amplifier (Net Amps 400; Electrical Geodesics Inc., Eugene, OR, USA). Impedances were kept <50kΩ, and electrodes were referenced online to Cz. Preprocessing involved: offline filtering (0.1 Hz high-pass, 40 Hz low-pass, 50 Hz notch, and a second-order Butterworth filter with a linearly computed –12 dB/octave roll-off with forward and backward passes to eliminate phase-shift), segmentation into epochs around cue onset (-100ms to 500ms), semi-automated rejection of transient noise, eye movements, and muscle artefacts. Artefact rejection criteria were $\pm 150\ \mu V$, along with visual inspection (e.g., Melinder et al., 2010; Shimi et al., 2014). Artefact-contaminated electrodes were interpolated using three-dimensional splines (Perrin et al., 1987; average numbers of epochs removed, and electrodes interpolated in SOMs). Cleaned epochs were averaged, baseline corrected to the 100ms pre-cue time interval, and re-referenced to the average reference. An additional 50Hz notch filter was applied due to persistent environmental noise even after initial filtering. Only data from trials with correct responses and from blocks with >50% accuracy were used for statistical analyses.

Analyses of the EEG focused on Event Related Potentials (ERPs). The above steps were done separately for ERPs from the four cue conditions, as EEG/ERP analyses were anchored to the cue array, and separately for cues in the left and right hemifield. To analyse cue-elicited lateralised ERPs, data from both hemiscalps had to be anchored to the same 'reference' side. We thus changed the labels of single-trial data from all conditions with cues presented on the left so that they represent activity over the right hemiscalp, resulting in a set of veridical "cue-on-the-right" data and mirrored "cue-on-the-right" data. Next, these two types of data were averaged together, creating a single lateralised average ERP. As a result, we obtained 4 different cue ERPs. All preprocessing and analyses were conducted using CarTool (available at www.fbmlab.com/cartool-software/; Brunet et al., 2011).

Data analysis design

Behavioural analyses. Analyses were focused on mean reaction time (RT) attentional capture effects, following related adult (e.g., Folk et al., 1992) and child (e.g., Gaspelin et al., 2015) literature. We calculated attentional capture effects by subtracting the mean RTs for trials where the cue and target were in the same vs. different location, separately for the four cue conditions. RT data cleaning is described in SOMs. RT data were submitted to a mixed-design four-way ANOVA with one between-subject factor of Age (5th-graders vs. 3rd-graders vs. 1st-graders), and three within-subject factors: Cue Colour (target colour-cue - TCC vs. nontarget colour-cue - NCC), Cue Modality (Visual - V vs. AudioVisual - AV), and Cue-Target Location (Same vs. Different). All analyses were conducted using SPSS for Macintosh 26.0 (Armonk, NY: IBM Corp). Detailed behavioural results were reported as part of a study where the emergence of adult visual and multisensory attentional control skills over childhood was investigated (Turoman et al. 2020). Therefore, here we only reported the most relevant behavioural results.

EEG analyses. We combined an analysis of the N2pc component, a traditional marker of attentional selection of potential targets (Eimer 1996; Eimer et al., 2009), with a more novel electrical neuroimaging (EN) approach. Combining such EN measures with established ERP correlates of cognitive processes can help elucidate the cognitive and brain mechanisms underlying multisensory attentional control (Matusz et al. 2019b). As no reliable N2pcs were found in any of the tested age groups, here we report only the EN analyses of the cue-elicited N2pc's (Full details of the N2pc analysis design and results are reported in SOMs).

Electrical neuroimaging (EN) encompasses a set of multivariate, reference-independent analyses of global features of the electric field at the scalp, providing robust, directly neurophysiologically interpretable results about differences between sets of conditions, groups or timepoints (Lehmann & Skrandies 1980; Murray et al., 2008; Tivadar & Murray 2019). Detailed, tutorial-like explanations of EN measures and how they capture the mechanisms underlying the N2pc are given in both Matusz et al. (2019b), and Murray et al. (2008). We also provide a concise description of the measures and how they were derived for the present study in SOMs. Briefly, in order to obtain *global* EN measures of *lateralised* N2pc effects, now the contralateral-ipsilateral difference ERPs created for N2pc analyses were mirrored onto the other hemiscalp, constructing ‘fake’ 129-electrode datasets. From these ‘fake’ datasets, Global Field Power (GFP) and topographical measures were extracted in order to investigate whether modulations of cue-elicited lateralised ERPs by visual and multisensory control mechanisms were a result of differential response strength within statistically indistinguishable brain generators, or differential recruitment of underlying brain generators, respectively.

For GFP analyses, each age group’s mean GFP values over the respective N2pc time-windows were extracted from group-averaged ‘fake’ ERPs per condition and submitted to separate two-way repeated measures ANOVAs (rmANOVAs) with factors: Cue Colour (TCC vs. NCC) and Cue Modality (V vs. AV). For topographical analyses, topographical map durations (in milliseconds) over each group’s N2pc time-window were submitted to separate three-way rmANOVAs, with factors: Cue Colour, Cue Modality , and Map (different levels due to different numbers of maps in each age group: 9 in 5th-graders, 5 in 3rd-graders, and 8 in 1st-graders).

Results

Behavioural analyses

A significant main effect of Age, $F_{(2,\ 89)}=32.8$, $p<0.001$, $\eta_p^2=0.4$ revealed that mean RTs sped up from 1st graders (1309ms) through 3rd graders (1107ms) to 5th graders (836ms). Each age group was reliably faster than its preceding group (all p's<0.001, see SOMs). Although Age did not interact with other factors (all F's<2, p's>0.1), RT data were analysed per age group in order to investigate in more detail multisensory distraction effects across age groups.

Fifth-graders showed a reliable task-set contingent attentional capture (TAC) effect, as demonstrated by a 2-way Cue-Target Location x Cue Colour interaction, $F_{(1,\ 25)}=19.5$, $p<0.001$, $\eta_p^2=0.4$, which was driven by significant capture effects for TCC (56ms), but not NCC distractors (6ms; Fig.2A). However, 5th-graders showed no evidence for a reliable multisensory enhancement (MSE) effect (Cue-Target Location x Cue Modality, $F_{(1,\ 25)}=1.4$, $p=0.3$; Fig.2B). Similarly, 3rd-graders also showed TAC, evidenced by a Cue-Target Location x Cue Colour interaction, $F=6.4$, $p=0.02$, $\eta_p^2=0.2$, which was driven by reliable capture effects elicited by TCC (55ms), but not NCC distractors (7ms). Again, however, there was no evidence for MSE in 3rd-graders (Cue-Target Location x Cue Modality, $F_{(1,\ 37)}=2.1$, $p=0.2$; Fig.2B). In the youngest group, there was no evidence for either TAC (Cue-Target Location x Cue Colour, $F_{(1,\ 27)}=1.4$, $p=0.2$), or MSE (Cue-Target Location x Cue Modality, $F_{(1,\ 27)}=0.4$, $p=0.5$) (Fig.2D).

---FIG.2 HERE---

EEG analyses

Separate 2 x 2 rmANOVAs on the average GFP values over each age group's N2pc time-window revealed no significant main effects or interactions in the 5th- and 1st-graders (p's>0.1), while in 3rd-graders, the main effect of Cue Colour only trended ($F=3.07$, $p=0.09$, $\eta_p^2=0.08$). Full results can be found in SOMs.

Segmentation of the post-cue period per age group resulted in the following: in 5th-graders, 14 clusters explained 91.5% of GEV, in 3rd-graders, 11 clusters explained 88.3% of GEV, and in 1st-graders, 11 clusters explained 84.9% of GEV in the group-averaged difference ERPs. Next, the fitting procedure revealed the template maps that characterised each age group's N2pc time window: in 5th-graders, there were 9 template maps over 144-290ms post-cue, in 3rd-graders there were 5 template maps over 151-275ms post-cue, and in 1st-graders, there were 8 template maps over 110-302ms post-cue. The fitting time-windows differ slightly from N2pc time-windows, as we included template the maps that appeared to characterise cue-elicited N2pc's. For example, if the duration of a map was longer than the N2pc-based cut-off, we extended the fitting time-window beyond the cut-off and until the end of the map duration.

In 5th-graders, a 2 x 2 x 9 rmANOVA revealed a main effect of Map, $F=2.7$, $p=0.009$, $\eta_p^2=0.1$, confirming the presence of stable patterns of lateralised N2pc-like ERP activity in 5th-graders.

Hereafter, we did not follow up this main effect with post-hoc tests, as it was not informative as to the presence of TAC or MSE in topography. To help differentiate topographic maps between groups, map numbers were given a prefix indicating the group they belonged to, i.e., '5' for 5th-graders, '3' for 3rd-graders, and '1' for 1st-graders. There was a 2-way interaction between Map and Cue Colour, $F_{(8,\ 200)}$=3.4, p=0.001, η_p^2=0.1, which was driven by modulations in the durations of three maps: Map54, Map56, and Map59 (Fig.3A, left panel). Specifically, Map56 was present for a longer portion of the N2pc time-window in response to TCC distractors (28ms) than NCC distractors (8ms), $t_{(25)}$=3.7, p=0.005. Meanwhile, Map54 was present longer in response to NCC distractors (20ms) than for TCC distractors (7ms), $t_{(25)}$=3.4, p=0.008, and Map59 was also present longer for NCC (16ms) than TCC distractors (9ms), $t_{(25)}$=3, p=0.04 (Fig.3A, left panel). Hereafter, map duration differences that were not reported were not statistically significant (p's > 0.1). The Map x Cue Modality interaction was also significant, $F_{(8,\ 200)}$=2.4, p=0.02, η_p^2 =0.1, and was driven by modulations of durations of Map52 and Map53 (Fig.3A, right panel). Here, Map53 was present longer for AV cues (29ms) than for V cues (13ms), $t_{(25)}$=3.7, p=0.007, while Map52 was present longer for V cues (17ms) than for AV cues (5ms), $t_{(25}$ =3.1, p=0.03. Finally, the 3-way Map x Cue Colour x Cue Modality interaction was also significant, $F_{(8,\ 200)}$=2.2, p=0.048, η_p^2 = 0.1. For brevity, we report the post-hoc analyses in SOMs. Taken together, not only did 5th graders' topography show reliable TAC and MSE, but also the interaction of these factors.

In 3rd-graders, there was also a main effect of Map, $F_{(3.2,\ 117.6)}$=9.8, p<0.001, η_p^2=0.2, as revealed by a 2 x 2 x 5 rmANOVA. Like in 5th graders, the 2-way Map x Cue Colour interaction was significant, $F_{(2.9\ 107)}$=2.8, p=0.04, η_p^2=0.1, and was driven by Map35 being longer for TCC (21ms) than NCC distractors (5ms), $t_{(37)}$=3.7, p=0.002 (Fig.3B, left panel). The Map x Cue Modality interaction was also significant $F_{(3.1,\ 114.3)}$=8, p<0.001, η_p^2=0.2, and was driven by modulations of Maps 31–34. Here, Map32 was present longer for AV cues (44ms) than for V cues (26ms), $t_{(37)}$=3.5, p=0.002, and Map34 was also present longer for AV cues (37ms) than for V cues (21ms), $t_{(37)}$=3.2, p=0.01. Conversely, Map31 was present longer for V cues (43ms) than for AV cues (26ms), $t_{(37)}$=3.5, p=0.004, and Map33 was present longer for V cues (16ms) than for AV cues (6ms), $t_{(37)}$=2.8, p=0.006. Again, the 3-way Map x Cue Colour x Cue Modality interaction was significant, F=3.2, p=0.03, η_p^2=0.1, and its follow-up analyses are presented in SOMs.

In 1st-graders also a 2 x 2 x 8 rmANOVA revealed a main effect of Map, $F_{(4.7,\ 127.2)}$=4, p=0.003, η_p^2=0.1. Like in the older groups, the Map x Cue Colour interaction was significant, $F_{(7,\ 189)}$=4.2, p<0.001, η_p^2=0.1, driven by modulations in the durations of Maps 13-15. Map13 was present longer for TCC cues (47ms) than NCC cues (13ms), $t_{(27)}$=4.3, p=0.003. Conversely, Map14 was present longer for NCC cues (52ms) than TCC cues (27ms), $t_{(27)}$=3.1, p=0.007, as was Map15 (19ms vs. 5ms), $t_{(27)}$=2.8, p=0.05. Unlike in the older groups, however, map presence was not modulated by Cue Modality ($F_{(7,\ 189)}$=0.9, p=0.4). That said, the 3-way Map x Cue Colour x Cue Modality interaction was significant, $F_{(7,\ 189)}$=2.2, p=0.04, η_p^2=0.1, and its follow-up analyses are presented in SOMs.

---FIG.3 HERE---

Discussion

We investigated how attentional control skills develop using a rigorous paradigm that incorporated the multi-stimulus and multisensory characteristics of real-world contexts, such as the classroom. We compared first-graders, third-graders and fifth-graders on how they attend to visual objects and to more naturalistic, audiovisual objects. We enriched behavioural and classical EEG/ERP analyses with multivariate analyses.

Top-down visual attentional control is present even at school entry

In behaviour, robust task-set contingent attentional capture (TAC) was observed in 3rd grade and continued into 5th grade. However, EN topographical analyses demonstrated distinct stable patterns of global brain activity that were sensitive to TAC already at school entry.

In older groups, EN results clarified the brain mechanisms underlying the behaviourally-observed TAC. In 3rd-graders, behavioural TAC may emerge from enhancement of target-matching (TCC) distractor processing, via recruitment of brain networks that preferentially process goal-relevant information. Meanwhile, in 5th-graders, TAC may be driven by a combination of enhanced processing of target-matching information and suppressed processing of nonmatching (NCC) information. This is supported by one map predominating the processing of target-matching distractors, and two maps predominating the processing of target-nonmatching distractors. It cannot be ascertained if increased presence of NCC-preferring maps indexed the inhibition of NCC information, however, this is suggested by the concomitant behavioural inhibition of NCC distractors. Consistently, in a study isolating N2pc to visual distractors, a similar pattern of results was found in adults (Hickey et al., 2009). Though the relationship between topographic map modulations and distractor processing is not clear-cut, it is clear that differential brain network engagement, and not differences in brain response strength, are the mechanism that guides TAC in children.

In 1st-graders, EN revealed nascent visual attentional control processes. Despite no behavioural TAC, 1st-graders showed two maps that were preferentially recruited by target nonmatching cues, and one map recruited for target-matching cues. This mirrors the TAC-related findings in 5th-graders. These results support Gaspelin et al.'s (2015) findings that TAC is present as early as age 4 and extend them with novel mechanistic insights. Indeed, most studies on control processes in 4-year-olds utilised only behavioural measures (e.g., Burrage et al., 2008; Roebers et al., 2011, but see Brod et al., 2017). That separable sets of brain networks are preferentially active in response to potentially goal-relevant and goal-irrelevant information even earlier than the 5-to-7-year shift is the first finding of its kind.

Attentional control over multisensory objects develops after two years of schooling

Behavioural analyses did not detect MSE in any group. However, in the older groups, EN revealed distinct brain networks over 200ms post-stimulus, recruited by visual and audiovisual distractors. The presence of topographical, but not behavioural, modulations by MSE is important for two reasons. First, this finding supports the idea that salient multisensory stimuli, even when task-irrelevant, can control attention independently of top-down goal-relevance influences (Matusz & Eimer, 2011). No evidence for behavioural MSE could be explained by behavioural (but not EN measures) being 'contaminated' by processing speed and motor response processes, which develop slowly over time (e.g. Kail & Ferrer, 2007) Second, this finding suggests that multisensory distraction may emerge earlier than previously thought. Studies using exclusively behavioural measures have suggested that optimal multisensory integration (Gori et al., 2008, 2012), and even interference by multisensory objects (Matusz et al., 2015; 2019a) occur later than the ages we tested here using EN (6-7 years). Thus, distraction from multisensory objects emerges already after two years of schooling, and 6-7-year-olds are not protected from it, as suggested by research on multisensory interference.

First-graders showed no evidence for topographic modulations by MSE. If anything, the addition of sound to visual distractors seemed to have had a disruptive effect on ERPs overall, as shown by suppressed contralateral ERP responses. It may be that in 1st-graders, but not older children, attentional resources are separably allocated to vision and audition (e.g., Welch & Warren, 1980), although evidence for such an account has been mixed (Jacoby et al., 2012). It remains unclear how children at school-entry processed the additional distractor sound. We are currently testing an additional auditory-only condition in our paradigm to address this issue.

Implications for education

Our study demonstrated that with less than a year of schooling experience, children showed a brain-network-level sensitivity to the goal-relevance of visual stimuli, while children with two additional years of schooling showed a sensitivity to the (audio)visual nature of stimuli. This corroborates the idea of schooling experience as a “neurocognitive developmental institution” (Baker et al., 2012), although our study design could not ascertain whether schooling influenced attentional control independently of experience-unrelated (maturation-like) development. Further, our findings suggest that the unisensory distraction by classroom clutter in children as young as 5 (Fisher et al., 2014) and noise in ages 5-8 (Massonnié et al., 2019) may be further compounded by children’s sensitivity to distraction by audiovisual objects from age 6 onwards. Simultaneously, children with little school experience are protected from multisensory distraction. Therefore, in classrooms, but not kindergartens, there may be merit in minimising the risks of multisensory distraction by reducing decoration or use of new technologies – unless they are goal-relevant, i.e., related to the subject of instruction.

References

Bach, S., Richardson, U., Brandeis, D., Martin, E., & Brem, S. (2013). Print-specific multimodal brain activation in kindergarten improves prediction of reading skills in second grade. *Neuroimage*, *82*, 605-615.

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental psychology*, *36*(2), 190-201.

Baker, D. P., Salinas, D., & Eslinger, P. J. (2012). An envisioned bridge: Schooling as a neurocognitive developmental institution. *Developmental Cognitive Neuroscience*, *2*, S6-S17.

Bakos, S., Mehlhase, H., Landerl, K., Bartling, J., Schulte-Körne, G., & Moll, K. (2020). Naming processes in reading and spelling disorders: An electrophysiological investigation. *Clinical Neurophysiology*, *131*(2), 351-360.

Barutchu, A., Crewther, D. P., & Crewther, S. G. (2009). The race that precedes coactivation: development of multisensory facilitation in children. *Developmental science*, *12*(3), 464-473.

Baum, G. L., Ciric, R., Roalf, D. R., Betzel, R. F., Moore, T. M., Shinohara, R. T., ... & Cook, P. A. (2017). Modular segregation of structural brain networks supports the development of executive function in youth. *Current Biology*, *27*(11), 1561-1572.

Bavelier, D., & Green, C. S. (2019). Enhancing Attentional Control: Lessons from Action Video Games. *Neuron*, *104*(1), 147-163.

Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., ... & Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proceedings of the National Academy of Sciences*, *107*(17), 7939-7944.

Brinch, C. N., & Galloway, T. A. (2012). Schooling in adolescence raises IQ scores. *Proceedings of the National Academy of Sciences*, *109*(2), 425-430.

Brod, G., Bunge, S. A., & Shing, Y. L. (2017). Does one year of schooling improve children's cognitive control and alter associated brain activation?. *Psychological science*, *28*(7), 967-978.

Brunet, D., Murray, M. M., & Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. *Computational intelligence and neuroscience*, *2011*, 2.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental neuropsychology*, *33*(3), 205-228.

Burrage, M. S., Ponitz, C. C., McCready, E. A., Shah, P., Sims, B. C., Jewkes, A. M., & Morrison, F. J. (2008). Age-and schooling-related effects on executive functions in young children: A natural experiment. *Child Neuropsychology*, *14*(6), 510-524.

Casey, B. J., Tottenham, N., Liston, C., & Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development?. *Trends in cognitive sciences*, *9*(3), 104-110.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental psychology*, *27*(5), 703.

Cepeda, N. J., Kramer, A. F., & Gonzalez de Sather, J. (2001). Changes in executive control across the life span: examination of task-switching performance. *Developmental psychology*, *37*(5), 715.

Cliffordson, C., & Gustafsson, J. E. (2008). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence*, *36*(2), 143-152.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, *3*(3), 201.

Couperus, J. W., & Quirk, C. (2015). Visual search and the N2pc in children. *Attention, Perception, & Psychophysics*, *77*(3), 768-776.

De Beer, J., Engels, J., Heerkens, Y., & van der Klink, J. (2014). Factors influencing work participation of adults with developmental dyslexia: a systematic review. *BMC public health*, *14*(1), 77.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience, 18(1*), 193-222.

Donnelly, N., Cave, K., Greenway, R., Hadwin, J. A., Stevenson, J., & Sonuga-Barke, E. (2007). Visual search in children and adults: Top-down and bottom-up mechanisms. *Quarterly Journal of Experimental Psychology*, *60*(1), 120-136.

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and clinical neurophysiology*, *99*(3), 225-234.

Eimer, M., Kiss, M., Press, C., & Sauter, D. (2009). The roles of feature-specific task set and bottom-up salience in attentional capture: an ERP study. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(5), 1316.

Fisher, A. V., Godwin, K. E., & Seltman, H. (2014). Visual environment, attention allocation, and learning in young children: When too much of a good thing may be bad. *Psychological science*, *25*(7), 1362-1370.

Fiske, A., & Holmboe, K. (2019). Neural substrates of early executive function development. *Developmental Review*, *52*, 42-62.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human perception and performance*, *18*(4), 1030.

Gaspelin, N., Margett-Jordan, T., & Ruthruff, E. (2015). Susceptible to distraction: Children lack top-down control over spatial attention capture. *Psychonomic bulletin & review*, *22*(2), 461-468.

Gawrilow, C., Fäsche, A., Guderjahn, L., Gunzenhauser, C., Merkt, J., & von Suchodoletz, A. (2014). The impact of self-regulation on preschool mathematical achievement. *Child Indicators Research*, *7*(4), 805-820.

Godwin, K., & Fisher, A. (2011). Allocation of attention in classroom environments: consequences for learning. In *Proceedings of the Annual Meeting of the Cognitive Science Society* 33(33).

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., . . . Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences, USA*, *101*, 8174–8179. doi:10.1073/ pnas.0402680101.

Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Current Biology*, *18*(9), 694-698.

Gori, M., Sandini, G., & Burr, D. (2012). Development of visuo-auditory integration in space and time. *Frontiers in integrative neuroscience*, *6*, 77.

Hickey, C., Di Lollo, V., & McDonald, J. J. (2009). Electrophysiological indices of target and distractor processing in visual search. Journal of cognitive neuroscience, 21(4), 760-775.

Hwang, K., Velanova, K., & Luna, B. (2010). Strengthening of top-down frontal cognitive control networks underlying the development of inhibitory control: a functional magnetic resonance imaging effective connectivity study. *Journal of Neuroscience, 30(46),* 15535-15545.

Isbell, E., Calkins, S. D., Cole, V. T., Swingler, M. M., & Leerkes, E. M. (2019). Longitudinal associations between conflict monitoring and emergent academic skills: An event-related potentials study. *Developmental psychobiology*, *61*(4), 495-512.

Jacoby, O., Hall, S. E., & Mattingley, J. B. (2012). A crossmodal crossover: opposite effects of visual and auditory perceptual load on steady-state evoked potentials to irrelevant visual stimuli. Neuroimage, 61(4), 1050-1058.

Kail, R. V., & Ferrer, E. (2007). Processing speed in childhood and adolescence: Longitudinal models for examining developmental change. Child development, 78(6), 1760-1770.

Kiss, M., Van Velzen, J., & Eimer, M. (2008). The N2pc component and its links to attention shifts and spatially selective visual processing. *Psychophysiology*, *45*(2), 240-249.

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and clinical neurophysiology*, *48*(6), 609-621.

Lewkowicz, D. J., & Turkewitz, G. (1980). Cross-modal equivalence in early infancy: Auditory–visual intensity matching. *Developmental psychology*, *16*(6), 597.

Luck, S. J., & Hillyard, S. A. (1994). Spatial filtering during visual search: evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1000.

Massonnié, J., Rogers, C. J., Mareschal, D., & Kirkham, N. Z. (2019). Is classroom noise always bad for children? The contribution of age and selective attention to creative performance in noise. *Frontiers in psychology*, *10*, 381.

Matusz, P. J., & Eimer, M. (2011). Multisensory enhancement of attentional capture in visual search. *Psychonomic Bulletin & Review*, *18*(5), 904.

Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkley, R., & Scerif, G. (2015). Multi-modal distraction: Insights from children's limited attention. *Cognition*, *136*, 156-165.

Matusz, P. J., Merkley, R., Faure, M., & Scerif, G. (2019a). Expert attention: Attentional allocation depends on the differential development of multisensory number representations. *Cognition*, *186*, 171-177.

Matusz, P. J., Turoman, N., Tivadar, R. I., Retsa, C., & Murray, M. M. (2019b). Brain and cognitive mechanisms of top–down attentional control in a multisensory world: Benefits of electrical neuroimaging. *Journal of cognitive neuroscience*, *31*(3), 412-430.

Melinder, A., Gredeback, G., Westerlund, A., & Nelson, C. A. (2010). Brain activation during upright and inverted encoding of own- and other-age faces: ERP evidence for an own-age bias. *Developmental Science, 13,* 588–598.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & 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.

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: a step-by-step tutorial review. *Brain topography*, *20*(4), 249-264.

Murray, M. M., Lewkowicz, D. J., Amedi, A., & Wallace, M. T. (2016). Multisensory processes: a balancing act across the lifespan. *Trends in Neurosciences*, *39*(8), 567-579.

Perrin, F., Pernier, J., Bertnard, O., Giard, M. H., & Echallier, J. F. (1987). Mapping of scalp potentials by surface spline interpolation. *Electroencephalography and clinical neurophysiology*, *66*(1), 75-81.

Plude, D. J., Enns, J. T., & Brodeur, D. (1994). The development of selective attention: A life-span overview. *Acta psychologica*, *86*(2-3), 227-272.

Roebers, C. M., Röthlisberger, M., Cimeli, P., Michel, E., & Neuenschwander, R. (2011). School enrolment and executive functioning: A longitudinal perspective on developmental changes, the influence of learning context, and the prediction of pre-academic skills. *European Journal of Developmental Psychology*, *8*(5), 526-540.

Rohr, C. S., Vinette, S. A., Parsons, K. A., Cho, I. Y., Dimond, D., Benischek, A., ... & Bray, S. (2016). Functional Connectivity of the Dorsal Attention Network Predicts Selective Attention in 4–7 year-old Girls. *Cerebral Cortex*, *27*(9), 4350-4360.

Scerif, G., Kotsoni, E., & Casey, B. J. (2006). Functional Neuroimaging of Early Cognitive Development. *Handbook of Functional Neuroimaging of Cognition*, 351.

Shimi, A., Kuo, B. C., Astle, D. E., Nobre, A. C., & Scerif, G. (2014b). Age group and individual differences in attentional orienting dissociate neural mechanisms of encoding and maintenance in visual STM. *Journal of Cognitive Neuroscience*, *26*(4), 864-877.

Shimi, A., Nobre, A. C., & Scerif, G. (2015). ERP markers of target selection discriminate children with high vs. low working memory capacity. *Frontiers in systems neuroscience*, *9*, 153.

Soto-Faraco, S., Kvasova, D., Biau, E., Ikumi, N., Ruzzoli, M., Morís-Fernández, L., & Torralba, M. (2019). *Multisensory interactions in the real world*. Cambridge University Press.

Sun, M., Wang, E., Huang, J., Zhao, C., Guo, J., Li, D., ... & Song, Y. (2018). Attentional selection and suppression in children and adults. *Developmental science*, *21*(6), e12684.

Tamnes, C. K., Walhovd, K. B., Grydeland, H., Holland, D., Østby, Y., Dale, A. M., & Fjell, A. M. (2013). Longitudinal working memory development is related to structural maturation of frontal and parietal cortices. *Journal of cognitive neuroscience*, *25*(10), 1611-1623.

Tivadar, R. I., & Murray, M. M. (2019). A primer on electroencephalography and event-related potentials for organizational neuroscience. *Organizational Research Methods*, *22*(1), 69-94.

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of school psychology*, *40*(1), 7-26.

Trick, L. M., & Enns, J. T. (1998). Lifespan changes in attention: The visual search task. *Cognitive Development*, *13*(3), 369-386.#

Turoman, N., Tivadar, R. I., Retsa, C., Maillard, A. M., Scerif, G., & Matusz, P. J. (2020). The development of attentional control mechanisms in multisensory environments. *BioRxiv*.

Valås, H. (1999). Students with learning disabilities and low-achieving students: Peer acceptance, loneliness, self-esteem, and depression. *Social psychology of education*, *3*(3), 173-192.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. Psychological Bulletin, 88, 638–667. http:// dx.doi.org/10.1037/0033-2909.88.3.638.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception & Performance*, *16*, 121-134.

Yoder, P. J., Molfese, D., Murray, M. M., & Key, A. P. (2013). Normative topographic ERP analyses of speed of speech processing and grammar before and after grammatical treatment. *Developmental neuropsychology*, *38*(8), 514-533.

Figure legends

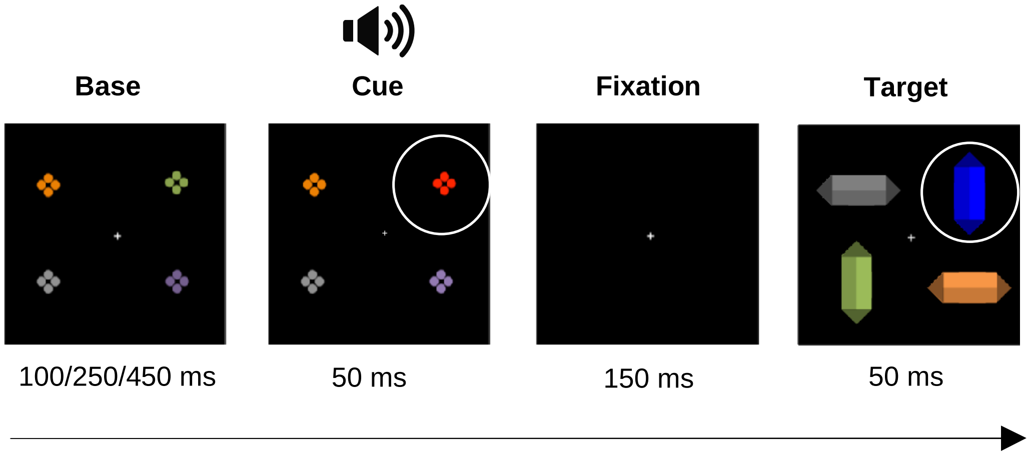

Fig.1. Experimental trial sequence for the child-friendly version of Matusz & Eimer (2011, Experiment 2). Each trial was composed of a series of consecutive arrays: the base array, cue array, fixation array, and target array. The cue array, to which EEG/ERP analyses were anchored to, involved one of the four differently coloured elements of the base array changing colour. The colour change was randomised and not spatially predictive of the location of the subsequent target (same cue-target location on 25% of trials). Here, a target-nonmatching (or nontarget colour) cue is shown, with an accompanying tone, constituting a NCCAV trial. The white circles around the target blue diamond and preceding target-nonmatching red cue serve to highlight the cue and target in the current figure and did not appear in the experimental task.

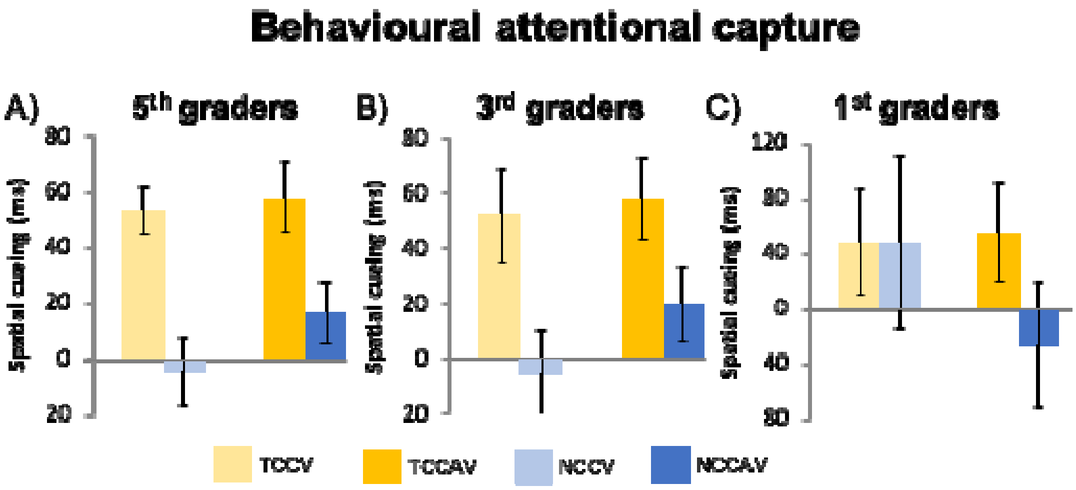

Fig.2. Bars coloured according to the figure legend in the image represent mean behavioural RT attentional capture, and error bars represent the standard error of the mean. From 3rd grade onwards, behavioural RT capture was larger for TCC than NCC, whereas enhancements of capture by sound presence were not significant.

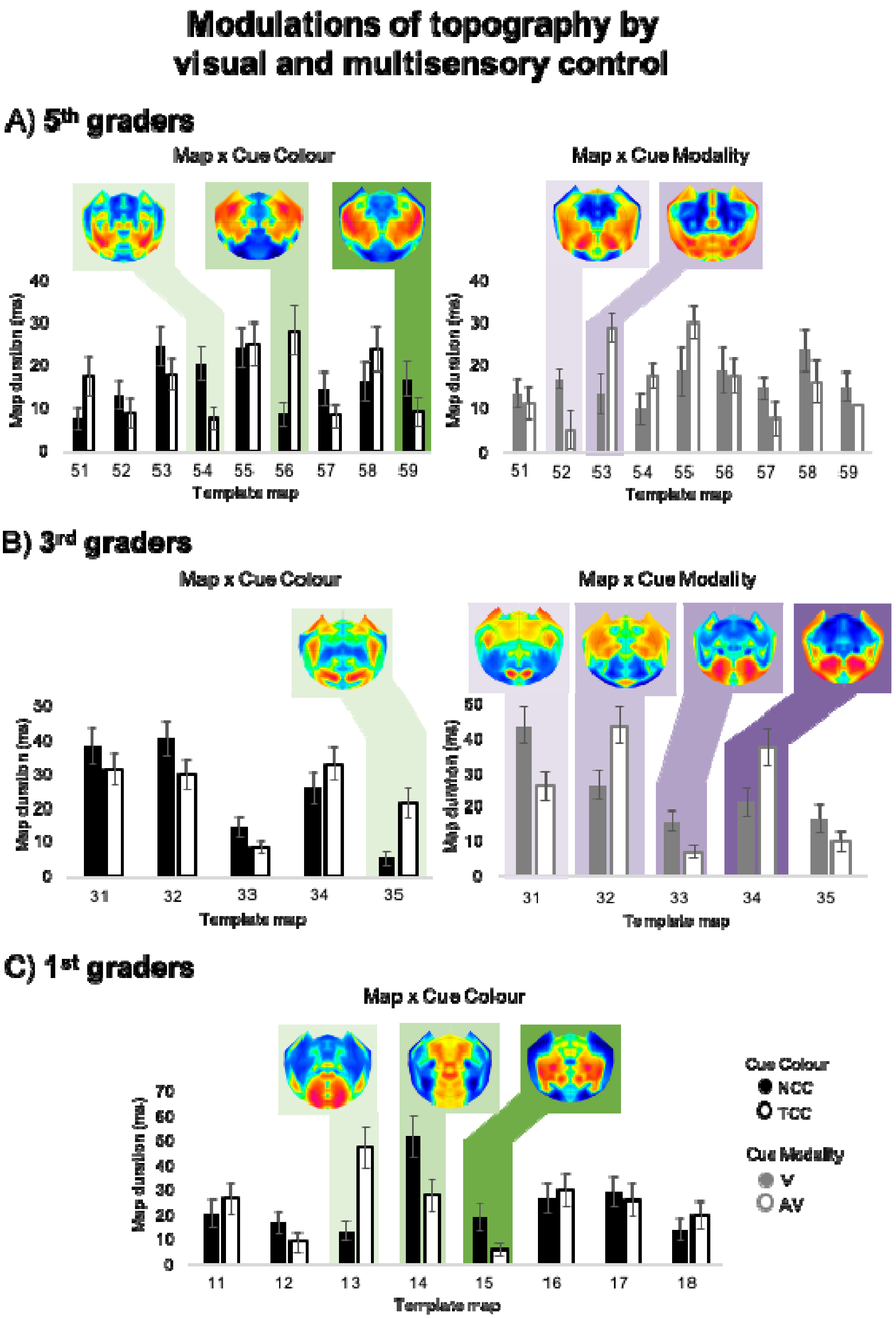

Fig.3. Results of the two-way interactions of the topographical EN analyses. Bars represent map durations over each group's time-window, where black and black-outlined bars represent the interaction with Cue Colour, and grey and grey-outlined bars represent the interaction with Cue Modality. Error bars represent the standard errors of the mean. Template maps that were significantly modulated by Cue Colour are highlighted in shades of green, while maps that were modulated by Cue Modality are highlighted in shades of purple.