The development of attentional control mechanisms in multisensory environments

Nora Turoman[a,f], Ruxandra I. Tivadar[a,b], Chrysa Retsa[a], Anne M. Maillard[c], Gaia Scerif[d], Pawel J. Matusz[a,e,f*]

[a] The LINE (Laboratory for Investigative Neurophysiology), Department of Radiology and Clinical Neurosciences, University Hospital Center and University of Lausanne, Lausanne, Switzerland
[b] Department of Ophthalmology, Fondation Asile des Aveugles, Lausanne, Switzerland
[c] Service des Troubles du Spectre de l'Autisme, Department of Psychiatry, University Hospital Center and University of Lausanne, Lausanne, Switzerland
[d] Department of Experimental Psychology, University of Oxford, Oxfordshire, UK
[e] Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA
[f] Information Systems Institute at the University of Applied Sciences Western Switzerland (HES-SO Valais), 3960, Sierre, Switzerland

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

Author contributions:
Nora Turoman: Investigation, Formal analysis, Data curation, Software, Visualisation, Writing - original draft.
Ruxandra I. Tivadar: Software, Writing - review & editing.
Chrysa Retsa: Software, Writing - review & editing.
Anne Maillard: Methodology, Writing - review & editing.
Gaia Scerif: Methodology, Writing - review & editing.
Pawel J. Matusz: Conceptualization, Funding acquisition, Methodology, Resources, Formal analysis, Data curation, Software, Supervision, Writing - review & editing.

Conflict of interest statement:
The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Highlights

- Behaviourally, visual attentional control, as measured by task-set contingent attentional capture, was found to reach an adult-like state as early as age 7

- Behaviourally, attentional enhancement from multisensory stimuli was not found in 9-year-olds, but their EEG topographic patterns were different for multisensory vs. purely visual distractors

- Traditional visual attentional event-related analyses, such as the N2pc, are not sensitive to detect attentional enhancement for multisensory objects in adults, and visual or multisensory attention in children

- Sensitive, multivariate analyses of the event-related potential signal, such as electrical neuroimaging, are adept at revealing the neural underpinnings of attentional control processes over development

Abstract

Attentional control outside of the laboratory operates in multisensory settings, but the development of mechanisms subserving such control remains poorly understood. We investigated when, over the course of childhood, adult-like attentional control mechanisms begin to emerge. Children aged five, seven, and nine were compared with adults on behavioural performance in a computer game-like multisensory spatial cueing task, while simultaneous 129-channel EEG was recorded. Markers of attentional control were behavioural spatial cueing effects and the N2pc ERP component (analysed both traditionally and using a novel, multivariate electrical neuroimaging framework). In behaviour, adult-like visual attentional control was present from age 7 onwards, whereas multisensory control was not seen in children. In EEG, multivariate analyses of the activity over the N2pc time-window, revealed stable patterns of brain activity in children. Adult activity patterns linked to visual attentional control were present age 7 onwards, mirroring behaviour. Activity patterns linked to multisensory control were found in 9-year-olds, even though such patterns were not noted in behaviour. By combining rigorous yet naturalistic paradigms with multivariate signal analyses, we provided new insights into the development of visual attentional control skills vis-à-vis multisensory attentional control, thus generating a more complete account of attentional development.

Keywords: attentional control, development, multisensory, visual attention, N2pc, electrical neuroimaging

Introduction

Everyday environments are inundated with information and necessitate preferential selection of currently important information. However, such environments are also multisensory in nature. Here, we investigated whether adults and children pay attention similarly to visual and multisensory stimuli and through similar brain mechanisms.

1. Everyday environments are multisensory

The brain processes multisensory objects differently than typically studied unisensory stimuli, as it integrates information across the senses. Multisensory processes afford faster and more accurate behavioural responses (Stein & Meredith 1993; Murray & Wallace 2012), and benefits for learning and memory (Bahrick & Lickliter 2000; Lewkowicz, 2014). However, it remains unclear if attentional control mechanisms, derived from purely visual or auditory studies, operate similarly on unisensory and multisensory stimuli. The relative primacy between multisensory integration and attentional control is a topic of ongoing debate (e.g. Talsma et al., 2010; Buchan & Munhall, 2012; Matusz et al. 2015, 2019a,b). Our group's study involving audiovisual adaptation of the Folk et al. (1992) task has found that both target-colour and nonmatching distractors captured adults' visual attention more strongly, when paired with spatially diffused tones, despite the strong visual focus of participants' top-down task-set (Matusz & Eimer 2011). This suggests that purely unisensory attentional research may be limited in explaining how spatial attention is controlled in real-world, multisensory settings. While we know relatively little about how adults control their attention towards multisensory objects, we know even less about how children do so, and whether similar control mechanisms are present at different ages.

2. Developing attentional control is poorly understood

Behaviourally, children have weaker visual attentional control skills than adults (Donnelly et al., 2007; Trick & Enns, 1998; Gaspelin et al., 2015). There is mounting evidence that the prefrontal cortex and parietal cortex (i.e., the main hubs of the frontoparietal network) show protracted development (Casey, et al., 2005; Tsujimoto, 2008), as does the connectivity between them (e.g., Konrad et al., 2005; Hwang et al., 2010). However, our understanding of the mechanisms underpinning developmental differences remain limited. For instance,

children's weak attentional skills may be driven by slowly maturing facilitatory, memory-driven attentional control (Shimi et al. 2014a), but also, weak inhibitory control (e.g. Hommel et al., 2004). Reliable N2pc's have been found in 10-year-olds, evidencing their ability to direct attention to both upcoming visual objects (Couperus & Quirk, 2015) and objects held in visual memory (Shimi et al. 2015). Thus, the N2pc could offer valuable insights into the development of top-down attentional control mechanisms.

3. Are children more sensitive to multisensory distraction?

Like all neurocognitive processes, multisensory processes, too, undergo development. Sensitivity to congruence of onset, intensity, or identity are already present in infancy (Lewkowicz & Turkewitz, 1980; Bahrick & Lickliter, 2000; Neil et al. 2006). Other processes related to perceptual estimation or sensorimotor skills seem to mature slowly (8 and 10-11 years, Gori et al. 2008, 2012; and Barutchu et al. 2009, respectively). Yet, multisensory benefits for incidental learning have been reported in children as young as 5 (Broadbent et al. 2018, 2019). As such, existing research offers no clear resolution on whether children are more sensitive to multisensory distraction than adults. At what age do children and adults start to utilise similar neurocognitive control mechanisms when attending to visual and multisensory information? As indirect evidence, we previously showed that audiovisual distractors interfere with visual search in adults and 11-year-olds, but not 6-year-olds, for both colour (Matusz et al. 2015) and numerals (Matusz et al. 2019). In these studies, however, distractors were always task-relevant, in that on a significant proportion of trials, they shared the target's identity. Thus, it is as yet to be established whether children are more or less sensitive to completely task-irrelevant stimuli.

4. The present study

We developed a child-friendly version of the multisensory adaptation of Folk et al.'s spatial cueing task (Matusz & Eimer 2011), and tested it systematically on 5-, 7-, and 9-year-olds, as well as young adults. This way, we could investigate the differences between adults and children in controlling their attention towards visual and audiovisual objects. We wanted to investigate which of the two attentional control processes (visual or multisensory) reaches an adult-like state earlier. Secondly, we also investigated whether traditional or multivariate electrical neuroimaging (EN) analyses of the N2pc component were more sensitive to

studying more naturalistic attentional control (attention gauged by multisensory stimuli), and to capturing changes therein from childhood into adulthood. We had the following predictions. Behaviourally, in adults, we expected to replicate task-set contingent visual attention capture ("TAC"; see Folk et al., 1992; Matusz & Eimer, 2011) and multisensory enhancement of attention capture ("MSE"; see Matusz & Eimer, 2011). In children, we expected to find TAC in older groups (Gaspelin et al. 2015), without clear age-group predictions for MSE. For canonically measured N2pc in adults (posterior contralateral electrodes, like PO7/8), we expected attenuated/eliminated N2pc for non-target matching distractors. For MSE in N2pc, we did not have strong predictions, as the only related study to date showed little evidence for N2pc to audiovisual distractors (Van der Burg et al. 2011). In children, N2pc, TAC and MSE effects were investigated in an exploratory fashion. w. This is because, first, our oldest child group was younger than the youngest children where N2pc was reported (Couperus & Quirk 2015). Second, the N2pc in our study is recorded to distractors (cf. to targets in all previous studies), which may further create sub-optimal conditions for its detection. Contrastingly, we predicted that the sensitive, multivariate EN analyses will reveal modulations of brain responses by visual or multisensory control in adults, and at least in older groups of children. We have recently shown that combining the added benefits of EN measures and well-understood EEG correlates of cognitive processes allows for distinguishing between different cognitive accounts of multisensory attentional control (Matusz et al. 2019b).

Methods

1. Participants

A total of 115 primary school children participated in the study, 28 of whom were enrolled in fifth grade, 46 in third grade, and 41 in first grade of primary school in the canton of Vaud, Switzerland. In the local school system, children enter formal education at age 4, but only begin to sit in benches and receive less play-oriented instruction in third grade, when they are aged 6-7. To reduce confusion due to school system specificities, each child group is referred to in this manuscript by their members' majority age, that is: '5-year-olds', '7-year-olds', and '9-year-olds'. Children were recruited from local schools, nurseries, public events and entertainment facilities. Recruitment took place in the period from March 2017 to May

2019. Of the total number of children recruited, 18 were excluded for failure to initiate the testing session or failure to complete the task with above chance-level accuracy (50%), thus excluding one 9-year-old, six 7-year-olds, and eleven 5-year-olds respectively. Finally, 5 additional participants (one 9-year-old, two 7-year-olds, and two 5-year-olds) were excluded because of unusable EEG signals due to excessive noise even after the two-step filtering process detailed below. Data for adult "controls" was taken from one task that was part of a larger study. Therefore, the final sample consisted of 92 children: 9-year-olds (*N* = 26, 10 male, M_{age}: 8y 10mo, *SD:* 5mo, range: 8y 1mo – 10y 1mo), 7-year-olds (*N* = 38, 18 female, M_{age}: 6y 10mo, *SD:* 4mo, range: 6y 1mo, 7y 9mo), and 5-year-olds (*N* = 28, 13 female, M_{age}: 5y, *SD:* 4mo, range: 4y– 5y 7mo), and 39 adults (14 male, M_{age}: 27y 6mo, *SD:* 4y, range: 22–38y). Participants of all ages had normal or corrected-to-normal vision and normal hearing, and had no history of sensory problems (e.g., related to vision or audition), neurological problems (e.g., epilepsy), neurodevelopmental disorders (e.g., autism, ADHD), or learning difficulties (e.g., dyslexia), as indicated by parental report for children, or by direct report for adults. No children had an FSIQ under 70 which would warrant exclusion, as confirmed by an overall cognitive functioning assessment (see below).

All research procedures were approved by the Cantonal Commission for the Ethics of Human Research (CER-VD). Informed consent was obtained from parents/caregivers and verbal assent was obtained from children before participating in the study.

2. Stimuli and procedure

All participants were tested at the Lausanne University Hospital Centre (CHUV). For children, the EEG session lasted between 1h and 1h30mins, including briefing, obtaining consent, the testing protocol, and breaks. For adults, the session took approximately 3h (part of a larger study), but the data used as part of this study were acquired within the first 1h-1h30, akin to the child protocol. Children's baseline overall cognitive level was also assessed during a separate session on a different day. Overall cognitive functioning was assessed with the Wechsler scale of intelligence for school-age children (WISC-V, Wechsler, 2014) and preschool (WPPSI-IV, Wechsler, 2012) depending of the child's age. We used the abbreviated full-scale intellectual quotient (FSIQ) that includes 4 subscales: Vocabulary, Matrix reasoning, Blocks and Similarities. After completing both sessions, children received a 30 Swiss franc voucher for a media store and parents/caregivers' travel costs were reimbursed.

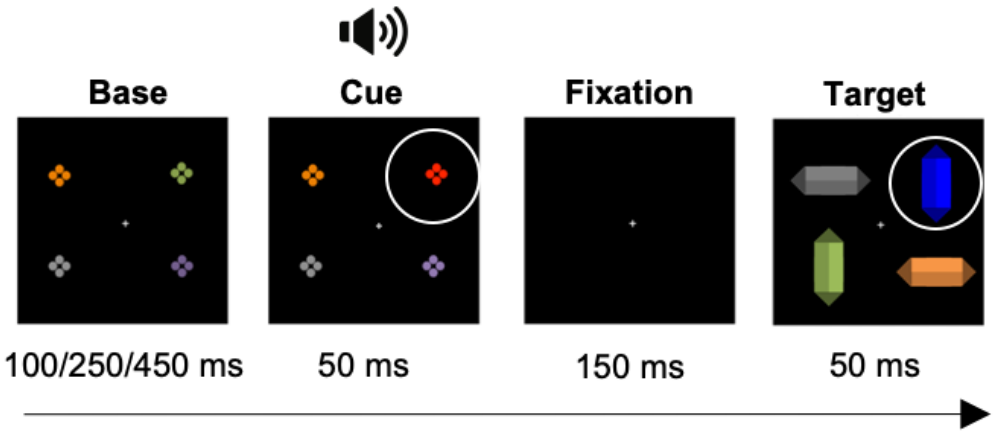

Figure 1. Experimental trial sequence for our child-friendly adaptation version of Matusz & Eimer (2011, Exp.2). The exemplary target blue diamond is preceded by a nontarget colour (NCC), i.e., red cue, both highlighted here by white circles (that did not appear in the experimental task). A spatially diffuse sound was presented together with the onset of the cue colour change (on 50% of all trials), creating an audiovisual nontarget colour distractor (NCCAV).

The EEG experiment was a child-friendly version of Matusz and Eimer's (2011, Exp.2) multisensory spatial cueing paradigm. Just as in the latter study, participants searched for a target defined by a single colour (e.g., a red bar) in a search array, which was preceded by an array containing spatially uninformative distractors. The distractor sets of 4 dots could match the target colour (red set of dots) or be another, nontarget colour (blue set of dots), and be presented alone or with spatially diffuse tones. We further adapted the task to be more age-appropriate and engaging for children in the following ways: 1) We introduced a game-like narrative, where participants had to help a pirate captain find treasure on a deserted island and moved along a treasure map after each completed block, 2) target bars were reshaped into "diamonds", 3) the number of elements in all arrays was reduced from 6 to 4 by removing 2 elements on the meridian (Figure 1). Each experimental trial consisted of the following sequence of arrays: base array, followed by cue array, followed by a fixation point, and finally a target array (Figure 1). The base array contained four differently coloured sets of closely aligned dots, each dot subtending 0.1° × 0.1° of visual angle. Each set element could be one of four possible colours (according to the RGB scale): green (0/179/0), pink (168/51/166), gold (150/134/10), silver (136/136/132). In a notable difference from the

original paradigm (Matusz & Eimer 2011, Exp.2), base array duration was no longer 450ms but rather it varied across trials (between 100, 250 and 450ms) to avoid building stimulus regularity-based predictions that could influence attentional control (Schwartze et al., 2011). In the cue array, one of the base array elements changed colour to either a target colour, or a nontarget colour that was not present in any of the elements before. The other 3 distractors remained one of the above four colours. The 2 cue colours were randomly selected with equal probability before each trial, and the colour change was not spatially predictive of the subsequent target location (same cue–target location on 25% of trials). On half of all trials, cue onset coincided with the onset of a pure sine-wave tone (2000Hz), presented from two loudspeakers on the left and right side of the monitor. Sound intensity was 80 dB SPL, as measured using a sound pressure meter as measured at the distance of the head using a CESVA SC-L sound pressure meter (CESVA, Barcelona, Spain). Thus, there were 4 cue conditions: TCCV (target colour-cue, Visual), NCCV (nontarget colour-cue, Visual), TCCAV (target colour-cue, AudioVisual), NCCAV (nontarget colour-cue, AudioVisual). The target array contained 4 rectangles where 1 was always the colour-defined target. The target "diamonds" and their preceding cues could have either a blue (RGB values: 31/118/220) or red (RGB values: 224/71/52) colour, and the target colour was counterbalanced across participants. The original uniformly coloured targets (Matusz & Eimer 2011, Exp.2) were given a diamond-like appearance by adding triangle shapes on the short sides of the bars and increasing and decreasing the luminance of certain sides of the bars by 20%.

Experimental sessions were conducted in a dimly lit, sound-attenuated room, with participants seated at a distance of 90 cm from a 23" LCD monitor with a resolution of 1080 × 1024 (60-Hz refresh rate, HP EliteDisplay E232). All visual elements were approximately equiluminant (~20cd/m^2), as determined by a luxmeter placed at a position adjacent to participants' eyes, measuring the luminance of the screen filled with each respective element's colour. The means of three measurement values were averaged across colours and transformed from lux to cd/m^2 in order to facilitate comparison with the results of Matusz & Eimer (2011). All elements were spread equidistally along the circumference of an imaginary circle against a black background, at an angular distance of 2.1° from a central fixation point.

Participants were instructed to find the target diamond of a predefined colour and to respond to its orientation (horizontal or vertical; randomly determined for each trial) by

pressing one of two horizontally aligned round buttons (Lib Switch, Liberator Ltd.) that were fixed onto a tray bag on their lap. Participants were told to answer as quickly and accurately as possible. To familiarise children with the task, a training block of 32 trials at 50% of regular task speed was administered. The subsequent full experimental session consisted of 8 blocks of 64 trials each, resulting in 512 trials in total. If participants did not respond within 5000ms of the target presentation, the next trial was initiated, otherwise the next trial was initiated immediately after a button press. Feedback on accuracy was given after each block, followed by the 'progress (treasure) map' which informed participants of the number of blocks remaining until the end, and during which participants could take a break and parents/caregivers could enter the testing room. To maintain motivation in younger participants, stickers on diamond-shaped sheets were offered during breaks following each session.

3. EEG acquisition and preprocessing

A 129-channel HydroCel Geodesic Sensor Net connected to a NetStation amplifier (Net Amps 400; Electrical Geodesics Inc., Eugene, OR, USA) was used to record continuous EEG data sampled at 1000Hz. Electrode impedances were kept below 50kΩ, and electrodes were referenced online to Cz. Offline filtering involved a 0.1 Hz high-pass and 40 Hz low-pass as well as 50 Hz notch and a second-order Butterworth filter (–12 dB/octave roll-off, computed linearly with forward and backward passes to eliminate phase-shift). Next, the EEG was segmented into peri-stimulus epochs from 100ms before cue onset to 500ms after cue onset. Epochs were then screened for transient noise, eye movements, and muscle artefacts using a semi-automated artefact rejection procedure. It has been noted previously that due to physiological differences between children and adults' skulls and brains, these two groups require different artefact rejection criteria to prevent discarding clean EEG signal (Scerif et al., 2006; Shimi et al., 2015). Therefore, as in previous event-related potentials (ERP) research on developing populations (e.g., Melinder et al., 2010; Shimi et al., 2014b), we used an automatic artefact rejection criterion of $\pm 100\ \mu V$ for adults and $\pm 150\ \mu V$ for children, along with visual inspection. For children, additionally, only EEG data from trials with correct responses, and from blocks with over 50% accuracy were used, to fit behavioural data. Data from artefact contaminated electrodes across all groups were interpolated using three-dimensional splines (Perrin et al., 1987). Average numbers of epochs removed, and

electrodes interpolated per participant in each age group can be found in Supplementary materials: EEG preprocessing.

Cleaned epochs were averaged, baseline corrected to the 100ms pre-cue time interval, and re-referenced to the average reference. Next, due to persistent environmental noise present in the majority of the child and adult datasets even after initial filtering, an additional 50Hz notch filter was applied. All of the above steps were done separately for ERPs from the four cue conditions, separately for cues in the left and right hemifield. To analyse cue-elicited lateralised ERPs, single-trial data from all conditions with cues presented on the left were relabelled to have electrodes over the left hemiscalp represent activity over the right hemiscalp, and vice versa. After relabelling, the "mirror cue-on-the-right" single-trial data and the veridical "cue-on-the-right" data were averaged together, creating a single lateralised average ERP for each of the 4 cue conditions. As a result of this, we obtained 4 different ERPs, one for each of the 4 conditions (TCCV, NCCV, TCCAV, NCCAV). All preprocessing and EEG analyses, unless otherwise stated, were conducted using CarTool software (available for free at www.fbmlab.com/cartool-software/; Brunet, Murray, & Michel, 2011).

4. Data analysis design

As we previously found both task-set contingent visual attention capture (TAC) and multisensory enhancement of attention capture (MSE) in adults (Matusz & Eimer, 2011), we used these as behavioural markers of top-down visual and bottom-up multisensory control processes. Next, we combined traditional N2pc component analyses with an electrical neuroimaging framework (EN).

4.1. Behavioural analyses. Analyses were focused on reaction-time (RT) spatial cueing effects. This measure was derived by subtracting the mean RTs for trials where the cue and target were in the same location from the mean RTs for trials where the cue and target location differed, separately for each of the 4 cue conditions, following Matusz & Eimer (2011). Error rates were also analysed, in the form of percentages. Before the analysis, RT data were cleaned following a two-step procedure. First, blocks with mean accuracy below chance level (50%) were removed. Thus, in children, 15% of all blocks were removed (3% for 9-year-olds, 7% for 7-year-olds, and 37% for 5-year-olds respectively). In adults, all blocks were used due to high overall accuracy (>95%). Next, RT data from the remaining

blocks was cleaned following the procedure of Gaspelin et al. (2015). Specifically, incorrect and missed trials were discarded, as were trials with RTs below 200ms and above 1000ms for adults, and below 200ms and above 5000ms for children. Moreover, all RTs above 2.5 *SD*s from individual participant's mean RTs were also removed. Overall, 26% of all trials were removed (6% in adults, 28% in 9-year-olds, 29% in 7-year-olds, and 40% in 5-year-olds). Next, to verify if RT spatial cueing modulations were preserved after correcting for children's general cognitive slowing, each individual's RT's per condition was divided by their average overall RT, and then converted to a percentage (following Gaspelin et al., 2015, see also Maylor & Lavie, 1998). 'Raw' and scaled RT data were normally distributed, and thus submitted to separate mixed-design four-way repeated-measures ANOVAs with one between-subject factor of Age (adults vs. 9-year-olds vs. 7-year-olds vs. 5-year-olds), 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). Next, as part of follow-up analyses, data for each age-group were submitted to separate repeated-measures ANOVAs with within-subject factors: Cue Colour, Cue Modality, and Cue-Target Location. Error data were not normally distributed, and thus analysed using separate three-way Friedman tests for each child group, with factors Cue Colour, Cue Modality, and Cue-Target Location. In the case of adult control data, we conducted a three-way Durbin test instead, with factors Cue Colour, Cue Modality, and Cue-Target Location. All analyses, including post-hoc paired *t*-tests, were conducted using SPSS for Macintosh 26.0 (Armonk, NY: IBM Corp).

4.2. Overview of ERP analyses. Given that the N2pc is a well-studied and well-understood correlate of attentional control in visual contexts and in adult populations, we began our ERP analyses by conducting a canonical N2pc analysis on the contralateral and ipsilateral average ERPs elicited across the 4 cue conditions. This way, we could compare N2pc's elicited by our visual and audiovisual distractors with other extant research in adults and children. Additionally, these analyses helped us better bridge previous and present traditional N2pc results with the present study's electrical neuroimaging (EN) analyses of the N2pc. Briefly, EN encompasses a set of multivariate, reference-independent analyses of global features of the electric field measured at the scalp (Lehmann & Skrandies 1980; Murray et al., 2008; Tivadar & Murray 2019), providing robust, direct neurophysiologically interpretable results

about differences between sets of conditions, groups, or timepoints. Thus, we used data from the entire 129-channel electrode montage in our analyses of the contralateral versus ipsilateral ERP voltage gradients within an EN approach.

Since the aim of the current study was to identify the emergence of adult-like attentional control mechanisms in childhood, all ERP analyses in developmental groups followed a 'normative' framework. That is, the parameters typically used for canonical analyses of the N2pc in adult visual attention research were applied to analyse children's ERP data. Below we detail the two types of analyses and how they were conducted across age groups.

*4.2.1. Canonical N2pc analysis***.** We extracted the mean amplitude for each of the 4 cue conditions within a prescribed time-window, separately for the contralateral and the ipsilateral posterior electrodes. Regarding electrode pair choice, EGI 129-channel equivalents of the canonical PO7/8 electrodes (e.g., Eimer et al., 2009; Kiss et al., 2008a, 2008b) were electrodes e65 and e90. Regarding the time-window, we used the period of 180-300ms after stimulus onset (e.g., Luck & Hillyard, 1994; Eimer, 1996; Eimer & Kiss, 2008; Eimer 2014). We used these criteria to extract mean amplitudes for each of the 8 ERPs (4 cue conditions for ipsilateral and contralateral electrode each) for each of the four age groups. We then submitted these mean amplitude values to separate 3-way repeated measures ANOVAs, with within-subject factors: Cue Colour (TCC vs. NCC), Cue Modality (V vs. AV), and Contralaterality (Contralateral vs. Ipsilateral). For comparison, we also analysed the same data, choosing the electrode sites and time-window for extraction of mean amplitude values following a more data-driven approach (see Supplementary materials: Supplemental N2pc results).

*4.2.2. Electrical neuroimaging of the N2pc component***.** Important spatially-selective modulations of brain responses by attentional control mechanisms might be present across different conditions within the contralateral-to-ipsilateral gradients of voltage potentials (see Matusz et al. 2019b). These gradients (and their modulations) are not captured by canonical N2pc analyses as the latter analyse the mean difference in voltage between 1 contralateral electrode/electrode subset and 1 ipsilateral electrode/electrode subset (out of a set of >20 or >100 electrodes). To analyse the global mechanisms governing these response gradients instantiated by visual control and multisensory control, we first computed a difference ERP, by subtracting the voltages over the contralateral hemiscalp and

the voltages over the ipsilateral hemiscalp, separately for each of the 4 cue conditions. This resulted in a 59-channel lateralised ERP (as the midline electrodes from the 129-electrode montage were not informative). Next, this difference ERP was mirrored onto the other side of the scalp, recreating a "fake" 129-electrode montage (with values on midline electrodes set to 0). It was on these mirrored "fake" 129-electrode lateralised difference ERPs that we performed the EN response strength and topography analyses, across the 4 cue conditions, for each of the 4 age groups.

4.2.2.1. Strength-based modulations of the difference "N2pc-like" ERPs. The first step in our EN analyses was using Global Field Power (GFP) 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 networks within any of the age groups. Visual attentional effects observed in N2pc are traditionally assumed to arise from a modulation in response strength, i.e., as a result of a "gain-control" mechanism. In an EN framework, such differences would be readily detected as GFP differences between experimental conditions over the N2pc time-window (for more info on how GFP can grasp differences in the brain mechanisms behind the N2pc, see Matusz et al. 2019b).

To reiterate, we analysed the mean GFPs in the 129-channel "fake" difference ERPs from the 4 cue conditions, separately for each age group. GFP is a timepoint-to-timepoint standard deviation of voltage across the scalp and can be plotted as a single waveform, just like a regular single waveform. To mirror our canonical N2pc analyses in adults, we extracted the average GFP amplitudes measured within the canonical adult N2pc time-window of 180-300ms post-cue. We then submitted each age group's 4 averages to separate 2 x 2 repeated-measures ANOVAs with Cue Colour (TCC vs. NCC) and Cue Modality (V vs. AV) as within-subject factors.

4.2.2.2. Topographic modulations of the difference "N2pc-like" ERPs. Next, we investigated whether there were differences across the 4 cue-elicited lateralised difference ERPs that were driven by changes in electric scalp field topography. Such changes would indicate that visual or multisensory control modulate spatially-selective brain responses by activating distinct configurations of brain generators. Differences in two electric fields, independent of their strength, are indexed by DISS, and applying data clustering methods onto DISS over the time course of an ERP can reveal periods of tens to hundreds of milliseconds of stable topographic activity, i.e., topographic "maps", within that ERP

(elsewhere referred to as "functional microstates", e.g., Michel & Koenig, 2018). The present topographic analyses used a hierarchical clustering method, specifically, Topographical Atomize and Agglomerate Hierarchical Clustering (TAAHC) in order to identify topographic maps ("template maps" henceforth) in the group-averaged mirrored difference ERPs data in the time-interval following cue onset. Briefly, the TAAHC produces, in a series of iterative steps, configurations of clusters that explain certain amounts of global explained variance (GEV) in the ERP data (for a more detailed explanation see Matusz et al. 2019b; Murray et al., 2008). The optimal configuration of clusters is the smallest number of template maps accounting for the largest amount of variance in the grand-averaged electrophysiological responses. There are several criteria that help identify this number. In the present study we used the modified Krzanowski–Lai criterion as well as the Cross Validation index, and Dispersion (Murray et al., 2008).

As part of our "normative" age-based EN analyses, first we applied the TAAHC to the group-averaged *adult* ERP data and identified the optimal number of clusters that explain most of the adult ERP variance. Next, we tested to what extent the stable pattern of EEG activity ("template maps") seen in adults are present within the ERPs in the younger groups, and how this involvement differs by age group. That is, for each age group separately, we investigated whether, and if so, how strongly each of the clusters identified in the adult group-averaged difference ERPs were present in the single-subject ERP developmental data (the so-called "fitting" procedure). Specifically, every time-point over the adult N2pc time-window in the cue-induced mirror difference ERPs of each tested child was labelled by that adult topographical map with which it best correlated spatially. The final output for each participant was the number of timeframes (here in milliseconds) that each adult topographical map characterised the child's ERP in adult canonical N2pc time-window. These durations (in milliseconds) were submitted to separate three-way 2 x 2 x 4 repeated-measures ANOVAs in each age group, with factors: Cue Colour (TCC vs. NCC) and Cue Modality (V vs. AV), and Map (Map1 vs. Map2 vs. Map3 vs. Map4) followed by post-hoc *t*-tests. Maps with durations under 10 contiguous timeframes were not included in the analyses. Greenhouse-Geisser corrections were applied where necessary to correct for violations of sphericity. Unless otherwise stated in the results, map durations were statistically different from 0ms (as confirmed by post-hoc *t*-tests), meaning that they were reliably present across the time-windows of interest. Throughout the results, Holm-

Bonferroni corrections were used to correct for multiple comparisons between map durations. Comparisons passed the correction unless otherwise stated.

Results

1. Behavioural analyses

We first report the results of the 'raw' RT data analysis for each age group, followed by the results of the analysis of RTs corrected for slower processing speed.

1.1. 'Raw' reaction times. Mean RTs sped up progressively from 5-year-olds (1309ms) through 7-year-olds (1107ms) and 9-year-olds (836ms) to adults (594ms), which was reflected in a significant main effect of Age, $F_{(3, 127)} = 94.7$, $p < 0.001$, $\eta_p^2 = 0.7$. Here, 5-year-olds were reliably slower than 7-year-olds ($t_{(33)} = 4.4$, $p < 0.001$), who were slower than 9-year-olds ($t_{(32)} = 5.7$, $p < 0.001$), who were in turn, slower than adults ($t_{(32.5)} = 5.1$, $p < 0.001$). However, Age did not interact with any other factors (all *F's* < 2, *p*'s > 0.1). Nonetheless, to adequately investigate differences between adults and children, and the developmental trajectory of attentional control processes, we analysed the raw RT data from each age group separately.

Firstly, in adults, there was a significant main effect of Cue Colour, $F_{(1, 38)} = 36.9$, $p < 0.001$, $\eta_p^2 = 0.5$, driven by faster responses on trials with target colour-cues (TCC, 607ms) than on trials with nontarget colour-cues (NCC, 618ms). Adults also showed generally faster responses on trials with sounds (AV, 605ms) than with no sounds (V, 620ms), $F_{(1, 38)} = 76.1$, $p < 0.001$, $\eta_p^2 = 0.7$. Overall behavioural capture effects in adults were reliable, i.e. responses were faster for trials where the cue and target location were the same (600ms) versus when they were different (624ms), $F_{(1, 38)} = 110.9$, $p < 0.001$, $\eta_p^2 = 0.8$. Further, as in the original Matusz and Eimer's (2011) study, the adults' overall behavioural capture effects differed depending on the colour of the cue, as shown by a 2-way Cue-Target Location x Cue Colour interaction, $F_{(1, 38)} = 161.5$, $p < 0.001$, $\eta_p^2 = 0.8$ (this is the TAC effect). This effect was driven by statistically significant behavioural capture effects for the TCC condition (48ms, $t_{(38)} = 16.7$, $p < 0.001$), but not the NCC condition (1ms, $t_{(38)} = 0.2$, $p = 0.8$; Figure 2, top left panel, and Figure 3 top left panel). Again, as in the original 2011 study, behavioural capture effects

also differed when elicited by visual and audiovisual distractors, as shown by a two-way interaction between Cue-Target Location and Cue Modality, $F_{(1, 38)} = 4.9$, $p = 0.03$, $\eta_p^2 = 0.1$ (this is the MSE effect). This effect was driven by larger behavioural capture effects elicited by AV (26ms, $t_{(38)} = 10.8$, $p < 0.001$) than by V cues (21ms, $t_{(38)} = 8.9$, $p < 0.001$; Figure 2, top left panel, and Figure 3 top left panel). The Cue Colour by Cue Modality interaction ($F < 1$) was not significant, and neither was the Cue-Target Location x Cue Colour x Cue Modality interaction ($F < 3$, $p > 0.1$). These results demonstrated that adults showed both reliable TAC and MSE in behaviour, replicating Matusz and Eimer (2011).

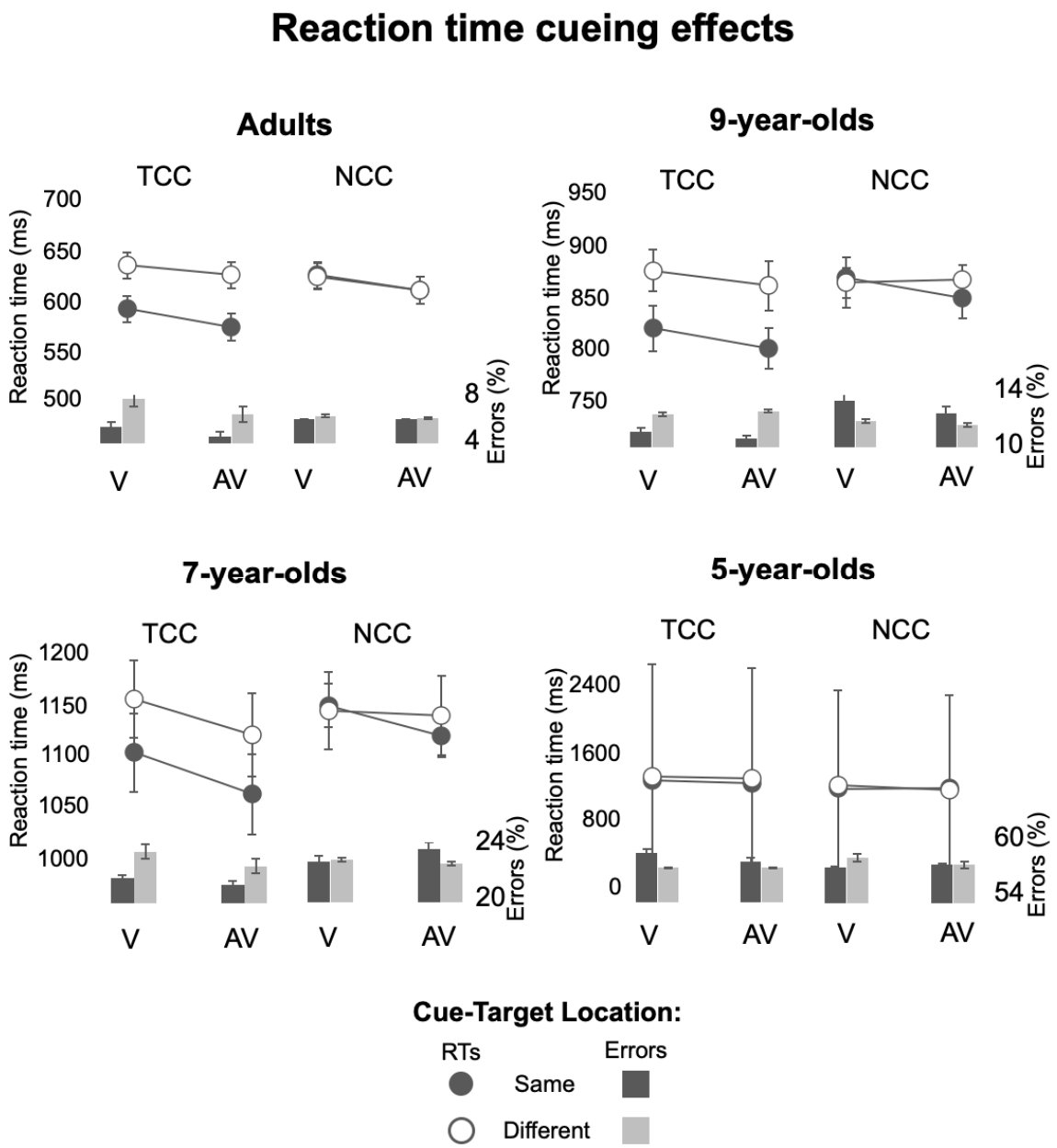

Figure 2. Mean reaction times shown for each of the 4 age groups on trials where Cue-Target Location was the same versus different, shown separately for target colour-cue (TCC) and nontarget colour-cue (NCC) trials, as well as visual (V) and audiovisual (AV) trials. Line graphs show the mean RTs, bar graphs show error rates (in percentages), and error bars represent the standard error of the mean.

Like adults, 9-year-olds responded faster on TCC trials (843ms) than on NCC trials (865ms), $F_{(1, 25)} = 28.4$, $p < 0.001$, $\eta_p^2 = 0.5$. Their overall behavioural capture effects were also reliable, with faster RTs for trials where the cue and target location were the same (839ms) versus

when they were different (870ms), $F_{(1, 25)} = 68.9$, $p < 0.001$, $\eta_p^2 = 0.7$. Overall speeding up of responses on AV compared to V trials now showed the level of a nonsignificant trend ($F_{(1, 25)} = 0.3$, $p = 0.08$, $\eta_p^2 = 0.1$). However, the main question was whether behavioural capture effects in 9-year-old children would be modulated by the cues' matching of the target colour, as well as the audiovisual nature of the cues. Notably, like in adults, did indeed show TAC, as evidenced by a 2-way interaction between Cue-Target Location and Cue Colour, $F_{(1, 25)} = 19.5$, $p < 0.001$, $\eta_p^2 = 0.4$. This interaction was driven by significant capture effects for the TCC distractors (56ms, $t_{(25)} = 8.3$, $p < 0.001$), but not for the NCC distractors (6ms, $t_{(25)} = 0.9$, $p = 0.7$; Figure 2, top right panel, and Figure 3 top right panel). However, in contrast with adults, 9-year-olds did not show MSE, with no evidence for visually-elicited capture effects enlarged on AV vs. V trials, i.e., no 2-way Cue-Target Location x Cue Modality interaction, $F_{(1, 25)} = 1.4$, $p = 0.3$. Other interactions failed to reach statistical significance (All F's < 2, p's > 0.1). With this, we can conclude that 9-year-olds showed reliable TAC, but not MSE, in behaviour.

In 7-year-olds, like in adults, responses were faster for trials with TCC cues (1112ms) than for NCC cues (1138ms), $F_{(1, 37)} = 18.7$, $p < 0.001$, $\eta_p^2 = 0.3$, and were also faster for trials with AV cues (1111ms) than V cues (620ms), $F_{(1, 37)} = 8.6$, $p = 0.006$, $\eta_p^2 = 0.2$. Further, overall capture effects were again reliable, with faster responses on cue-target location same (1109ms) versus different (1140ms) trials, $F_{(1, 37)} = 14$, $p < 0.001$, $\eta_p^2 = 0.4$. Just as in the two older groups, 7-year-olds, did show TAC, as shown by a Cue-Target Location x Cue Colour interaction, $F = 6.4$, $p = 0.02$, $\eta_p^2 = 0.2$. This was driven by reliable cueing effects elicited by TCC distractors (55ms, $t_{(37)} = 4.8$, $p < 0.001$), but not by NCC distractors (7ms, $t_{(37)} = 0.6$, $p = 1$; Figure 2, bottom left panel, and Figure 3 bottom left panel). However, as in 9-year-olds, 7-year-olds' visually-induced attentional capture effects did not show MSE, with no 2-way Cue-Target Location x Cue Modality interaction failing to reach significance, $F_{(1, 37)} = 2.1$, $p = 0.2$. Other interactions also did not reach statistical significance (All F's < 2, p's > 0.1). It thus appeared that 7-year-olds, like 9-year-olds before them, showed reliable TAC, but not MSE.

In 5-year-olds, as in the other age groups, we observed reliable overall attentional capture effects $F_{(1, 27)} = 14$, $p < 0.001$, $\eta_p^2 = 0.4$, driven by faster responses for cue-target location same (1312ms) versus different (1343ms) trials. However, there was no evidence for either of the two key interactions, specifically, the Cue-Target Location x Cue Colour interaction ($F_{(1, 27)} = 1.4$, $p = 0.2$), or the Cue-Target Location x Cue Modality interaction ($F_{(1,}$

$_{27)}$ = 0.4, p = 0.5). In further contrast with the older age groups, overall RTs were not affected by the colour of the cue, as shown by a nonsignificant main effect of Cue Colour, $F_{(1, 27)}$ = 2.6, p = 0.1. In one final contrast, faster responses on AV versus V trials showed only a nonsignificant trend, $F_{(1, 27)}$ = 3.5, p = 0.07, η_p^2 = 0.1. No other interactions reached statistical significance (All F's < 2, p's > 0.1). The 5-year olds, therefore, did not show reliable TAC nor MSE in behaviour.

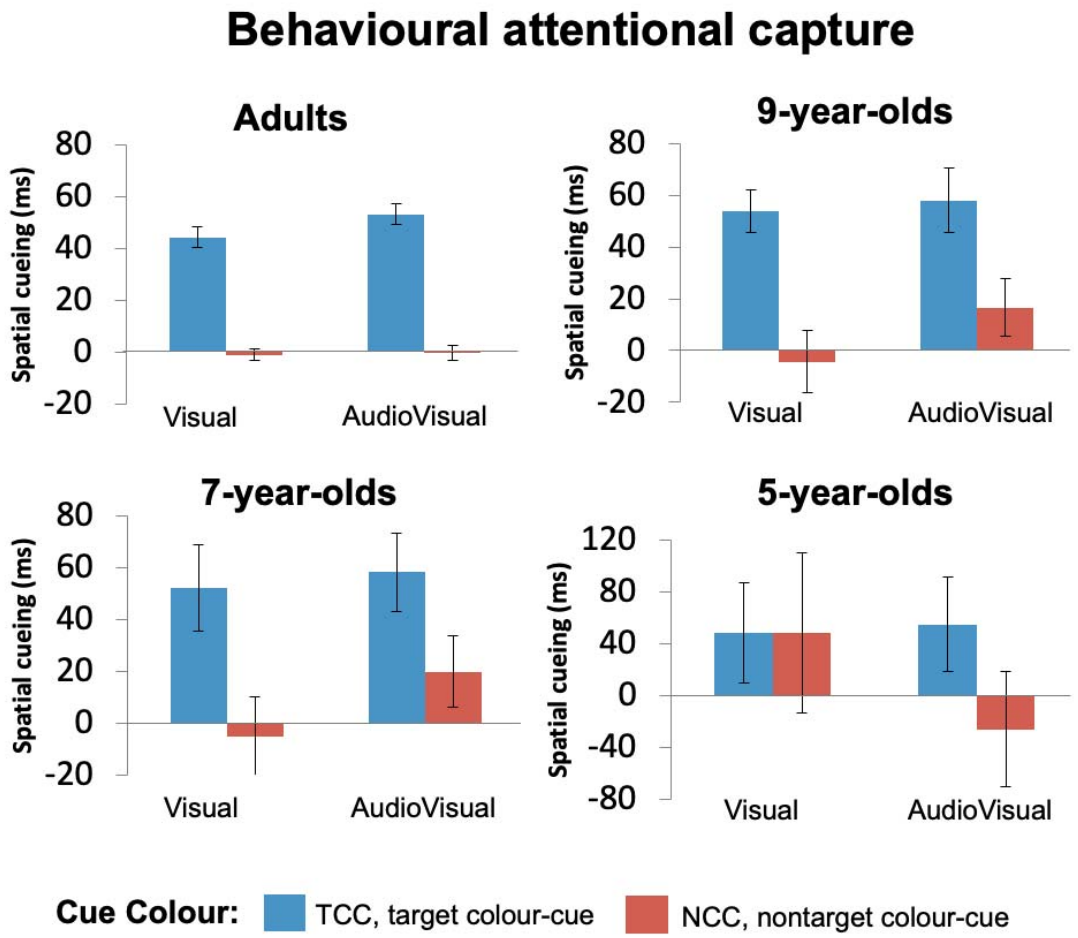

Figure 3. Bars coloured according to the figure legend in the image represent behavioural attentional capture indexed by mean RT spatial cueing effects, and error bars represent the standard error of the mean. Adults, 9-year-olds, and 7-year-olds all showed presence of top-down visual attentional control, exemplified by TAC. Specifically, all 3 age groups showed reliable attentional capture effects for target colour-cues, but not for nontarget colour-cues. In contrast, only in adults, attentional capture showed MSE.

1.2. RTs corrected for children's cognitive slowing. All of the child groups showed the same patterns of results as in the raw RT analyses. That is, 9-year-olds and 7-year-olds showed TAC but not MSE, and 5-year-olds did not show TAC or MSE. For brevity, we have relegated the full results of these statistical analyses to Supplementary materials: Supplemental behavioural results. These analyses demonstrated that, even after having corrected for children's overall cognitive slowing, no children exhibited MSE, and only older children exhibited TAC.

1.3. Error rates. Since error data were not normally distributed, we conducted a 1-way Kruskal–Wallis *H* test to test for differences between groups, and 3-way Friedman tests (or Durbin tests where there were no errors for a given condition) to test for differences within each age-group. Overall, error rates were highest in the youngest children (57%), and steadily reduced in 7-year-olds (23%), followed by 9-year-olds (12%), culminating in the smallest error rates in adults (6%), $\chi^2(3) = 81.4$, $p < 0.001$. In adults, error rates were modulated by Cue-Target Location $\chi^2(1) = 8.7$, $p = 0.003$, such that fewer errors were made on trials where the cue and target location was the same (5.5%) than when they were different (6.6%). Error rates were not significantly modulated by Cue Colour or Cue Modality (all *p*'s > 0.1). In 9-year-olds, 7-year-olds, and 5-year-olds alike, error rates were not significantly modulated by Cue-Target Location, Cue Colour or Cue Modality (all *p*'s > 0.1).

2. ERP analyses

2.1. Canonical N2pc analysis. In adults, the presence of reliable overall N2pc's across the canonical electrodes and canonical time-window was supported by a statistically significant main effect of Contralaterality, $F_{(1, 38)} = 17.8$, $p < 0.001$, $\eta_p^2 = 0.3$, where the mean contralateral amplitude, over the PO7 equivalent (-0.4µV), was larger than the ipsilateral amplitude, over the PO8 equivalent (0.1µV). This result suggested the presence of a reliable N2pc. Consequently, the contra-ipsilateral difference had a mean overall amplitude of -0.5µV. As expected, cue-elicited N2pc's differed in their magnitude depending on the cue colour, as supported by a Contralaterality x Cue Colour 2-way interaction, $F_{(1, 38)} = 17$, $p < 0.001$, $\eta_p^2 = 0.3$. This interaction was driven by a reliable N2pc for target colour-cues (-0.69µV; **Error! Reference source not found.**A, top and bottom left panels) but not for nontarget colour-cues (-0.25µV; Figure 4, top left and right panels). This result demonstrated presence of TAC in adult N2pc's. However, there was no evidence for a reliable difference in

mean N2pc amplitudes across V and AV cues, with no Contralaterality x Cue Modality 2-way interaction ($F < 1$), and therefore, no evidence for MSE.

Interestingly, the N2pc amplitudes elicited by TCC and NCC cues were modulated by whether they were presented alone or with sounds, as shown by a 3-way interaction between Contralaterality, Cue Colour, and Cue Modality, $F_{(1, 38)} = 8$, $p = 0.007$, $\eta_p^2 = 0.2$. We first analysed this interaction as a function of Cue Modality. First, for AV cues, mean N2pc amplitudes elicited by TCCAV were larger (-0.8µV) than mean amplitudes for NCCAV cues (-0.2µV), $t_{(38)} = 5$, $p < 0.001$. In contrast, for V cues, there was no statistically significant difference in mean N2pc amplitudes elicited by NCCV cues (-0.3µV) and TCCV cues (-0.6µV), $t_{(38)} = 1.8$, $p = 0.2$. When we analysed the 3-way interaction as a function of Cue Colour, for both TCC and NCC distractors, differences in mean N2pc amplitude between AV and V were at the level of a nonsignificant trend ($t_{(38)} = 1.8$, $p = 0.06$, and $t_{(38)} = 1.4$, $p = 0.07$, respectively). Other effects did not reach statistical significance (All F's < 1), except the main effects of Cue Colour, $F_{(1, 38)} = 8.4$, $p = 0.006$, $\eta_p^2 = 0.2$ (driven by larger ERP amplitudes for TCC -0.3µV, than for NCC -0.03µV, and Cue Modality, $F_{(1, 38)} = 7.1$, $p = 0.011$, $\eta_p^2 = 0.2$ (driven by larger ERP amplitudes for V, -0.3µV, than for AV, 0.06µV). Thus, although MSE was not observed in N2pc's, adult's overall ERP data was jointly modulated by visual and multisensory attentional control. This effect seemed to be driven by reliable difference between TCC and NCC distractors on trials where distractors were AV but not V.

For the child age groups, 2 x 2 x 2 repeated-measures ANOVAs were conducted on mean amplitude values from adult electrodes over the adult time-window. In no child group was there a significant main effect of Contralaterality (9-year-olds: $F_{(1, 25)} = 0.4$, $p = 0.6$; 7-year-olds: $F_{(1, 37)} = 0.04$, $p = 0.8$; 5-year-olds: $F_{(1, 27)} = 0.2$, $p = 0.6$; Figure 4, 2nd, 3rd, and 4th panels), and therefore, no N2pc. For this reason, we will not report other results unrelated to Contralaterality (they are available in Supplementary Materials: Supplemental N2pc results). To rule out the possibility that a lack of effects in children was due to literature-based values being suboptimal, we conducted an additional analysis where the N2pc time-window and electrode sites were selected from the adult data in a more data-driven fashion. We report the details of the procedure and results in Supplementary Materials: Supplemental N2pc results. Crucially however, this approach also showed no significant main effect of Contralaterality (All F's < 1), and thus no presence of an N2pc.

Contralateral-Ipsilateral waveforms

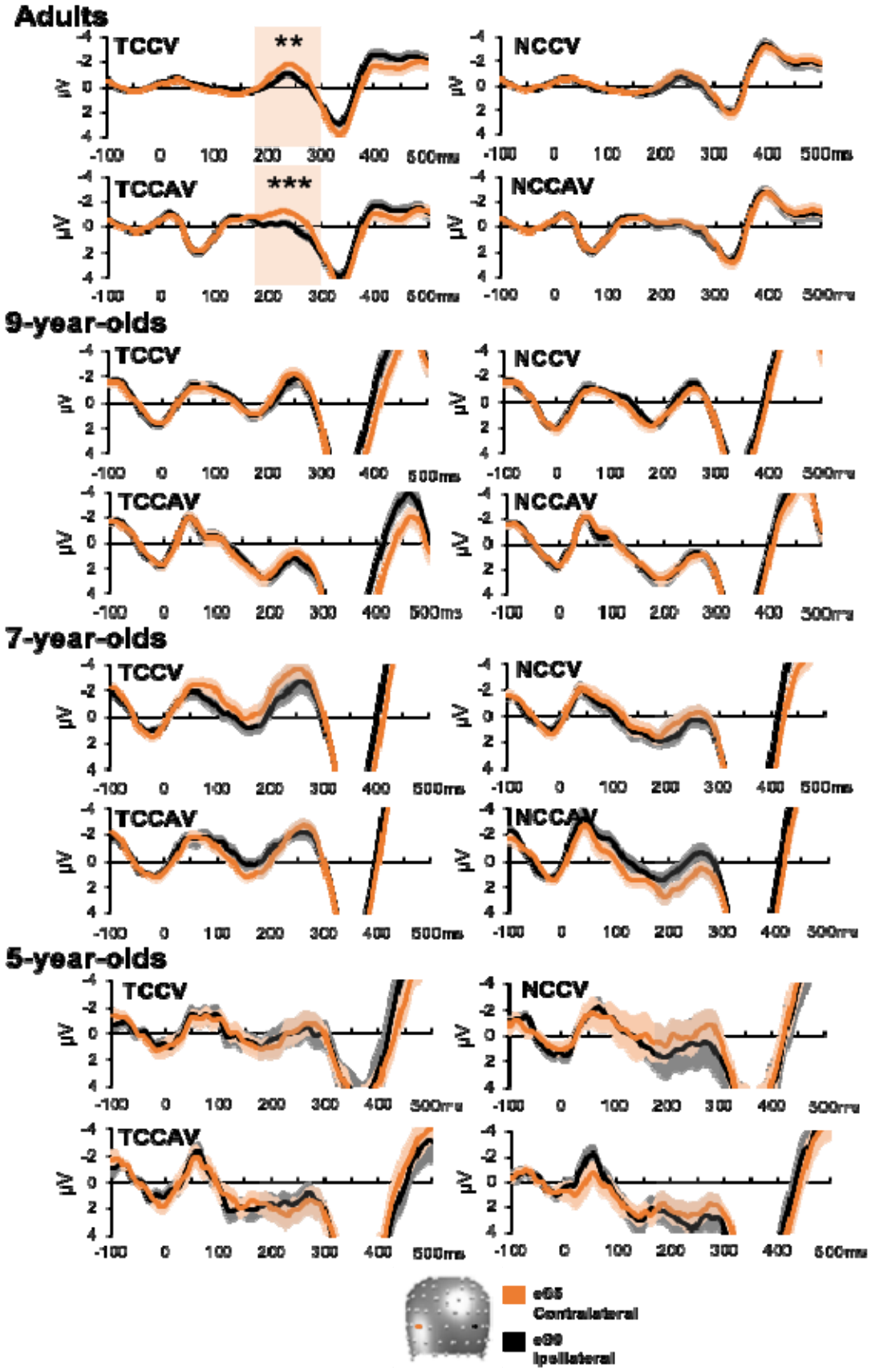

Figure 4. N2pc waveform results. Mean amplitude values are shown at contralateral and ipsilateral electrode sites, indicated in orange and black, per the head model and legend on the figure. The N2pc time-window of 180-300ms is highlighted in light orange, where the contra-ipsi difference is significant, and light grey where it is not. Significance levels are denoted as follows: ** > .01, *** > .001. Adults show significant contra-ipsi differences, that is reliable N2pc's, for target-colour cues (TCC) but not nontarget colour-cues (NCC). In children, there was no reliable N2pc in any of the four conditions.

2.2. Electrical neuroimaging of the N2pc component. An ANOVA on the average GFP values per condition revealed no significant main effects or interactions in adults, 9-year-olds, 7-year-olds, or 5-year-olds (All *F*'s < 1). Full results can be found in Supplementary Materials: Supplemental GFP results. For graphical representations of the GFP results, we direct the reader to Supplemental Figure 1 in the Supplementary Materials: Supplemental Figures.

The segmentation of the post-cue period of the adult data revealed 9 clusters which explained 82.8% of the GEV in the group-averaged ERPs. We remind the reader that topographical analyses (unless otherwise stated) were conducted on difference ERPs, which accounts for the lower rates of GEV. Next, a fitting procedure on the adult single-subject data revealed 4 template maps which characterised the N2pc time-period of 180-300ms post-cue. A 2 x 2 x 4 ANOVA on the mean durations of the 4 maps identified in the adult data revealed a main effect of Map, $F_{(3, 114)} = 18.3$, $p < 0.001$, $\eta_p^2 = 0.3$, where Map4 predominated (i.e. had the longest duration of all maps) the N2pc time-window across conditions (Figure 5, middle left panel). This demonstrated that adults had stable patterns of lateralised ERP activity. Hereafter, we did not follow up the main effect of Map with post-hoc tests, as it was not informative as to the presence of TAC or MSE in topography.

There was a 2-way interaction between Map and Cue Colour, $F_{(2.4, 89.1)} = 12$, $p < 0.001$, $\eta_p^2 = 0.2$. Following up this interaction by the factor of Cue Colour showed that Map4 was present longer in response to TCC (67ms) than to NCC distractors (40ms), $t_{(38)} = 5.2$, $p < 0.001$, while Map2 was present longer in response to NCC (34ms) than to TCC distractors (13ms), $t_{(38)} = 3.9$, $p = 0.004$. Other map durations did not differ significantly between TCC and NCC cues (all *p*'s > 0.1). Hereafter, map duration differences that are not reported here were not statistically significant (*p*'s > 0.1). Following up the interaction by the factor of Map revealed that for TCC cues, Map4 (67ms) was significantly longer than all other maps – Map1 (15ms), $t_{(38)} = 7.7$, Map2 (13ms), $t_{(38)} = 8$, and Map3 (25ms), $t_{(38)} = 6.3$, all *p*'s < 0.001, while

no map durations differed for NCC distractors (all p's > 0.1). These results suggest that Map4 drove the processing of TCC distractors, while no particular map was more implicated than others in the processing of NCC distractors. Finally, the map modulations by Cue Colour demonstrated here support the presence of TAC in adult ERP topography. Thus, it appeared that the Map x Cue Colour interaction was driven by modulations of Map2 and Map4 presence for different cue colours, where Map4 is especially implicated in the processing of target-colour cues.

In contrast to canonical N2pc analysis results, topographic map presence over the N2pc time-window interacted with Cue Modality, as evidenced by a 2-way interaction, $F_{(3, 114)} = 3.2$, $p = 0.027$, $\eta_p^2 = 0.1$. A follow-up by Cue Modality revealed that Map2 was present longer when cues were purely visual (V, 30ms) than when cues were accompanied by a sound (AV, 17ms) at the level of a nonsignificant trend, $t = 2.8$, $p = 0.08$. However, a follow-up by Map revealed that Map4 was present longer than any other maps for both AV cues (Map4 [53ms] vs. Map1 [25ms], $t_{(38)} = 4.3$, Map2 [17ms], $t_{(38)} = 5.6$, Map3 [26ms], $t_{(38)} = 4.2$, all p's < 0.001), and for V cues (Map4 [54ms] vs. Map1 [18ms], $t_{(38)} = 5.7$, Map2 [30ms], $t_{(38)} = 3.7$, Map3 [19ms], $t_{(38)} = 5.5$, all p's < 0.001). Taken together, it appeared that Map2 may be implicated in topographic modulations of lateralised ERPs by Cue Modality, whereas Map4 was the main map driving the processing of both AV and V cues.

Finally, the 3-way Map x Cue Colour x Cue Modality interaction was significant, $F_{(3, 114)} = 5.4$, $p = 0.002$, $\eta_p^2 = 0.1$. When followed up as a function of Cue Colour, for NCC distractors, Map2 presence for V cues was longer (50ms) than for AV cues (18ms), $t_{(38)} = 4.7$, $p < 0.001$. Yet, for TCC distractors, all map durations were comparable between V and AV cues (all p's > 0.1). Next, when following up as a function of Cue Modality, for AV cues, Map4 duration for TCC distractors was longer (67ms) than for NCC distractors (40ms), $t_{(38)} = 3.8$, $p = 0.004$. Likewise, for V cues, Map4 duration was longer for TCC (67ms) than NCC (39ms) distractors, $t_{(38)} = 3.6$, $p = 0.003$. However, Map2 was also longer for NCC (50ms) than TCC distractors (10ms), for V cues $t_{(38)} = 5.4$, $p < 0.001$. Thus, maps that are sensitive to TAC and MSE appear to interact, suggesting that top-down visual attentional control and bottom-up multisensory attentional control may share neural generators.

To explore if and when the above adult topographical EEG patterns are present in children, we submitted each child age-groups' data within the 180-300ms time-window to a

fitting procedure, where child topographical data were labelled according to the adult template maps with they which they best correlated spatially.

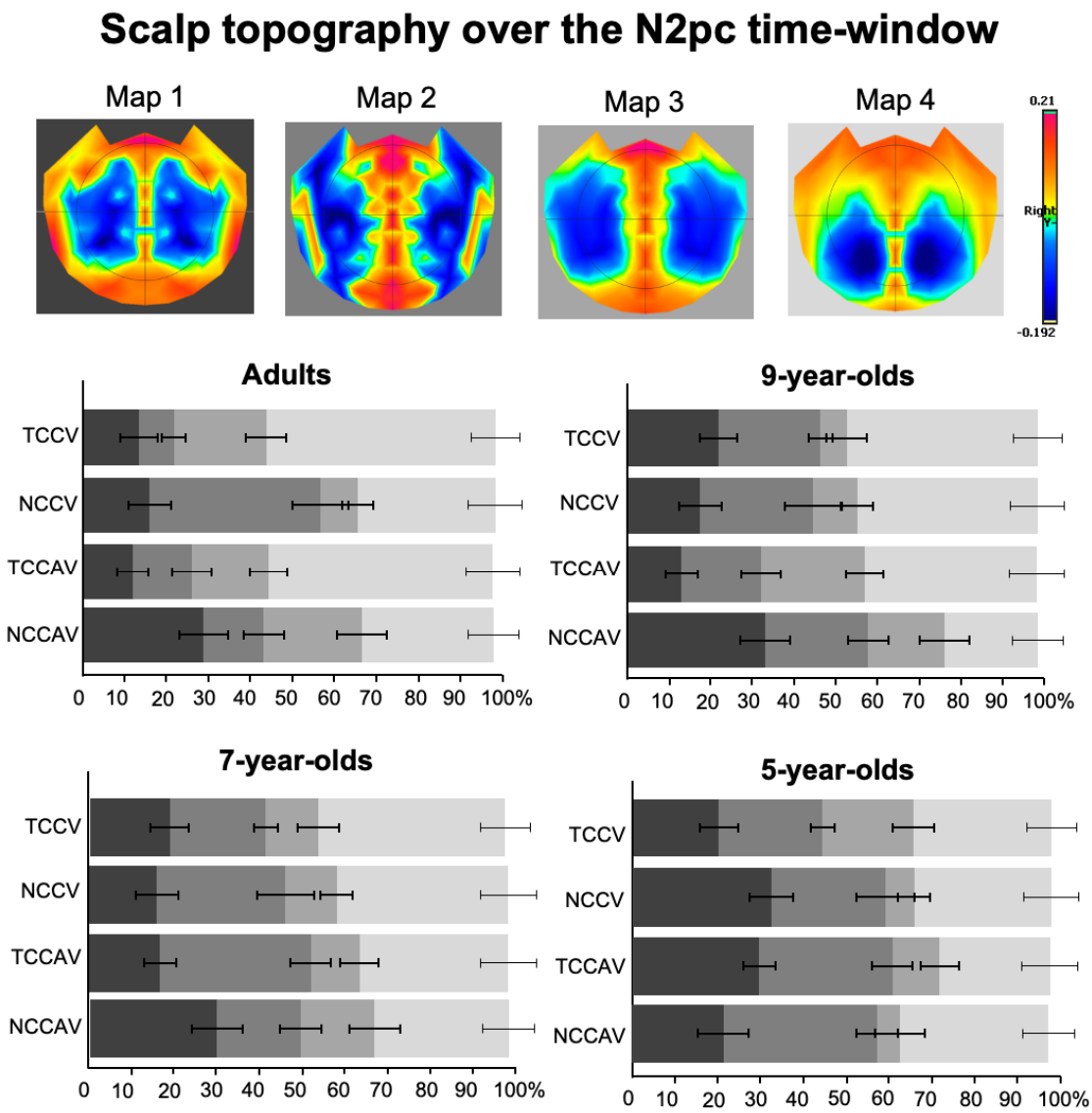

Figure 5. Scalp topography of the 4 lateralised difference template maps elicited over the N2pc time-window as a function of cue condition and observer age group. The four template maps resulting from the segmentation of the adult lateralised 'mirrored' difference ERP data are shown in the upper row. The bar graphs below represent each difference template map's relative duration (% ms) over the N2pc time window, shown separately for the adults and the 3 younger groups, and for each of the V and AV cue conditions separately. Bars in the graphs are coloured according to their map's backgrounds in the top row, and error bars denote the standard error of the mean. As visible in the lower graphs, Map 4 was the most dominant in adults, 9-year-olds, and 7-year-olds, while 5-year-olds did not have a clear map dominance pattern. Only in adults' duration of Map 4 was modulated by cue type that is whether cue colour matched that of the target-colour.

For 9-year-olds, the ANOVA revealed a main effect of Map, $F_{(3, 75)} = 9.2$, $p < 0.001$, $\eta_p^2 = 0.3$, and, like in adults, Map4 predominately characterised ERPs during the N2pc time-window (Figure 5, middle right panel). Map presence was modulated only by Cue Modality, as evidenced by a 2-way interaction between Map and Cue Modality, $F_{(3, 75)} = 3.4$, $p = 0.04$, $\eta_p^2 = 0.1$. A follow up by Cue Modality found that Map3 was longer for AV (27ms) than V cues (11ms), $t_{(25)} = 2.6$, $p = 0.02$, while Map4 was longer for V (55ms) than AV cues (39ms), $t_{(25)} = 2.5$, $p = 0.02$. However, the map that was sensitive to the (audio)visual nature of the cues in adults, Map2, was comparably present for V cues (31ms), and AV cues (27ms), $t_{(25)} = 0.7$, $p = 1$. In a follow-up as a function of Map, there were no significant differences between map durations for AV cues (all p's > 0.1). For V cues, however, Map4 (55ms) was longer than all other maps (Map1 [24ms], $t_{(25)} = 4$, Map2 [32ms], $t_{(25)} = 3.7$, Map3 [11ms], $t_{(25)} = 5.8$, all p's < 0.001). In a marked contrast to adults, 9-year-olds did not show the other 2-way interaction of interest, Map x Cue Colour ($F_{(3, 75)} = 1.3$, $p = 0.3$). Other interactions failed to reach statistical significance (all F's < 2, p's > 0.1). Taken together, 9-year-olds seemed to show *adult*-like MSE (a Map x Cue Modality 2-way interaction). Even though they did not show a modulation of the adult MSE-sensitive map, 9-year-olds' overall topographic pattern was like that of adults, with a predominance of Map4 across conditions.

In 7-year-olds, there was also a main effect of Map, $F_{(2.3, 85.5)} = 9.7$, $p < 0.001$, $\eta_p^2 = 0.2$, with a predominance of Map4, akin to the two older age groups (Figure 5, bottom left panel). Unlike in older age groups, however, no other main effects or interactions reached statistical significance (all F's < 3, p's > 0.1). This included the 2-way interactions of interest, Map x Cue Colour ($F_{(3, 111)} = 0.7$, $p = 0.6$) and Map x Cue Modality ($F_{(2.4, 87.3)} = 1.3$, $p = 0.3$). We can therefore conclude that 7-year-olds' topography did not show adult-like TAC or MSE, although their overall topographic pattern could be considered adult-like.

Finally, 5-year-olds also showed a main effect of Map, $F_{(3, 81)} = 6.3$, $p < 0.001$, $\eta_p^2 = 0.2$, but here, there was no clear map dominance pattern (Figure 5, bottom right panel). No other main effects or interactions reached statistical significance (All F's < 1), including the two 2-way interactions of interest, Map x Cue Colour ($F_{(2.1, 57)} = 0.8$, $p = 0.4$) and Map x Cue

Modality ($F_{(2.3, 61.6)}$ = 0.7, p = 0.5). With this, 5-year-olds seemed not to show adult-like TAC, MSE, or overall pattern of map presence.

Discussion

Learning environments such as classrooms are cluttered and necessitate children to control their attention, i.e., to focus their resources on relevant information and ignore unimportant information. There is modest research on how adults and children differ with regards to the brain and cognitive mechanisms of attentional control engaged by purely visual (and less so, auditory) information. In contrast, and despite clear educational relevance, little is known about the development of neuro-cognitive mechanisms governing attention to multisensory information. Our study aimed to provide insights into this important issue by clarifying how behavioural and brain mechanisms of attentional control engaged by multisensory objects develop vis-à-vis nascent control towards visual stimuli. In so doing, we made the first step towards understanding when, in school-aged children, visual and audiovisual attentional control reach an adult-like state. Moreover, with our combination of traditional behavioural (RT spatial cueing) and ERP measures of attentional selection (the N2pc component,) with robust and sensitive multivariate electrical neuroimaging analyses, we provided novel insights on the similarities in the neurophysiological mechanisms underlying naturalistic attentional control across adults and children.

1. Developmental trajectory of visual attentional control

Behaviourally, we replicated both task-set contingent visual attention capture (TAC) and multisensory enhancement of attention capture (MSE) henceforth in adults, in a larger sample and with small changes to the paradigm with respect to the Matusz and Eimer (2011) study. Crucially, children as young as 6-7 and children aged 8-9 showed adult-like magnitudes of both facilitatory visual attentional control (as shown by large and reliable spatial cueing by target-colour cues) and inhibitory visual attentional control (as shown by null cueing effects for target nontarget-colour cues). This pattern of results held even after correcting for children's overall slower processing speed. This suggested that children may reach an adult-like state of visual, feature-specific attentional control the likes of TAC

already at the age of 6-7. Behavioural evidence is converging that typical development of visual control mechanisms such as TAC may indeed plateau around this time. One other study found that the magnitude of attentional capture by nontarget colour singleton stimuli was comparable between adults and 6-year-olds (Oh-Uchi et al. 2010), although this study did not account for developmental differences in overall reaction time. More generally, the proposed plateau may hold at least for colour distraction, as shown in a replication of the original Folk et al. study with onset and colour distractors in 11-year-olds (Greenaway & Plaisted 2005). Behavioural findings were extended by EN findings which revealed, in adults, two stable patterns of brain activity that were each modulated by TAC and by MSE. Importantly, the adult lateralised EEG activity pattern, i.e., template map, that was modulated by TAC, dominated the N2pc time-window overall, and this same dominance pattern was shown both in 7- and 9-year-olds. Interestingly, the presence of the adult TAC-sensitive map was not modulated by target-colour-matching in the child groups, as evidenced by nonsignificant Map x Cue Colour interactions in 7-and 9-year-olds. However, the finding that the brain network recruitment of those child groups that showed adult-like visual attentional control in behaviour followed adult network recruitment that was modulated by target-colour-matching, at least indirectly supports that the child groups in question could deploy their top-down attention in a way that could be considered adult-like.

In our youngest group of 5-year-olds, we did not find a presence of TAC. This result contrasts with the only other study on TAC in children, by Gaspelin et al. (2015). Although we used similar data processing procedures, including RT data cleaning, their data suggested that young children exhibit a degree of TAC, albeit smaller than that of adults. This difference to our study could be accounted for by age differences (their study's mean age: 4.2 years, our study's mean age: 5 years), though, arguably, such differences should have provided evidence *in favour of* TAC in our sample. However, in our study, a fast-paced experimental task which included fully irrelevant sound stimuli also incorporated an EEG recording, increasing the total testing time. Such factors may have also increased discomfort and fatigue in our participants, in turn contributing to null TAC effects. Nevertheless, we provided novel evidence with respect to visual attentional control in such young children. The young children effectively utilised the colour-change distractors to orient their spatial attention among other coloured shapes, leading to reliable attentional capture effects. These effects were found despite the large variability in this age group's RTs, which in turn

may have prevented a reliable group-wide TAC from emerging. These findings extend those of Gaspelin et al. (2015) where cues were colour singletons, likely facilitating attention capture (see Johnson & Tucker, 1996; Markant & Amso, 2016 for cueing effects in infants). In further support, our EN analyses revealed that 5-year-olds show stable spatially selective (and as such, indicative of attentional selection in space) patterns of EEG activity that were observed in adults. This result is novel and important as it supports the idea that behavioural data of young children's nascent top-down visual control are instantiated through similar neuro-cognitive mechanisms as those of adults.

We hasten to add that 5-year-olds in our study were relatively familiar with the school context. At this age, Swiss children learn learn how to interact appropriately with peers and teachers and receive training in foundational skills such as phonics and numerical awareness (CIIP, 2012). Thus, by age 6-7, children will have been in formal education for two years, and there is evidence to show that even one year of schooling experience can augment the involvement of attentional control networks in response to the same stimuli (Brod et al., 2017). It is thus tempting to interpret our results as schooling experience having had an extensive training on children's attentional control. However, the present study design could not explicitly disentangle the effects of schooling experience from experience-independent cognitive and/or motor development. Indeed, the variability in the children's behavioural and EEG data, and especially that of the youngest children, could have partly been driven by differences in processing speed or motor function development across children in this group, as children were tested across the span of the whole school year. Such skills develop rapidly in early childhood, and thus the differences between children at the beginning and end of the first grade could have been quite large. The involvement of motor function development could be clarified by comparing the behavioural data across 5-year-olds tested earlier versus later in the year, to verify if TAC would be stronger in the latter than the former group. Likewise, a comparison of EEG topographic patterns in 5-year-olds tested earlier versus later in the year could verify if the topographies of later-tested-5-year-olds are similar to those of 7- and 9-year-olds. However, disentangling all of the potential sources of variance as exhaustively as possible would require large-scale comparisons across children before and after school entry, and between school systems, which was outside the scope of the present study.

2. Development of attentional control processed engaged by multisensory stimuli

In contrast with visually-elicited attentional control effects, we found no behavioural evidence for MSE in any of the younger groups. This result was somewhat surprising given the protracted development of frontoparietal areas that would render children both weaker at top-down attentional control skills and more susceptible to distraction (e.g., Bunge et al., 2002; Casey et al., 2005; Konrad et al., 2005; Hwang et al., 2011). Perhaps the visual attentional demands of the experimental task were so high that they eliminated attentional capture by perceptually salient distractors. In that case, however, we would expect null MSE in adult as well, and by contrast, our adult data show a robust MSE despite changes to the original paradigm and experimental setup (like the addition of EEG-recording-related demands). This supports the idea that multisensory stimuli may make up a particularly salient category of distractors, as shown by our own results across different tasks (Matusz & Eimer 2011; Matusz et al. 2015; 2019a). Furthermore, there is converging evidence that multisensory integration can occur at stages preceding those affected by top-down processes, be it visual attentional control or even consciousness (Giard & Peronnet 1999; Cappe et al. 2010; reviewed in De Meo et al. 2015; Murray et al. 2016). How would we, then, explain the group-wise absence of MSE in our younger groups? First, our developmental null MSE results are consistent with other studies on development of multisensory processing. Since the oldest children in the current study were aged 8-9, it is possible that MSE was not detected, as multisensory simultaneity detection purportedly matures only after this age (e.g., Gori et al., 2008; 2012; Barutchu et al. 2009). However, it is important to note that the current research did not study multisensory integration per se, but rather crossmodal interactions between vision and audition, and there is evidence that such interactions are present from early ages, and even at the age of 5 (e.g. Bahrick, 2001, Broadbent et al., 2018). Therefore, it is improbable that undeveloped multisensory integration was the root of our results. Perhaps the variability of MSE was too high in children, as such variability also nullified multisensory interference effects in the highly demanding task set in Matusz et al. (2015). Alternatively, behavioural analyses may not have been sensitive enough to detect MSE in children. The null effect of behavioural MSE in children becomes less surprising if we consider that MSE was a much smaller effect than TAC even in the original adult cohort where the effect was first noted (Matusz & Eimer, 2011, Experiment 2), and in the present adult control cohort. In support of this possibility, our EN analyses demonstrated that

children's brains were indeed responsive to the multisensory nature of distracting information.

EN topographic analyses revealed that, at least in the oldest group (9-year-olds), stable brain networks that were identified in the adult lateralised EEG over the traditional N2pc time-window were statistically modulated by the multisensory nature of the distractors. This finding suggest that multisensory processes permeate goal-directed behaviour even when the eliciting stimulus is almost completely irrelevant to the task, from 9 years onwards. Our results are the first to demonstrate that multisensory distraction activates spatially-selective brain mechanisms in children as young as 8-9, with this process occurring even in the context of strong, unisensory and feature-specific visual control. These results, especially if replicated in a larger sample and across both EEG and behavioural measures, would have important applied implications. For one, such findings would suggest that more effort needs to be put into creating classrooms where the risks of multi-sensory distraction is minimised (for detrimental effects of unisensory noise on classroom learning see: Fisher et al., 2014; Massonnié et al., 2019), although children younger than 7 may be partly protected from it (Matusz et al. 2015, 2019a). Second, and perhaps most importantly, such results would highlight the immense potential of largely involuntary and attention-demand-independent facilitatory effects of multisensory processing on allocating attention to objects and symbols and, therefore, on encoding into memory (as known from the visual literature, e.g., Astle & Scerif 2011; Shimi & Scerif, 2017).

3. The N2pc as a marker of developing real-world attentional control

In adults, canonically analysed EEG showed TAC, where mean N2pc amplitudes were larger for target-colour than for nontarget-colour distractors, replicating patterns observed elsewhere in the visual attention literature (e.g., Eimer et al. 2009). However, the N2pc did not mirror the MSE that we found in adults in behaviour. In the youngest children, the sound seemed to have a disruptive effect overall, as shown by no speeding effect on RTs, but also somewhat suppressed (Pd-like; Sawaki & Luck, 2010) contralateral EEG responses. Notably, the only other comparable study (van der Burg et al. 2011) similarly showed a weak effect of 'multisensoriness' of distractors compared to a much larger, enhancing effect of visual targets on ERPs. As such, this is one line of evidence for limited validity of the N2pc in testing

attentional control in naturalistic settings, involving stimuli differing in their multi-sensory nature or task-relevance.

Crucially, no canonical adult-like N2pc was found in response to visual or audiovisual distractors in any of the 3 age groups, in contrast with extant visual developmental studies. There, a delayed but significant N2pc has been reported as early as age 9 (Couperus & Quirk, 2015; Shimi et al., 2015; see Sun et al. 2018 for N2pc in 9-15-year-olds). Albeit children's N2pc differed in amplitude and/or onset latency from adult N2pc's, suggesting some degree of development. It is possible that our paradigm did not create optimal conditions to test for developing N2pc's. First, the targets in our task appeared after 200ms of cue onset, and thus we had little leeway in allowing child data to determine the optimal latency for investigation. Thus, it is probable that the strong target-related perceptual (and later) components effectively overwrote any distractor N2pc's that might have occurred later than approximately 180-300ms. Furthermore, we recorded N2pcs to distractors, rather than targets like the above studies. There is evidence that N2pc may not be optimal to detect distractor processing even in adults (Eimer et al., 2009). This, coupled with a wealth of evidence for the viability of the N2pc as a marker of target-based attentional selection that is sensitive to healthy development (Wiegand et al. 2017; Shimi et al., 2015), aging (Pagano et al., 2015; Wiegand et al., 2013), as well as mental disorders (Fuggetta et al., 2015; Wang et al., 2016), prevents concluding that the N2pc is suboptimal to detect real-world attentional control. Nonetheless, had our analyses stopped at a canonical N2pc approach, one could have concluded that attentional control processes like TAC and MSE are simply not elicited by goal-irrelevant stimuli (cf. targets) in children. However, with the use of EN, we revealed that adult-like lateralised attentional control processes over the N2pc time-window were present at age 7 onwards, corroborating our behavioural results. To our knowledge, this is the youngest age group in which spatially-selective N2pc-like brain mechanisms have been reported. Similarly, there were stable N2pc-like lateralised EEG patterns that were modulated by MSE, as early as 8-9 years, with this effect otherwise not visible in the behavioural data.

In summary, our findings challenge the idea of the canonical N2pc as a viable marker of attentional selection of multisensory distractors in adults, or of distractors in general in children. However, when taking into account activity across the whole brain occurring over the time-period in which the N2pc is elicited, neurophysiological markers of attentional

control processes can be identified across different populations and type of stimuli. We argue that these brain mechanisms transpire within the lateralised voltage gradients, whose single points across contra and ipsilateral hemifields are indeed captured by canonical electrode-pair analyses. With this sensitivity of EN methods, our results showcase how lateralised ERP patterns reflect 1) the previously elusive sensitivity of visual attentional selection to bottom-up multisensory processing, and 2) attentional control processes in children as young as 7. We are currently testing the extent to which these spatially-selective patterns are robust across other types of multisensory processes, especially those based on semantic relationships or temporal expectations (Sarmiento et al. 2016; Matusz et al. 2016; see review in ten Oever et al. 2016). In yet unpublished data, we found evidence in support of the important role of these top-down processes in modulating adult multisensory distraction across both the brain and behaviour.

4. Conclusion

Taken together, our study revealed the developmental trajectory of a frequently studied visual attentional control mechanism that is task-set contingent attentional capture (TAC). We show, both behaviourally and using an EN analytical framework, that TAC develops early in childhood, likely after the age of 5, and reaches adult-like state already at age 7. Though MSE, present in adults, was undetected in children's behaviour or traditionally-analysed EEG signals, an EN framework again was crucial, as it revealed spatially-selective brain mechanisms sensitive to the multisensory nature of distracting stimuli in children as young as 8-9. Our findings underline the utility of combining traditional behavioural and EEG/ERP markers of visual attentional control with advanced EEG analytical techniques for investigating the development of attentional control and identifying developmental differences attentional control between adults and children, but also in revealing early similarities.

Acknowledgments

We thank Noémie Kirscher for assistance with child IQ data collection and Louise Vasa for assistance with child EEG data analysis. Funding: This work was supported by the Pierre Mercier Foundation, the Swiss National Science Foundation (grant PZ00P1_174150), and the Fondation Asile des Aveugles.

References

Astle, D. E., & Scerif, G. (2011). Interactions between attention and visual short-term memory (VSTM): What can be learnt from individual and developmental differences?. *Neuropsychologia*, *49*(6), 1435-1445.

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

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.

Broadbent, H. J., Osborne, T., Mareschal, D., & Kirkham, N. Z. (2019). Withstanding the test of time: Multisensory cues improve the delayed retention of incidental learning. Developmental science, *22(1),* e12726.

Broadbent, H. J., White, H., Mareschal, D., & Kirkham, N. Z. (2018). Incidental learning in a multisensory environment across childhood. *Developmental science*, *21*(2), e12554.

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.

Buchan, J. N., & Munhall, K. G. (2012). The effect of a concurrent working memory task and temporal offsets on the integration of auditory and visual speech information. *Seeing and Perceiving*, *25*(1), 87.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron*, *33*(2), 301-311.

Cappe, C., Thut, G., Romei, V., & Murray, M. M. (2010). Auditory–visual multisensory interactions in humans: timing, topography, directionality, and sources. *Journal of Neuroscience*, 30(38), 12572-12580.

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.

Conférence intercantonale de l'instruction publique de la Suisse romande et du Tessin (CIIP). Plan d'études romand (PER), Aperçu des contenus (June 2012). Available at: https://www.plandetudes.ch/documents/10136/19192/cycle_1_webCIIP.pdf

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

De Meo, R., Murray, M. M., Clarke, S., & Matusz, P. J. (2015). Top-down control and early multisensory processes: chicken vs. egg. *Frontiers in integrative neuroscience*, *9*, 17.

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. (2014). The neural basis of attentional control in visual search. *Trends in Cognitive Sciences*, *18*(10), 526-535.

Eimer, M., & Kiss, M. (2008). Involuntary attentional capture is determined by task set: Evidence from event-related brain potentials. *Journal of cognitive neuroscience*, *20*(8), 1423-1433.

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.

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.

Fuggetta, G., Bennett, M. A., & Duke, P. A. (2015). An electrophysiological insight into visual attention mechanisms underlying schizotypy. *Biological psychology*, *109*, 206-221.

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.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of cognitive neuroscience*, *11*(5), 473-490.

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.

Hommel, B., Li, K. Z., & Li, S. C. (2004). Visual search across the life span. *Developmental psychology*, *40*(4), 545.

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.

Kiss, M., Jolicœur, P., Dell'Acqua, R., & Eimer, M. (2008b). Attentional capture by visual singletons is mediated by top-down task set: New evidence from the N2pc component. *Psychophysiology*, *45*(6), 1013-1024.

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

Konrad, K., Neufang, S., Thiel, C. M., Specht, K., Hanisch, C., Fan, J., ... & Fink, G. R. (2005). Development of attentional networks: an fMRI study with children and adults. *Neuroimage*, *28*(2), 429-439.

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. (2014). Early experience and multisensory perceptual narrowing. *Developmental psychobiology*, *56*(2), 292-315.

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., Retsa, C., & Murray, M. M. (2016). The context-contingent nature of cross-modal activations of the visual cortex. *Neuroimage*, *125*, 996-1004.

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.

Maylor, E. A., & Lavie, N. (1998). The influence of perceptual load on age differences in selective attention. *Psychology and aging*, *13*(4), 563.

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.

Michel, C. M., & Koenig, T. (2018). EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. *Neuroimage*, *180*, 577-593.

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.

Murray, M.M. and Wallace, M.T., eds (2012) *The Neural Bases of Multisensory Processes,* CRC Press

Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental science*, *9*(5), 454-464.

Oh-uchi, A., Kawahara, J. I., & Sugano, L. (2010). Attentional capture and metaattentional judgment: a study of young children, parents, and university students. *Psychologia*, *53*(2), 114-124.

Pagano, S., Fait, E., Monti, A., Brignani, D., & Mazza, V. (2015). Electrophysiological correlates of subitizing in healthy aging. *PloS one*, *10*(6).

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.

Sarmiento, B. R., Matusz, P. J., Sanabria, D., & Murray, M. M. (2016). Contextual factors multiplex to control multisensory processes. *Human brain mapping*, *37*(1), 273-288.

Sawaki, R., & Luck, S. J. (2010). Capture versus suppression of attention by salient singletons: Electrophysiological evidence for an automatic attend-to-me signal. *Attention, Perception, & Psychophysics*, *72*(6), 1455-1470.

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

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2011). Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biological psychology*, *87*(1), 146-151.

Shimi, A., & Scerif, G. (2017). Towards an integrative model of visual short-term memory maintenance: Evidence from the effects of attentional control, load, decay, and their interactions in childhood. *Cognition*, *169*, 61-83.

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.

Shimi, A., Nobre, A. C., Astle, D., & Scerif, G. (2014a). Orienting attention within visual short-term memory: Development and mechanisms. *Child development, 85(2),* 578-592.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. The MIT 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.

Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in cognitive sciences*, *14*(9), 400-410.

Ten Oever, S., Romei, V., van Atteveldt, N., Soto-Faraco, S., Murray, M. M., & Matusz, P. J. (2016). The COGs (context, object, and goals) in multisensory processing. *Experimental brain research*, *234*(5), 1307-1323.

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.

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

Tsujimoto, S. (2008). The prefrontal cortex: Functional neural development during early childhood. *The Neuroscientist*, *14*(4), 345-358.

Van der Burg, E., Talsma, D., Olivers, C. N., Hickey, C., & Theeuwes, J. (2011). Early multisensory interactions affect the competition among multiple visual objects. *Neuroimage*, *55*(3), 1208-1218.

Wang, E., Sun, L., Sun, M., Huang, J., Tao, Y., Zhao, X., ... & Wang, Y. (2016). Attentional selection and suppression in children with attention-deficit/hyperactivity disorder. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1(4), 372-380.

Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence* (4th ed.). Bloomington, MN: Pearson.

Wechsler, D. (2014). Wechsler intelligence scale for children-fifth edition. Bloomington, MN: Pearson.

Wiegand, I., Finke, K., Müller, H. J., & Töllner, T. (2013). Event-related potentials dissociate perceptual from response-related age effects in visual search. Neurobiology of Aging, 34(3), 973-985.

Wiegand, I., Petersen, A., Finke, K., Bundesen, C., Lansner, J., & Habekost, T. (2017). Behavioral and brain measures of phasic alerting effects on visual attention. Frontiers in human neuroscience, 11, 176.