The development of inhibitory control: an averaged and single-trial Lateralized Readiness Potential study

Donna Bryce, Dénes Szűcs, Fruzsina Soltész, David Whitebread

Bryce, Donna, Dénes Szűcs, Fruzsina Soltész, and David Whitebread. 2011. “The development of inhibitory control: an averaged and single-trial Lateralized Readiness Potential study.” *NeuroImage* 57 (3): 671–85. https://doi.org/10.1016/j.neuroimage.2010.12.006.

The development of inhibitory control: An averaged and single-trial Lateralized Readiness Potential study

Donna Bryce [a,b,*], Dénes Szűcs [a,*], Fruzsina Soltész [a], David Whitebread [b]

[a] *Centre for Neuroscience in Education, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge, CB2 3EB, UK*
[b] *Faculty of Education, University of Cambridge, 184 Hills Road, CB2 8PQ, UK*

Introduction

The term 'executive functions' encompasses a range of skills used to monitor and control behavior flexibly in novel situations; these skills are important for education and child development. While executive functioning research originally emerged from research into adult clinical populations (e.g. Luria, 1973), these crucial skills have increasingly been investigated in children (Carlson, 2005; Epsy, 1997) and in an educational context (Blair and Razza, 2007; Meltzer, 2007). Further, in bridging the disciplines of neuroscience and education, executive functions may be fruitful skills to investigate as they are more specifically defined and discrete than the learning approaches they may contribute to or be precursors of, such as metacognitive skilfulness (Fernandez-Duque et al., 2000; Shimamura, 2000). Inhibitory control (IC), the ability to suppress thoughts or responses when they are inappropriate or no longer relevant (Deak and Narasimham, 2003), is one executive function which appears to be crucial to cognitive development. Here we provide an event-related brain potential (ERP) investigation of developmental changes of IC across three age groups (5-year-olds in Year 1 of primary school, 8-year-olds in Year 3 of primary school, and adults) using an adaptation of a Stroop task. Our aim was to understand how the temporal properties of covert response preparation, as well as the increasingly successful suppression of unwanted motor tendencies, contribute to the development of IC. Besides a thorough analysis of reaction time (RT) we exploited an ERP derivative, the Lateralized Readiness Potential (LRP), which is able to track motor response preparation before a behavioral response is given. Several other ERP parameters, usually associated with stimulus processing speed, were also examined as control measures.

Inhibitory control is a major contributor to educational performance. It has been associated with mathematical and reading ability in children at all stages of schooling: preschool (Deak and Narasimham, 2003; Espy et al., 2004), kindergarten (Blair and Razza, 2007; Brock et al., 2009), the beginning of primary school (Swanson, 2006) and the end of primary school (St Clair-Thompson and Gathercole, 2006). Aside from contributing to various academic subjects, inhibitory control appears to be crucial for coping with other demands of attending school. That is, Brock et al. (2009) found that a composite measure of 'cool' executive functions (that is, those with fewer emotional and more cognitive demands) predicted both a child's behavior in class, as reported by their teacher, and their engagement in learning, as observed by a researcher. IC has been associated with features of temperament important for formal school settings, such as activity level and impulsivity (Gonzalez et al., 2001) and social

---

* Corresponding authors. Centre for Neuroscience in Education, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge, CB2 3EB, UK. Fax: +44 1223 767602.

E-mail addresses: dlb36@cam.ac.uk (D. Bryce), ds377@cam.ac.uk (D. Szűcs).

emotional competence (Rhoades et al., 2009). IC has also been linked to other milestones in cognitive development that are important for educational gains, such as working memory which is required when manipulating and maintaining information simultaneously (Alloway et al., 2005). Both Roberts and Pennington (1996) and Wilson and Kipp (1998) have proposed that IC and working memory are directly related, albeit in slightly different ways, and there is some debate regarding how this relationship changes with age (Davidson et al., 2006; Roncadin et al., 2007; Tsujimoto et al., 2007). Further, IC skills, in particular as relating to conflict resolution and not delay, are superior in bilingual children (Bialystok and Martin, 2004; Carlson and Meltzoff, 2008; Martin-Rhee and Bialystok, 2008), and aging bilingual adults (Bialystok et al., 2004) as compared to their monolingual peers, suggesting this is a malleable skill which is somewhat dependent on environment.

As yet, the majority of research into inhibitory control in young children has focused on developing age-appropriate tasks and determining how the skill develops. IC is normally assessed in the laboratory using tasks such as the Stroop (Stroop, 1935), Eriksen flanker (Eriksen and Eriksen, 1974), Go/NoGo (e.g. Pfefferbaum et al., 1985), and Stop-Signal (Logan, 1994) tasks; in more real-to-life settings, tasks such as delay of gratification (Mischel et al., 1972), A-not-B (Piaget, 1954), and the bear–dragon task (Reed et al., 1984) have been used. Accordingly, depending on the specific tasks used, there is variation in findings regarding the 'age of onset'. It can be concluded fairly confidently that a basic form of IC is present even from one year of age (see Davidson et al., 2006; Garon et al., 2008 for reviews) but it continues to develop until adolescence (Stop-Signal and Eriksen flanker tasks) or adulthood (Stroop task, Huizinga et al., 2006). Many researchers consider the first few years of formal schooling to be a period of intense growth in IC and related skills (Hughes et al., 2010), making this an interesting stage to examine.

Typically, developmental studies rely on accuracy or raw RT data to draw conclusions, an approach which has recently been challenged (Span et al., 2004) as some researchers highlight the need for speed of processing to be taken into account when comparing participants of different ages. Different statistical methods have been adopted to achieve this, including the use of standardized RT scores (Christ et al., 2001), proportional transformations of RTs (Christ et al., 2001; Johnson et al., 2003), and logarithmic transformations of RTs. Faust et al. (1999) posit that standardized RTs are superior to both log and proportional transformed RTs in terms of controlling for differences in processing speeds. However, it may not always be appropriate to transform raw RTs into standard scores, depending on the shape of the distribution. For instance, when standardized RTs are used, it is assumed that the raw RTs are distributed normally, when in fact often they are skewed (Jensen, 1992). A solution to the above issue is provided by fitting an exponential-Gaussian (ex-Gaussian) distribution to RTs (Leth-Steensen et al., 2000; McAuley et al., 2006; Mewhort et al., 1992; Myerson et al., 2007). This approach produces three parameter values: mu (μ), which reflects mean performance, sigma (σ), which reflects variance in performance, and tau (τ), which reflects the skewness of the upper tail of the distribution. Ex-Gaussian analysis was recently applied to a classic Stroop task with children aged 10–12 years and young adults (Fagot et al., 2009). The authors concluded that 'whereas for young adults, interference is reflected in a slowing of responses (μ) and lapses of attention (τ) in the incongruent condition, interference in children seems to be entirely accounted for by a slowing of responses (μ)' (Fagot et al., 2009). However, there are two caveats regarding this interpretation. First, the participants were not all performing the task in their native language which would impact the level of interference the color–words asserted on the participants. Second, the classic Stroop task is not ideal for comparing IC in children and adults, as word reading is more automatic in adults than in children. Therefore, it is unclear whether the inhibitory demands were equivalent for each group. A simplified version of the Stroop which is not affected by reading experience would be better able to address developmental effects. For this reason, in our research we have used such a task, and adopted a range of RT transformations in an attempt to control for differences in processing speed.

An important issue concerning RT and accuracy measures is that they represent the culmination of many cognitive processes. In contrast, exploiting its virtually unlimited temporal resolution, electroencephalography (EEG) provides a detailed insight into the cognitive processes that occur before behavioral responses are given. With regard to IC, most adult EEG studies have focused on understanding the functional significance of the centro-parietal P3 (henceforth referred to as P3b, as defined by Dien et al., 2004) and fronto-central N2 components. A series of studies, using different tasks, showed that greater P3b peak amplitudes and latencies were elicited for conditions requiring IC (continuous performance task: Davis et al., 2003; Stop-Signal task: Dimoska et al., 2006; Eriksen flanker task: Ridderinkhof and van der Molen, 1995). Subsequently, two studies that used cueing Go/NoGo paradigms in adults (Bruin et al., 2001; Smith et al., 2007) found that the P3b amplitude was larger in 'NoGo' as opposed to 'Go' conditions, and for incorrectly as compared to correctly cued trials (i.e. when the participant was cued to prepare the incorrect response and therefore the level of IC required was increased). Ramautar et al. (2004) found a similar effect using the Stop-Signal paradigm with adults. These findings suggest that the P3b component may be directly or indirectly sensitive to IC, as altering the IC demands of a task results in changes in features of the P3b. The N2 is usually observed in conjunction with the P3b, and dissociating their functional significance has been the focus of a number of studies. For instance, while Bruin et al. (2001) and Smith et al. (2007) found that the P3b was affected by the validity of cues, both also found the N2 was not. Similarly, Donkers and Van Boxtel (2004) introduced a third condition (called GO) of the classic Go/NoGo task, whereby in the 'GO' condition the adult participants were required to respond with maximum force (so there was conflict, but no response inhibition). The N2 was elicited in both NoGo and GO trials, and N2 amplitudes were greater if these trials were less frequent. These, and many other studies, suggest that the N2 reflects conflict monitoring, and not response inhibition.

Developmental EEG studies that have tested the N2 and P3b components in children in order to assess IC report that the peak latencies and amplitudes of the N2/P3b are reduced in older relative to younger children (continuous performance task: Davis et al., 2003; Go/NoGo task: Lamm et al., 2006; Eriksen flanker task: Ridderinkhof and van der Molen, 1995). Developmental fMRI studies have largely attributed the development of IC to the maturation of the prefrontal cortex (Bunge and Wright, 2007; Durston and Casey, 2006; Marsh et al., 2006; van der Molen, 2000). For example, Adleman et al. (2002) found that in a Stroop task, parietal lobe activation reaches adult levels by adolescence, but prefrontal cortex activation continues to develop in this period. Rubia et al. (2007) found similar results with the Stop-Signal task — that prefrontal cortex activation increases with development. Schroeter et al. (2004) also found increased dorso-lateral prefrontal cortex activation with age in a functional near-infrared imaging (fNIR) study using a Stroop task, and participants aged 7–29 years. These results have been interpreted as indicating that the prefrontal cortex is crucial for IC, and that children's IC is limited by this brain region being late maturing. However, a limitation of many fMRI and P3b and N2 studies is that the functional interpretation of findings is not clear.

It has long been theorized that there are at least two processes that contribute to IC. These can generally be grouped as 'early' (referred to as interference suppression or stimulus interference control) and 'late' (referred to as response inhibition or response interference control) processes. Hereafter we refer to these early and late processes as interference suppression and response inhibition, respectively. Interference suppression is defined as 'the ability to filter out irrelevant information in the environment' (Bunge et al., 2002) and can be considered an 'early' stage process because when interference is successfully suppressed initially, less response inhibition is required at a later stage. A similar phenomenon is hypothesized to contribute specifically to performance on Stroop tasks, referred to as stimulus interference control (Jongen and Jonkman, 2008). This refers to the ability to suppress interference present in the stimulus (i.e. the printed word) at an early stage of inhibitory control. Two analogous concepts have also been proposed to exert an effect at 'later' stages of inhibitory control. These are response inhibition (Bunge et al., 2002) and response interference control (Jongen and Jonkman, 2008). Both these phenomena focus on inhibiting prepotent or automatic motor responses that are elicited by the stimuli. It is important to note that an additional 'early' process that may contribute to IC abilities is task preparation (Brass and Yves von Cramon, 2002; Chikazoe et al., 2009; Steinhauser et al., 2009). This refers to the phenomenon that responses are faster when cues regarding the nature of the upcoming task are provided, and it has been shown that greater preparation leads to more efficient inhibition in an adapted Go/NoGo task (Chikazoe et al., 2009).

Some evidence is accumulating for the separation of the two processes that contribute to IC (interference suppression and response inhibition). For instance, Bunge et al. (2002) used fMRI with children (aged 8–12-years) and adults (aged 19–33-years) and manipulated the Eriksen flanker task to establish that different brain regions are used for interference suppression and response inhibition. It was found that during interference suppression trials children showed prefrontal activity in the opposite hemisphere to adults, and when response inhibition was being used children activated only posterior regions, while adults activated both posterior and prefrontal regions. This study suggests that interference suppression and response inhibition are probably dissociable components of IC. More specifically related to the Stroop task, the finding that IC is composed of at least two processes is supported by Jongen and Jonkman (2008), who found that stimulus interference control is mature by 6–7 years of age, but response interference control continues to develop until 12 years of age. Similarly, de Houwer (2003) and Melcher and Gruber (2009) have identified different conflict factors within Stroop tasks, using different conditions. As far as we are aware, separating the processes that contribute to IC has not yet been achieved within one condition, as was the aim in our study. Therefore, in this study, we have compared the N2 and P3b components across three developmental stages but have focused our investigation on an EEG phenomenon that provides a different view of the development of IC, the Lateralized Readiness Potential (LRP).

The LRP is able to continuously track motor cortex activation, indicating selective motor preparation and response initiation before an overt response is given (de Jong et al., 1988; Gratton et al., 1988). The LRP was first used to study development in the pioneering research of Ridderinkhof and van der Molen (1995), which compared the onset latencies of correct response preparations as measured by the LRP in children aged 5–12 years and adults, produced by the Eriksen flanker task. It was found that responses were prepared faster with age, while another indicator of perceptual processing (the P3b peak latency) did not interact with age. They concluded that 'the current data favour an interpretation that posits the locus of developmental change in the ability to resist interference at the motor end of the reaction process' (page 1053; Ridderinkhof and van der Molen, 1995). However, this study only observed correct response preparation, and used only one measure, LRP onset latency.

We also used the LRP in our previous studies, and demonstrated that it was able to detect the activation of incorrect response tendencies in correctly responded trials in young children in Stroop tasks (Szucs et al., 2009a; Szucs et al., 2007). This approach was based on classical studies which showed that in correctly responded incongruent trials, the LRP can detect latent response preparation associated with incorrect responses in Stroop and flanker tasks (Gratton et al., 1988; Smid et al., 1990). These data have been interpreted as demonstrating that both task-relevant and task-irrelevant stimulus features undergo parallel processing, and motor processes are prepared in parallel. In the incongruent condition this can result in latent incorrect motor cortex activation even in the absence of overt incorrect response activity. These incorrect response tendencies are then corrected and participants ultimately give a correct response. The LRP is able to track both the initial incorrect response tendencies and the transition from incorrect to correct response preparation. Hence, numerous parameters of the LRP can serve as effective measures of IC because they provide an insight into the dynamics of organizing overt and suppressed response tendencies. These parameters are summarized in Fig. 1.

Here we extend our previous study by increasing the complexity of experimental design, using more dependent measures of the LRP (see Fig. 1), and by broadening the sample to include 5-year-olds (in Year 1 of primary school), 8-year-olds (in Year 3 of primary school) and adults so that the continuity of developmental effects can be examined. We focused on early childhood because the early school years are particularly interesting in the development of IC. Participants completed an animal Stroop task which required them to select which animal is larger in real-life, from two animals of differing physical sizes. The most important measure was the LRP in the incongruent condition. Based on our previous study we expected that in ultimately correctly responded incongruent trials (Fig. 1B) the LRP would initially demonstrate incorrect response activation (Fig. 1B, zone II), followed by correct response activation (Fig. 1B, zone III). Importantly, this latent incorrect response activation in the

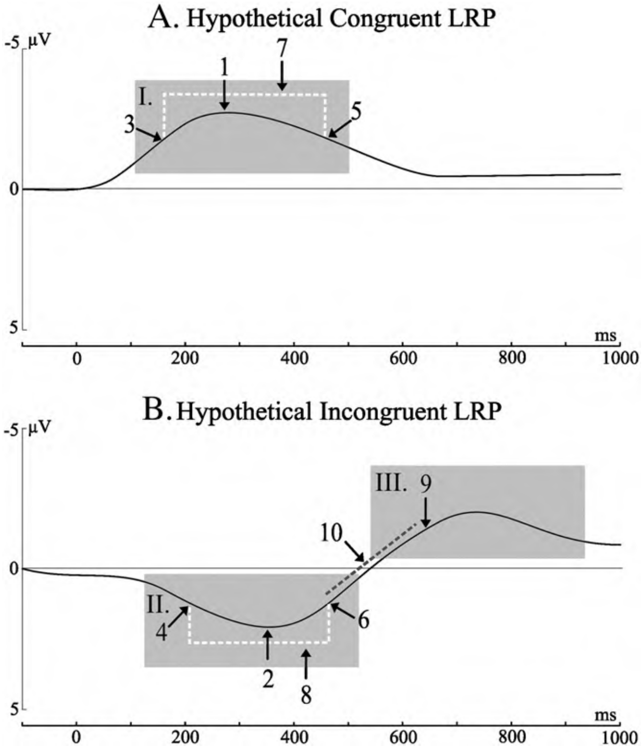

Fig. 1. Illustration of the Lateralized Readiness Potential (LRP) measures using hypothesized LRP graphs in (A) the congruent condition and (B) the incongruent condition. According to the traditional computation, a negative-going LRP indicates correct response preparation and a positive-going LRP indicates incorrect response preparation; negative is plotted upwards; stimulus was presented at time zero. The grey zones signify the expected deviations from zero, called 'response preparations' (I is the initial correct response preparation in the congruent condition, II is the initial incorrect response preparation in the incongruent condition, and III is the secondary correct response preparation in the incongruent condition). Within these grey zones, the specific measures of interest are indicated by arrows — (1, 2) initial response preparation peaks; (3, 4) initial response preparation onsets; (5, 6) initial response preparation cessations; (7, 8) initial response preparation durations; (9) onset of secondary response preparation in incongruent condition; (10) transition from incorrect to correct response preparation in incongruent condition.

incongruent condition cannot be observed by any behavioral methods. We expected that the developmental changes in the timing of incorrect and subsequent correct response activation would be highly informative regarding the development of IC. More specifically, we expected that the LRP would be able to detect stronger incorrect response activation, and the transition from incorrect to correct response activation would take longer (Fig. 1B, arrow 10), in younger than older participants. Additionally, how each age group processed the initial conflict could also prove informative, so we compared the onsets (Fig. 1, arrows 3 and 4) and durations (Fig. 1, arrows 7 and 8) of initial response preparation in each congruency condition (correct preparation in congruent, and incorrect preparation in incongruent trials). While the LRP has been examined in children previously, the functional significance of the phenomenon is yet to be systematically checked across development. For this reason, we conducted a single-trial analysis of the LRP to address whether the LRP is functionally similar in children and adults; that is, whether the three age groups show the same relationships among LRP variables and RTs. Conspicuous ERP waves typically related to early visual processing (the parieto-occipital P1, P2, N2) were examined to determine the effect of speed of processing on neural activity. The amplitude and latency of the centro-parietal P3b wave was examined because some have found it to reflect IC (e.g. Bruin et al., 2001; Ramautar et al., 2004; Smith et al., 2007), while others use it as a measure of stimulus processing speed (e.g. Duncan-Johnson and Kopell, 1981; Ilan and Polich, 1999; Kutas et al., 1977). ERPs were examined to consider the effects of real-life size difference of the animals, which served as an indicator of perceptual processing speed and attentional selection of the task-relevant stimulus dimension. Besides ERP measures, we applied several transformations to RT data in order to account for changes in speed of processing and consider qualitative differences in how RTs were distributed.

Materials and methods

Participants

Three age groups of participants took part in the experiment, and after artifact rejection 16 participants remained in each group. Initially, twenty-nine of the youngest group (5-year-olds in Year 1 of primary school, recruited from four schools in the local area) and 32 children in Year 3 of primary school (8-year-olds, recruited from five schools in the local area) visited the laboratory. Sixteen adults attended the laboratory and all participants' data were used. All analysis occurred after artifact rejection was complete in all participants. The mean age of the 5-year-old group was 5:09 (years: months; nine female; range of ages 5:06–6:07), the mean age of the 8-year-old group was 7:11 (four female; 7:07–8:05), and the mean age of the adult group was 27:00 (eight female; 23:02–31:07). Six 5-year-olds, four 8-year-olds and 7 adults spoke more than one language. Adults were graduate students and professionals, and received a £10 payment for taking part in the experiment. Children received a T-shirt as a token of gratitude for participation.

Stimuli and procedure

Stimuli were colored pictures of two animals, which in real-life are different sizes, presented on a screen. One animal image was physically larger than the other, and the task was to select which animal was larger in real-life. In the congruent condition (Fig. 2A), the larger in real-life animal (e.g. rhinoceros) was physically larger on the screen than the smaller in real-life animal (e.g. butterfly). In the incongruent condition (Fig. 2B), the larger in real-life animal was physically smaller on the screen than the smaller in real-life animal. The stimuli reported here are similar to our previous study (Szucs et al., 2009a), but with some important modifications. In this study, in addition to the congruency condition, there was also a real-life size difference manipulation. That is, there were three animal groupings based on their real-life size: small (e.g. butterfly), medium (e.g. sheep) and large (e.g. rhinoceros). Depending on which animals were presented in the pair, the size difference between members of a pair was either small (small-medium or medium-large animal pairings) or large (small-large animal pairings). This made it possible to measure ERP size difference effects as an indicator of 'speed of processing' and the attentional selection of the task-relevant stimulus dimension. It was also intended to prevent the participants from learning the animal pairings, as the medium-sized animals (e.g. sheep) could either be larger in real-life (if paired with a small animal) or smaller in real-life (if paired with a large animal). The size difference factor was examined in certain analyses where this issue was pertinent. Otherwise, only the congruency factor was analyzed, as this was the main focus of the study.

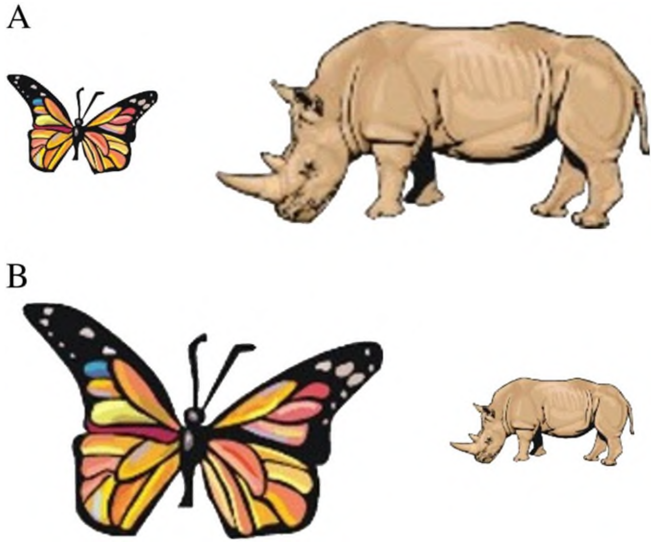

Fig. 2. The two conditions of the animal Stroop task, (A) congruent and (B) incongruent.

The experiment consisted of up to 8 blocks of 48 trials, with all participants completing at least 4 blocks. A practice block of 12 trials was completed initially, and all children also played a computer game beforehand to ensure they understood the concept of 'larger in real-life'. All children completed this preparatory game and the practice block satisfactorily. A trial began with an image of an eye presented for 500 ms. Participants were instructed to blink if necessary during this period, and refrain from blinking at other times. The eye was followed by a 1000 ms blank screen, and then the stimulus was presented for 4000 ms or until the participant gave a behavioral response. The inter-trial interval was 1000 ms. Participants gave their responses on specially designed button boxes, resting on a lap tray. Each block was balanced for congruent/incongruent trials, and left-hand/right-hand responses. In addition, the ordering of trials within a block ensured that congruent and incongruent epochs requiring either a left-hand or a right-hand response were preceded by epochs belonging to each response hand and congruency condition an equal number of times. Stimuli were presented by the Presentation software (Neurobehavioral Systems). Data were recorded in an acoustically and electrically shielded testing booth. Only correctly responded trials were included in RT and EEG analysis.

Behavioral data analysis

Accuracy and reaction time (RT) were analyzed by Congruency (congruent, incongruent) × Size Difference (large, small size difference) × Age Group (5-year-old, 8-year-old, adult) repeated measures ANOVAs. The distinctiveness of RT distributions was tested using the Kolmogorov–Smirnov (K–S) test. These comparisons were made within groups (congruent vs. incongruent conditions), and across groups (within each condition). Additionally, RT data were transformed using four methods that may provide a more in-depth insight into the data. Three methods have been used in the literature (Christ et al., 2001; Johnson et al., 2003; Pritchard and Neumann, 2009) to account for processing speed differences between age groups. First, standardized RTs were computed for each individual, according to individual mean and standard deviation values. Second, raw RTs were log transformed. Standardized and log transformed scores were entered into a Congruency (2) × Age Group (3) ANOVA. Third, mean RTs were proportionally transformed using the formula: (incongruent RT − congruent RT)/congruent RT. The proportionally transformed data were subjected to a one-way ANOVA with an Age Group factor. Kolmogorov–Smirnov tests comparing raw RT distributions to the normal Gaussian distribution were conducted (within each congruency condition and age group) in order to assess the shape of the RT distributions, as it has been suggested that the ex-Gaussian distribution is a better model of RT than the normal (Gaussian) distribution (Leth-Steensen et al., 2000; McAuley et al., 2006; Mewhort et al., 1992). Subsequently, ex-Gaussian distributions were fitted to each individual's congruent and incongruent RT data using the software package provided by Lacouture and Cousineau (2008). The ex-Gaussian distribution provides three parameters as opposed to the two parameters of the Gaussian distribution (mean and standard deviation). The three parameters are mu (μ) reflecting average performance, sigma (σ) reflecting variance, and tau (τ) reflecting the skewness of the upper tail. The three parameters were entered into Congruency (2) × Age Group (3) ANOVAs (following the procedure adopted by Fagot et al., 2009). Greenhouse–Geisser epsilon correction was not necessary in any behavioral and physiological ANOVAs (described below), as there were always fewer than three conditions. Post hoc Tukey–HSD tests were used for pair-wise comparisons.

ERP recording and pre-processing

EEG was recorded using the Electrical Geodesics system, with a 65-channel Geodesic Sensor Net for the children, and a 129-channel Geodesic Sensor Net for the adult participants. The sampling rate was 500 Hz. A 70 Hz on-line low-pass filter was used during data acquisition. During analysis, the data were band-pass filtered offline 0.03 to 30 Hz and recomputed to average reference. As mentioned previously, only correctly responded trials were included in EEG analysis. Epochs extended from −100 to 1000 ms relative to stimulus presentation, as this resulted in a satisfactory epoch retention rate after artifact rejection. Data for longer epoch lengths from the two child groups are provided for illustrative purposes in Supplementary Fig. 1. A baseline correction was applied based on activity −100 to 0 ms. Spline interpolation was carried out if necessary and the mean number of channels interpolated per participant were: three in the 5-year-olds, two in the 8-year-olds, and two in the adults. Artifact rejection was based on four criteria: the maximum gradient allowed for an epoch was 50 μV, the maximum absolute difference between the minimum and maximum voltages in an epoch was 200 μV, minimum/maximum amplitude was ± 120 μV, and the lowest activity allowed was 0.5 μV. If any of these limits were exceeded, the epoch was rejected. If more than 60% of epochs were rejected, that whole participants' data were rejected. Of those participants whose data remained in the analysis, the mean percentage of epochs accepted was 55% in the 5-year-old group (range of 41–82%), 59% in the 8-year-old group (range of 41–85%), and 79% in the adult group (range of 63–90%).

ERP analysis

LRP measures

The EEG data were subject to three types of analysis. First of all, the LRP was calculated using Coles' (1989) equation:

$$\left[(\mathrm{ER}-\mathrm{EL})_{\text{left hand response}} + (\mathrm{EL}-\mathrm{ER})_{\text{right hand response}}\right] / 2.$$

ER represents the activity from an electrode situated over the right motor cortex (usually C4 in the 10–20 electrode system), and EL represents the activity from an electrode situated over the left motor cortex (usually C3). In our study, for the children (wearing a 65-channel Geodesic Sensor net) EL was calculated as an average of two electrodes closest to the C3 position (electrodes 17 and 21), and ER was calculated as an average of electrodes 53 and 54 (see Fig. 3A for electrode locations). Other locations in the child EEG net were also investigated (one electrode, and a mean of four electrodes surrounding positions C3/C4). These did not provide significantly different LRPs, therefore the mean of two electrodes was selected as it provided a good spatial approximation to C3 and C4 and improved the signal-to-noise ratio. The adults' 129-channel Geodesic Sensor net had electrodes in the same positions as electrodes C3 and C4 (electrodes 36 and 104, respectively; see Fig. 3B for electrode locations).

According to the above calculation of the LRP, a negative deviation reflects correct response preparation, and a positive deviation reflects incorrect response preparation. Therefore, we expected an initial negative LRP deviation (reflecting correct response preparation) in the congruent condition (zone I in Fig. 1) and an initial positive LRP deviation (reflecting incorrect response preparation) in the incongruent condition (zone II in Fig. 1). In the incongruent condition, this would be followed by correct response preparation (zone III in Fig. 1). The LRP phenomenon of most interest, when considering IC, is the transition from incorrect response preparation to subsequent correct response preparation in the incongruent condition (arrow 10 in Fig. 1). In order to enhance the signal-to-noise ratio, each individual participant's raw LRP in each condition was smoothed by a 150 ms moving average window. An important initial step in analyzing the LRP was to test for significant deviations from zero by running point-by-point two-tailed one-sample t-tests against zero. This confirms whether the LRP shows any significant correct or incorrect response preparations. An LRP waveform deviation from zero was considered to reach significance if the p-value was less than 0.05 for more than 10 consecutive sampling points, 20 ms (similar to the procedures of Kao et al., 2009; Martens et al., 2010; Scheibe et al., 2009; Szucs and Soltesz, 2007; Szucs et al., 2009a, 2009b).

The early part of the LRP waveform is able to characterize the initial response preparation in the congruent/incongruent conditions (negative-going LRP, reflecting correct response preparation, in the congruent condition; positive-going LRP, reflecting incorrect response

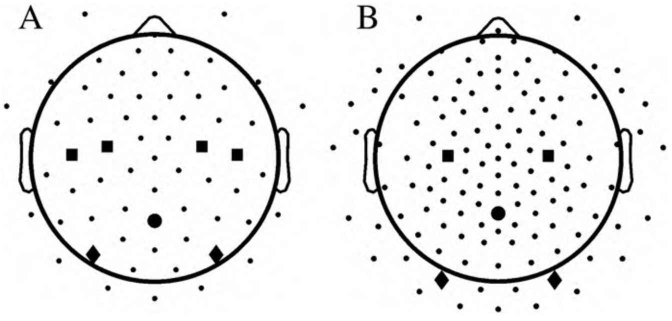

Fig. 3. The electrode nets used in the experiment for (A) children and (B) adults. Squares mark the electrodes used for LRP calculation; circles those used for the P3b; diamonds those used for P1, P2 and N2 waves.

preparation, in the incongruent condition). In both the congruent and incongruent conditions, initial peak amplitudes and latencies (arrows 1 and 2 in Fig. 1) were identified as the most extreme points (negative and positive, respectively) between 100 and 600 ms. Congruency (2) × Age Group (3) repeated measures ANOVAs were run on these peak amplitudes and latencies to examine group differences in initial LRP deviations. Independent samples t-tests were run separately on congruent and incongruent peak amplitudes to determine which groups differed from one another. Pearson correlations (controlling for age) were then run between LRP measures (peak amplitudes and latencies of initial response preparations) and behavioral measures (RTs and accuracy), to serve as an indication of the functional significance of these LRP measures.

Smoothed LRPs were jack-knifed (Miller et al., 1998, 2009; Mordkoff and Gianoros, 2000; Ulrich and Miller, 2001) to allow response preparation onsets and cessations to be calculated in a reliable way. This procedure is widely used in the LRP literature (e.g. Lansbergen and Kenemans, 2008; Scheibe et al., 2009; Schroeter and Leuthold, 2009; Wild-Wall et al., 2008). Jack-knifing is a statistical method for combining data from different participants, so that instead of individual participants, each data entry is a subaverage (calculated as the grand average with one participant removed). This way, for instance, each data entry is the mean of 15 participants, as opposed to one (when $N = 16$). This increases the signal-to-noise ratio, reduces variation, and makes additional types of analysis possible. A correction is used which compensates for the artificial reduction of variance. Initial response preparation peaks in jack-knifed LRPs were identified as the minimum or maximum (depending on condition) values within 150–500 ms. Initial response preparation onsets (arrows 3 and 4 in Fig. 1) were identified as the latencies at which the smoothed jack-knifed LRPs exceeded 75% of the relevant peak's amplitude. This conservative criterion was used to ensure that the onset related to the relevant peak in the LRP waveforms (as the children's data still contained a considerable amount of noise). Initial response preparation onset latencies were subjected to a Congruency (2) × Age Group (3) repeated measures ANOVA. Initial response preparation cessations (arrows 5 and 6 in Fig. 1) were calculated as the latency at which the LRP waveforms returned to 75% of the relevant peak's amplitude, after the peak. From onset and cessation data, durations of initial response preparations (arrows 7 and 8 in Fig. 1) were calculated as the time between the onset and the cessation of the LRP. A repeated measures Congruency (2) × Age Group (3) ANOVA was then run on initial response preparation durations.

Besides characterizing the onset, cessation and duration of initial response preparation, properties of the subsequent correct response preparation in the incongruent condition were also examined. This secondary deviation in the incongruent condition was of great interest because it reflects the transition from incorrect to correct response preparation. As such, this transition is considered the clearest measure of response inhibition, as it reflects the efficiency with which an incorrect response is ceased and replaced with a correct one. The peaks of the secondary correct response preparations in the incongruent condition were identified as the minimum values within the time range of 300–990 ms. The onset latency of the second, negative-going deviation in the incongruent LRP waveform (indicating correct response preparation, arrow 9 in Fig. 1) was then calculated in the same way as described above. A repeated measures Transition from Incorrect to Correct Response Preparation [called ‘Transition’ factor in the remainder of the paper] (2) × Age Group (3) ANOVA was run on the incongruent latency measures. The Transition factor had two levels (both measures of the LRP wave in the incongruent condition): the cessation latency of the initial incorrect response preparation (arrow 6 in Fig. 1), and the onset latency of the subsequent correct response preparation (arrow 9 in Fig. 1). For all ANOVAs using jack-knife derived variables, F-values and post hoc tests were corrected (to produce F_c) according to the criteria established by Ulrich and Miller (2001). The corrected F-value is calculated using the equation: $F_c = F/(n-1)^2$ where n is the number of observations in each cell.

Single-trial LRP

In order to assess the consistency of data across all trials, we compiled a single-trial analysis of the LRP. This analysis also allowed us to investigate the similarity in functional behavior of the LRP across age groups. More precisely, we assessed the assumed relationship between smaller incorrect and larger correct initial response preparations and faster reaction times in each age group. To reduce the volume of data presented, only the incongruent condition is presented here because this condition was our primary interest as it demonstrates inhibitory control. The single-trial LRP was calculated similarly to the normal (averaged) LRP except that ERPs recorded during left and right-hand responses were not summed and divided by two. At the level of single-trials this gives equivalent results to using the averaged LRP. That is, for trials where the correct response was a left-hand response, the ERP activity equivalent to ER–EL was used, and for trials where the correct response was a right-hand response, the ERP activity equivalent to EL–ER was used. Within each group, individual trials were pooled across all participants, and the amplitude was standardized separately for each individual. Standardization ensures that amplitudes are comparable across participants. Single-trial LRPs were sorted according to RT. For each trial, the incorrect peak latency (time of most positive point in the time period 100–700 ms) was identified. This analysis assessed the robustness and variability of single-trial data.

In order to assess the functional behavior of the LRP (i.e. the relationship between LRP amplitudes and performance, measured by RT) the 30% of trials (in each age group) with the fastest/slowest RT were averaged. The mean amplitudes of selected 100 ms windows were calculated for the fastest 30% and the slowest 30% of trials in each age group. The mean amplitudes of incorrect response preparation (a positive LRP deviation) were calculated from 225–325 ms in 5-year-olds, 300–400 ms in 8-year-olds, and 190–290 ms in adults; the mean amplitudes of secondary correct response preparation (a negative LRP deviation) were calculated from 600–700 ms in 5-year-olds, 525–625 ms in 8-year-olds, and 425–525 ms in adults. These 100 ms time windows were selected to best cover the peaks of the LRP produced by ‘fast trials’. An alternative analysis was run using windows selected to best cover the peaks of the group-averaged LRP waveform (Fig. 5B). In this case, the mean amplitudes of incorrect response preparation were calculated from 300–400 ms in 5-year-olds, 300–400 ms in 8-year-olds, and 200–300 ms in adults; and the mean amplitudes of secondary correct response preparation were calculated from 750–850 ms in 5-year-olds, 650–750 ms in 8-year-olds, and 425–525 ms in adults. Subsequent to this, a Transition (initial incorrect response preparation, subsequent correct response preparation) × Speed (fast trials, slow trials) × Age Group (5-year-olds, 8-year-olds, adults) ANOVA was run to consider effects on mean amplitudes.

ERP ANOVA

In addition to the LRP analysis, point-by-point ANOVAs were run on ERP data to examine whether differences in the speed of perceptual processing could account for the developmental LRP results observed. As real-life size difference of the animal pairings was a factor in this task (i.e. pairs of small–medium/medium–large animals constituted small size differences; small–large animal pairings constituted large size differences), it was possible to examine not only how each group responded to the congruency factor but also to this size difference factor. ANOVAs were processed separately for children and adults to avoid the analysis being conflated by general latency differences between children and adults. The children's data were analyzed using a 3-way ANOVA: Congruency (2) × Size Difference (2) × Age Group (2). The adult's data were analyzed using a 2-

way ANOVA: Congruency (2) × Size Difference (2). Time intervals that reached significance ($p<0.005$) over 10 consecutive datapoints (20 ms) at a minimum of four electrodes were considered to show significant effects. If the real-life size difference factor was processed faster with age, this would indicate that there were significant differences in speed of perceptual processing that could also account, to some extent, for the congruency effect. However, if the ERP ANOVA showed that the groups process the real-life size difference factor with similar speed, this would indicate that the congruency effect reflects the development of inhibitory control over and above differences in speed of processing.

ERP waves

Further, for consistency and to examine how the latency of major ERP phenomena may affect the data, the latency of prominent ERP peaks were analyzed. These were the peak latencies of parieto-occipital P1, N2, and P2, thought to be related to the earliest stages of visual processing, and the peak amplitude and latency of the centro-parietal P3b, thought to be sensitive to both inhibitory control and to the speed of stimulus processing (see Introduction). The early visual components were determined as the most positive (P1, P2) or negative (N2) values within specific time intervals (see below) at electrodes showing maximum amplitudes (32 and 45 in children, and 69 and 89 in adults). The P3b was determined as the most positive point within specific time intervals (see below) at electrodes showing maximum amplitudes (electrode 34 in children and 62 in adults). Electrode positions are shown in Fig. 3. Peak latencies were measured in the following intervals, P1: 100–180 ms in children, 70–140 ms in adults; N2: 150–300 ms in children, 100–200 ms in adults; P2: 300–420 ms in children, 180–260 ms in adults; and P3b: 400–700 ms in children, 300–600 ms in adults. P1, N2 and P2 peak latencies were then entered into a Congruency (2) × Hemisphere (2) × Age Group (3) ANOVA for analysis. P3b peak amplitudes and latencies were entered into a Congruency (2) × Age Group (3) ANOVA for analysis.

EEG data were processed using Brain Vision Analyzer (Brain Products, Munich), Matlab 7.9, SPSS 17.0 and Statistica 9.

Results

Behavioral results

Behavioral results are shown in Table 1. Incongruent trials were responded to less accurately than congruent trials (91 vs. 97%; $F(1, 45)=22.84$, $p<0.001$), and small size difference trials were responded to less accurately than large size difference trials (91 vs. 97%; $F(1,45)=69.24$, $p<0.001$). There was a Congruency × Age Group interaction ($F(2,45)=5.82$, $p=0.006$). Post hoc tests revealed that 5-year-olds showed a greater congruency effect than 8-year-olds (difference: 8%; $p=0.041$) and adults (9%; $p<0.001$). 8-year-olds were not significantly different from adults ($p=0.081$). There were two other significant interactions in accuracy: a Size Difference × Age Group interaction ($F(2, 45)=5.41$, $p=0.008$), showing that size difference had a greater effect on the youngest participants (driven by 5-year-olds being significantly different than both 8-year-olds, $p=0.042$, and adults, $p<0.001$). Further, a Congruency × Size Difference interaction ($F(1, 45)=44.13$, $p<0.001$) in accuracy showed there was a greater effect of congruency in the small size difference (congruency effect of 9.24%, $p<0.001$) than the large size difference (congruency effect of 1.5%, $p=0.270$) condition.

Table 1
Accuracy and reaction time in the three age groups (means and standard error of the means given).

	Accuracy (%)		Reaction time (ms)	
	Congruent	Incongruent	Congruent	Incongruent
5-year-olds	95 (1.25)	84 (3.25)	1416 (85.25)	1622 (68.5)
8-year-olds	96 (1.25)	93 (1.75)	1151 (57.25)	1274 (58)
Adults	99 (0.25)	97 (0.5)	752 (23.75)	803 (26.75)

An analysis of RTs in error trials indicated that none of the error trials were fast guesses (in the range 0–300 ms). The range of error RTs by age were 550–3760 ms in the 5-year-olds, 386–3088 ms in the 8-year-olds, 442–1138 ms in adults. In the 5-year-olds there were no error responses faster than 500 ms, there were 3 (2%) in the 8-year-old group, and 5 (9%) in the adult group. This demonstrates that the majority of error RTs were long in each group.

In raw RT, incongruent trials were responded to 127 ms slower than congruent trials ($F(1, 45)=54.42$, $p<0.001$), and small size difference trials were responded to 167 ms slower than large size difference trials ($F(1, 45)=178.08$, $p<0.001$). There was a Congruency × Age Group interaction ($F(2,45)=6.71$, $p=0.003$). Post hoc tests confirmed all contrasts were significant, and there was a congruency effect of 206 ms for the 5-year-olds, 123 ms for 8-year-olds, and 51 ms for adults. As in accuracy, there was also a significant Size Difference × Age Group interaction ($F(2, 45)=4.12$, $p=0.023$) and a Congruency × Size Difference interaction ($F(1, 45)=6.87$, $p=0.012$) in RT. These showed the same trends as in accuracy: size difference affected RTs less with age (all age groups were significantly different from one another, all $p<0.001$, size difference effect in 5-year-olds: 200 ms, 8-year-olds: 185 ms, and adults: 117 ms), and there were greater congruency effects in the small size difference (congruency effect of 153 ms, $p<0.001$) as compared to the large size difference condition (congruency effect of 101 ms, $p<0.001$).

The first analysis of RT distributions was a series of Kolmogorov–Smirnov (K–S) tests to compare the distributions of RTs. This suggested that the RT distributions were qualitatively different between age groups. Within groups, the congruent and incongruent RTs were distributed differently: comparing congruent and incongruent conditions in 5-year-olds $D=0.23$, $p<0.001$; in 8-year-olds $D=0.18$, $p<0.001$ and in adults $D=0.14$, $p<0.001$. Across groups, the 5-year-olds' incongruent condition RTs were distributed differently than both the 8-year-olds' ($D=0.06$, $p=0.047$) and the adults' ($D=0.08$, $p=0.004$). In the congruent condition, the 5-year-olds RTs were distributed differently than the adults ($D=0.07$, $p=0.005$). Within-subject variability in RTs reduced with age. In the congruent condition, the ranges for means and standard deviations for each group were: 5-year-olds mean 1002–1898 ms, standard deviation 289–864; 8-year-olds mean 824–1555 ms, standard deviation 187–448; and adults mean 601–863 ms, standard deviation 113–293. In the incongruent condition, the ranges for means and standard deviations for each group were: 5-year-olds mean 1199–2391 ms, standard deviation 301–812; 8-year-olds mean 991–1594 ms, standard deviation 207–423; and adults mean 628–994 ms, standard deviation 101–306.

Four alternative methods were used to transform the raw RTs (see Supplementary Table 1 for the mean and standard error of mean values). In standardized RTs there was a main effect of Congruency ($F(1, 45)=63.56$, $p<0.001$) but no significant Congruency × Age Group interaction ($p=0.383$). In contrast, the log transformed RTs showed a main effect of Congruency ($F(1,45)=56.08$, $p<0.001$) and a significant Congruency × Age Group interaction ($p=0.043$). Post hoc tests show that all groups differed from one another ($p<0.01$) A one-way ANOVA with an Age Group factor was run on the proportionally transformed RT data (calculated from the formula: [incongruent RT − congruent RT]/congruent RT) which showed a marginally significant Age Group effect ($F(2, 45)=3.16$, $p=0.052$). Post hoc tests showed that the difference was only marginally significant between 5-year-olds and adults ($p=0.054$).

Kolmogorov–Smirnov tests comparing raw RT distributions to the normal distribution indicated that the 5-year-old and the adult RTs in the congruent condition deviated from a normal distribution (both $p<0.05$). This, and an examination of the shape of raw RT distributions

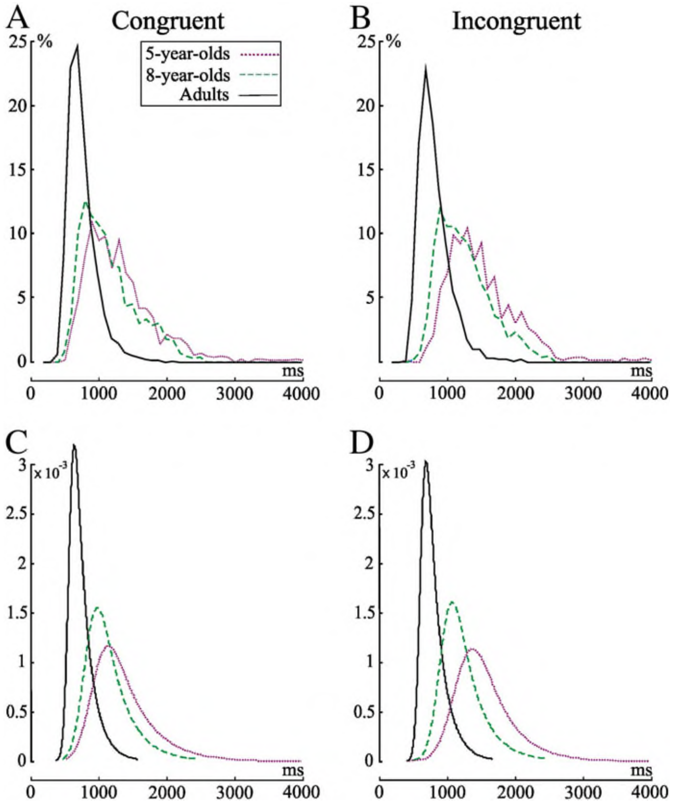

Fig. 4. (A–B) Raw reaction times and (C–D) an ex-Gaussian distribution fitted to reaction times, in (A, C) congruent and (B, D) incongruent conditions. Dotted purple lines depict the 5-year-olds, dashed green lines depict the 8-year-olds, solid black lines depict the adults. Raw RTs are expressed as a percentage of all RTs; ex-Gaussian distributions are expressed as probability densities.

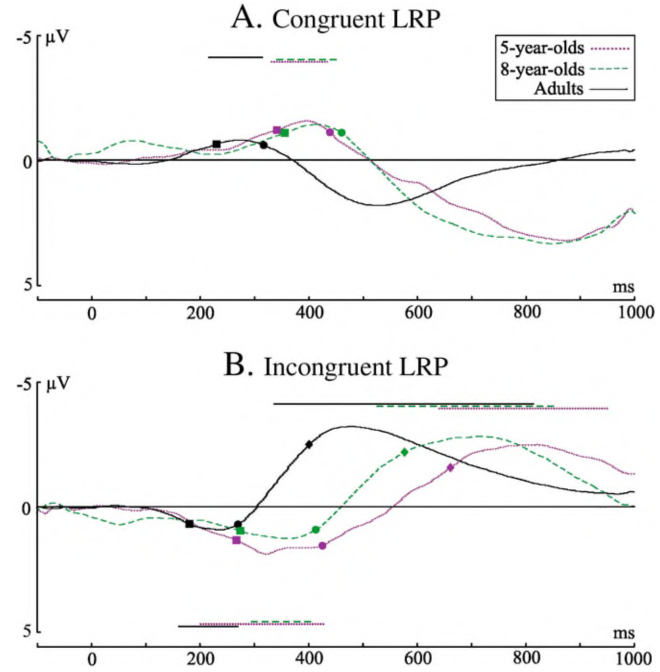

Fig. 5. The smoothed Lateralized Readiness Potential (LRP) in (A) congruent and (B) incongruent conditions. Dotted purple lines depict the 5-year-olds' LRP, dashed green lines depict the 8-year-olds' LRP, solid black lines depict the adults' LRP. Squares denote initial response preparation onsets, circles response preparation cessations, and diamonds response preparation secondary onsets. Horizontal lines indicate regions in which the LRP deviated significantly from zero.

(Figs. 4A–B), suggested that fitting an ex-Gaussian distribution may provide a better representation of the data. Fig. 4 depicts the distribution of raw RTs (A–B) and an ex-Gaussian distribution fitted to the RTs (C–D). There was a significant main effect of Congruency ($F(1, 45) = 29.73$, $p < 0.001$) and a significant Congruency × Age Group interaction ($F(2, 45) = 5.83$, $p = 0.006$) on the measure representing the mean (μ). Post hoc tests indicated that adults differed significantly from 5-year-olds ($p < 0.001$) and 8-year-olds ($p = 0.001$) but the two child groups were not significantly different from one another ($p = 0.093$). There were no main effects or interactions on the other parameters (σ and τ).

LRP measures

LRPs are shown in Fig. 5, split by congruency. As expected, initially there was a significant negative deviation in the congruent LRP (5-year-olds: 330–434 ms, 8-year-olds: 338–454 ms; adults: 216–316 ms) and a significant positive deviation in the incongruent LRP (5-year-olds: 200–428 ms, 8-year-olds: 294–404 ms; adults: 162–272 ms). These initial LRP deviations reflect correct response preparation in the congruent condition, and incorrect response preparation in the incongruent condition. First, an ANOVA was run on the peak amplitude of the initial LRP deviation. This ANOVA showed the expected main effect of Congruency ($F(1, 45) = 105.83$, $p < 0.001$) and demonstrated a significant Congruency × Age Group interaction ($F(2, 45) = 8.95$, $p = 0.001$), whereby the peak amplitudes became less extreme with age. Independent samples t-tests were run separately on congruent and incongruent peak amplitudes to determine which groups differed from one another. These results indicated that the two child groups' peak amplitudes did not differ from one another in either condition, 5-year-olds differed from adults on both conditions (congruent: $t(30) = -2.344$, $p = 0.026$; incongruent: $t(21.7) = 2.766$, $p = 0.011$), and 8-year-olds showed a marginally significant difference from adults only on the congruent condition ($t(23.4) = -2.039$, $p = 0.053$). There were no significant effects in the ANOVA run on peak latencies, suggesting that the timing of the initial peaks of the LRP were comparable across congruency conditions and age groups.

As all of the behavioral and peak measures correlated with age (see Supplementary Table 2), Pearson correlations, controlling for age, among peak measures (amplitude and latency) and behavioral measures (reaction time and accuracy) were run on the whole

Table 2
Lateralized Readiness Potential onsets, cessations and durations (means and standard error of the means given).

	LRP measures in congruent condition			LRP measures in incongruent condition			
	Correct onset	Correct cessation	Duration	Incorrect onset	Incorrect cessation	Duration	Correct onset
5-year-olds	340 (2.56)	437 (0.81)	97 (2.53)	268 (1.07)	425 (3.10)	158 (3.17)	662 (2.13)
8-year-olds	355 (1.47)	462 (0.84)	106 (1.38)	277 (2.48)	413 (1.26)	136 (2.86)	575 (1.84)
Adults	230 (0.74)	317 (0.93)	87 (1.33)	184 (0.54)	272 (0.69)	89 (0.93)	401 (0.63)

dataset. The congruent LRP peak amplitude showed a significant correlation with congruent RT ($r = -0.35$, $p < 0.001$) and the incongruent LRP peak latency correlated significantly with accuracy in the congruent condition ($r = -0.45$, $p = 0.002$).

The mean onsets and cessations (defined as the latency at which the smoothed jack-knifed LRP reached or returned to 75% of its peak amplitude) are given in Table 2 and represented on Fig. 5 as squares (initial response preparation onsets), circles (initial response preparation cessations) and diamonds (secondary response preparation onsets).

An ANOVA on the initial response preparation onset latencies showed both a main effect of Congruency ($F_c(1, 45) = 13.41$, $p < 0.001$) and Age Group ($F_c(2, 45) = 9.18$, $p < 0.001$), but the Congruency × Age Group interaction was not significant ($F_c(2, 45) = 0.30$, $p = 0.742$; see Fig. 6A). Onsets were earlier in the incongruent condition (243 ms) than in the congruent condition (309 ms). Tukey post hoc tests on the Age Group factor indicated that both child groups had significantly later initial response preparation onset latencies than the adult group (both $p < 0.05$) but they were not significantly different from one another (5-year-olds: 304 ms, 8-year-olds: 316 ms, adults: 207 ms).

The duration of initial response preparations were calculated in both the congruent and incongruent conditions as the time interval between onset and cessation of initial (correct/incorrect) response preparation. A repeated measures ANOVA run on duration values showed no main effects of Congruency ($F_c(1, 45) = 1.14$, $p = 0.292$) or Age Group ($F_c(1, 45) = 0.94$, $p = 0.400$), nor a significant Congruency × Age Group interaction ($F_c(2, 45) = 0.35$, $p = 0.704$). However, the data (see Fig. 6B) suggest a trend for improved discrimination between the congruency conditions with age.

As predicted in the incongruent condition, incorrect response preparation was followed by subsequent correct response preparation represented by a significant negative deviation relative to baseline (5-year-olds: 640–952 ms, 8-year-olds: 524–854 ms; adults: 334–816 ms). As previously mentioned, the transition from incorrect to correct response preparation in the incongruent condition can be considered a reflection of successful response inhibition. Therefore, a repeated measures ANOVA was run using two measures from the incongruent LRP: cessation of the incorrect response preparation and onset of the correct response preparation. There was the expected main effect of Transition ($F_c(1, 45) = 105.85$, $p < 0.001$), a main effect of Age Group ($F_c(2, 45) = 21.97$, $p < 0.001$) and a significant Transition × Age Group interaction ($F_c(2, 45) = 3.48$, $p = 0.039$; see Fig. 5C). Post hoc tests on the Age Group factor indicated that both child groups had significantly longer Transition latencies than the adult group (both $p < 0.05$) but they were not significantly different from one another (transition time in 5-year-olds: 237 ms, 8-year-olds: 162 ms, adults: 129 ms).

The LRP produced in this animal Stroop task also contained some unexpected incorrect response preparation in the congruent condition, in all three groups. As we did not see this pattern in our previous study (Szucs et al., 2009a), it can only be assumed that the incorrect response preparation is a task-specific phenomenon. We can identify three possible explanations of this phenomenon: first, a prolonged period of self-checking before responding caused by the three real-life size groups of animals making the task less predictable; second, a range of different strategies being employed by participants, supported by the considerable variation in reaction times; third, additional motor cortex activity following the initial response preparation that is not response-related, perhaps caused by using a different response device in our current study. However, the longer epochs provided for illustrative purposes (Supplementary Fig. 1) show that the LRP in the incongruent condition is only composed of a significant positive deviation followed by a significant negative deviation, supporting the use of the Transition measures.

Single-trial LRP

In order to assess inter-trial variability and the functional behavior of the LRP across age groups, single-trial analysis (Fig. 7) was used. Figs. 7A–C show trials in the incongruent condition, for each age group, sorted by RT. Visual inspection reveals that in the incongruent condition, a strong negative LRP deviation precedes the response in most of the trials (most visible in the adult data, Fig. 7C). The image of the adult's data demonstrates that the incorrect response preparation in the incongruent condition (red in color) is extremely consistent in duration, whereas the subsequent correct response preparation (blue in color) increases in length when RT is longer. While the child data are clearly noisier and show greater trial-by-trial variability, a similar pattern is discernable (Figs. 7A and B), which suggests that the LRP is functionally similar in children and adults. This is confirmed by the following analyses.

Variability in the peak latency of the initial positive deviations of the LRP in each individual trial (indicating incorrect response preparation) is represented by boxplots in Figs. 7D–F. The mean latencies of the incorrect peaks in single-trials were 301 ms in 5-year-olds, 313 ms in 8-year-olds, and 247 ms in adults. This representation corroborated the group-averaged data, as the equivalent values in group-averaged LRPs were 362 ms, 324 ms and 247 ms. As expected, the youngest group, with the noisiest data, showed the greatest discrepancy between methods (61 ms). Remarkably, the single-trial

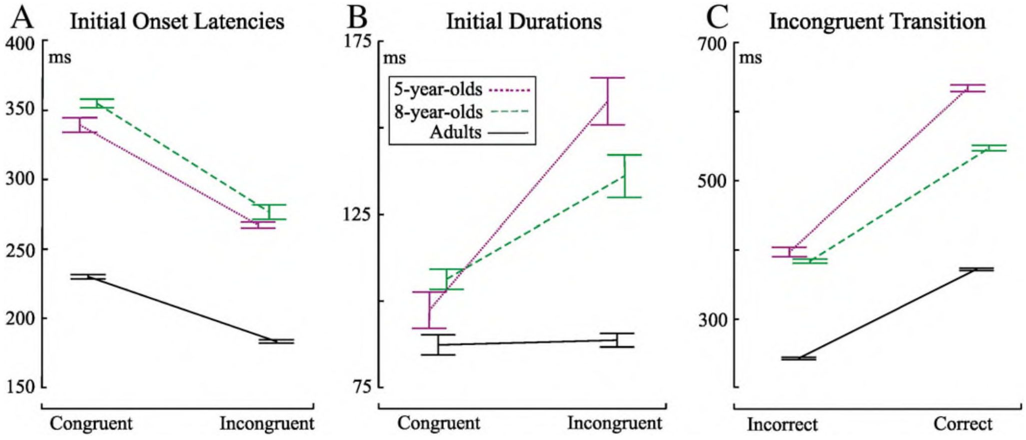

Fig. 6. Graphs showing interaction effects on (A) initial response preparation onset latencies, (B) initial response preparation durations and (C) the transition from incorrect to correct response preparation in the incongruent condition. Dotted purple lines depict the 5-year-olds' data, dashed green lines depict the 8-year-olds' data, solid black lines depict the adults' data. Error bars represent 95% confidence intervals.

latency data are in excellent agreement with the averaged group data, despite the commonly held view that single-trial LRPs are unreliable (Stahl et al., 2010).

The functional similarity of LRP behavior was assessed by comparing the mean incongruent LRPs for the fastest and slowest 30% of trials from each group (Figs. 7G–I). In comparison to the slowest trials, the fastest trials in each age group have a less extreme (in amplitude) incorrect response preparation, followed by a more extreme (in amplitude) correct response preparation. This is confirmed by an analysis of the mean amplitudes in selected 100 ms periods (based on the peaks of the 'fast trials' LRP; highlighted in Figs. 7G–I). A Transition (2) × Speed (2) × Age Group (3) ANOVA was run on these mean amplitudes. This showed, as well as the predicted main effect of Transition ($F(1, 2262) = 10152.89$, $p < 0.001$), a

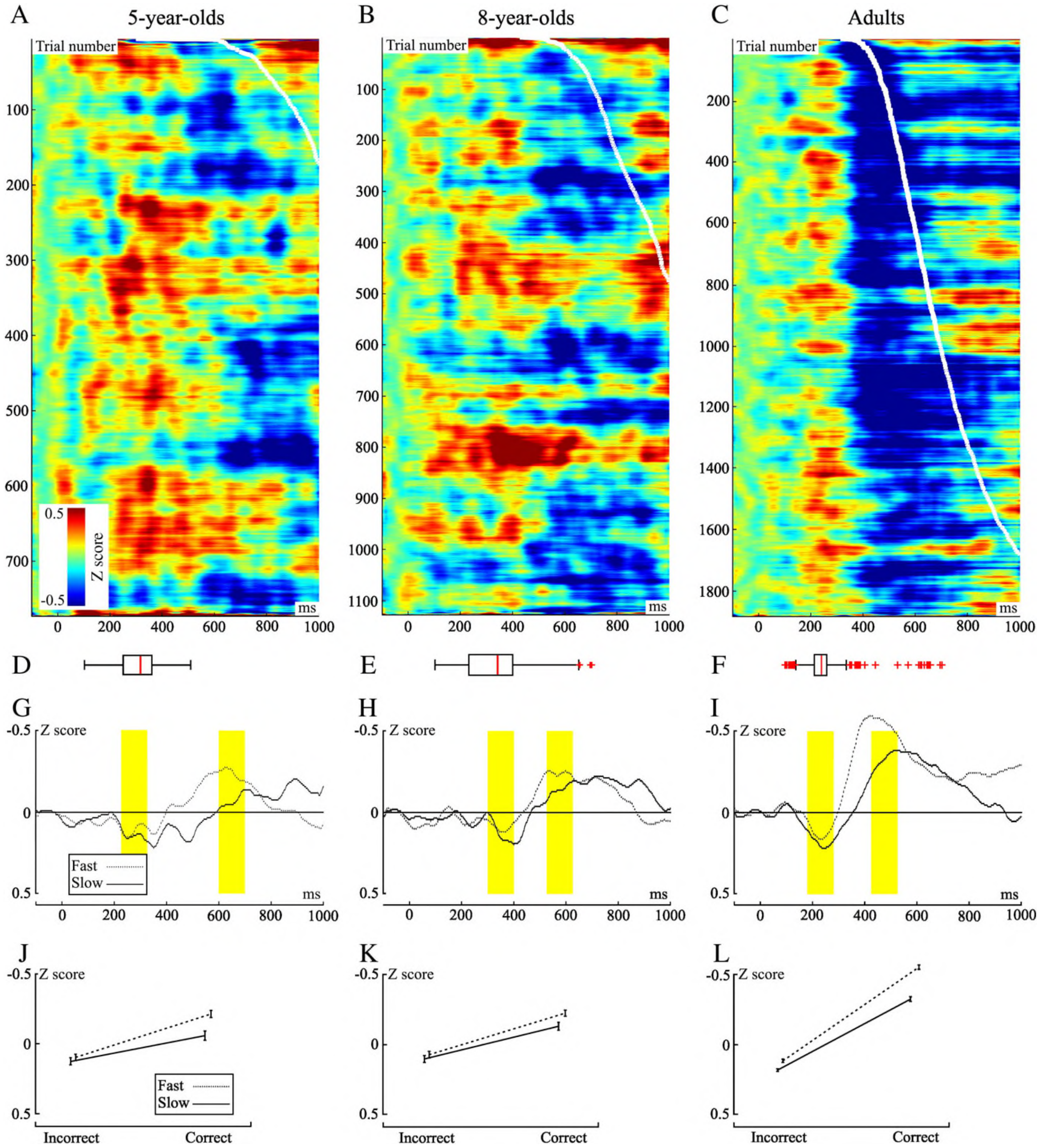

Fig. 7. (A–C) Single-trial Lateralized Readiness Potential (LRP) analysis results for the incongruent condition in 5-year-olds, 8-year-olds and adults. Trials are sorted according to RT (white line) and red coloring represents positive LRP amplitudes (incorrect response preparation), blue represents negative LRP amplitudes (correct response preparation). (D–F) Boxplots summarizing the positive peak latencies (incorrect response preparation). (G–I) LRP of the 30% fastest and 30% slowest trials in each age group, with the highlighted areas indicating time periods used to calculate mean amplitudes. (J–K) Mean amplitudes (from the highlighted 100 ms windows), averaged for fast and slow trials, for incorrect and correct response preparations. Error bars represent 95% confidence intervals.

significant Transition × Speed interaction ($F(1, 2262) = 275.03$, $p < 0.001$), a significant Transition × Age Group interaction ($F(2, 2262) = 881.55$, $p < 0.001$) and a significant three-way Transition × Age Group × Speed interaction ($F(2, 2262) = 14.29$, $p < 0.001$), represented in Figs. 7J–L. The alternative method for selecting windows within which to calculate mean amplitudes (based on the peaks of the group-averaged LRP) produced the same result. The interaction of most interest (Transition × Speed interaction) shows that the fast trials had a significantly less extreme mean amplitude in the incorrect response preparation than slow trials (fast: 0.11 vs. slow: 0.15 μV), and a significantly more extreme mean amplitude in the correct response preparation than slow trials (fast: −0.35 vs. slow: −0.18 μV). This shows that those trials that resulted in fast RTs demonstrated a weaker initial incorrect response preparation, and a stronger subsequent correct response preparation, providing evidence for the assumed relationship between LRP deviations and RT.

To summarize, the single-trial analysis of the incongruent LRP has supported the assumed relationship between response preparations as measured in the LRP and RT. This analysis has demonstrated that despite inter-trial variability the pattern of the LRP is easily detected even in the (relatively) noisy child data. Further, it has provided robust evidence that the LRP is functionally similar in 5 and 8-year-old children and adults, and is therefore a valid measure in developmental EEG research.

ERP ANOVA

ERP amplitude was tested by point-by-point ANOVAs. The most important objective of this analysis was to examine when real-life size differences were processed, and whether this differed between groups as the timing of real-life size effects can serve as a measure of perceptual processing speed and attentional selection of the task-relevant stimulus dimension. In the point-by-point ANOVA on ERP activity from the two child groups, crucially, there was no Age Group × Size Difference interaction (even when the significance threshold was reduced to $p < 0.05$) which suggests that 5-year-olds and 8-year-olds process the size difference with similar speed. In addition, size difference main effects were not delayed in children (80–170 ms, 250–500 ms, 570–670 ms, 740–790 ms) relative to adults (200–230 ms, 400–460 ms, 550–700 ms). These data suggest that the processing of size difference was not slower in children than in adults.

ERP waves

In order to check whether any congruency related LRP effects were driven by congruency effects on the latencies of ERP waves, the latencies of characteristic ERP phenomena were tested by Congruency × Hemisphere × Age Group ANOVAs. As the P3b has been associated with inhibitory control, this component was examined for consistency with the literature. The P1, N2 and P2 have also been considered to be indicators of visual processing. The ANOVAs on the peak latencies of P1, N2, and P2 showed no significant effects. Likewise, the Congruency × Age Group ANOVA on the P3b wave showed no significant effects in latency or amplitude.

Discussion

We examined how IC develops across 5-year-olds, 8-year-olds, and adults. To this end, we analyzed various RT measures, the LRP and ERPs. We found that the LRP revealed substantial developmental differences in two processes that contribute to IC: first, children were slower to initiate their response preparations in general than adults, but each group showed a similar congruency effect (as indexed by initial LRP onsets); second, children also differed from adults in how they overcame conflict in the incongruent condition, which became faster with age (as indexed by the transition from incorrect to correct response preparation in the incongruent condition). The analysis of RT and ERPs demonstrated that the developments in IC across these age groups cannot be accounted for by changes in speed of processing.

Congruency effects and speed of processing

As expected, performance (accuracy and raw RT) was superior in the congruent as compared to the incongruent condition, and this congruency effect lessened with age. Two out of three transformations that aimed to account for speed of processing (proportional and logarithmic transformations) also confirmed that the congruency effect decreases with age, whereas standardizing RTs according to a Gaussian distribution did not. These data are similar to Johnson et al. (2003) and Christ et al. (2001). However, as Faust et al. (1999) highlight, standardized RTs should be interpreted with caution when the RT distributions differ between groups and conditions. This was the case in this study, and therefore transforming the RTs according to the Gaussian distribution is considered the least reliable approach. The fourth transformation applied to the RT data, fitting an ex-Gaussian distribution to raw RTs, is thought to be a better model of RT distributions than the Gaussian distribution (Fagot et al., 2009; Leth-Steensen et al., 2000). The data supported this proposal, as Kolmogorov–Smirnov tests indicated some RT distributions deviated from the Gaussian distribution. There was a robust Congruency × Age Group interaction on the ex-Gaussian parameter μ, indicating that there were greater congruency effects in children than adults. These ex-Gaussian RT findings replicate Fagot et al.'s (2009) results, using a simpler Stroop task in younger children. Overall, four out of five RT measures indicated that Stroop congruency effects in RT do decrease with age. Hence, we conclude that it is unlikely that age-related changes in congruency effects could be explained by changes in general stimulus processing speed. This conclusion is in line with our ERP data (see below).

It is important to note that there were variations in our findings (the results of pair-wise comparisons) regarding what drives the interaction in RT data. In raw and log transformed RTs all groups differed from one another, in the proportional transformation only 5-year-olds and adults differed, and in the ex-Gaussian parameter μ adults differed from both child groups. This highlights the issue that different transformations of raw data may lead to slightly different results and interpretations. Faust et al. (1999) suggest that proportional transformations are more appropriate than log transformations, as log transforms rescale variables in a non-linear fashion. However, they also note that proportional transformations have the disadvantage of being more susceptible to Type 1 errors in Age Group × Condition interactions than standardized RTs (Faust et al., 1999). Further, while these transformations aim to account for differences in processing speed, it must be noted that in fact they only account for task-specific processing speed (Faust et al., 1999) as all the RT measures are from the animal Stroop task. Taking into account the drawbacks of using these transformations, and the strong case for representing RTs using an ex-Gaussian distribution (Leth-Steensen et al., 2000; McAuley et al., 2006; Mewhort et al., 1992; Myerson et al., 2007) rather than Gaussian or logarithmic distributions, we suggest that more confidence could be placed in the ex-Gaussian results in this case. Our findings highlight the need to consider RT distributions more closely in developmental studies. Indeed, a comparison of RT distributions established that there are differences in how RTs are distributed between the youngest and the other two groups, within conditions. This shows age-related differences in the shape of RT distributions, from 5 to 8-year-olds. Other studies have also shown that RT distributions are qualitatively different in children with disorders, such as ADHD (Leth-Steensen et al., 2000). This suggests that thorough exploration of RT data holds great potential to shed more light on the development of IC, and other skills.

Besides the RT analysis, two ERP analyses were also used to test whether congruency effects could be explained by differences in speed of processing between age groups. First, size difference effects (ERP amplitude effects as a function of real-life size difference between the two animal stimuli) served to measure the speed of attentional capture of the task-relevant stimulus dimension. Significant real-life size difference effects appeared with the same latency in adults and children, suggesting that adults were no more efficient in focusing on the task-relevant dimension. Second, the timing of early visual processing wave peaks (parieto-occipital P1, N2, P2) coincided across all groups. In addition, the latency and amplitude of the P3b peak which is frequently used as a measure of stimulus processing speed was also similar in all groups (Duncan-Johnson and Kopell, 1981; Ilan and Polich, 1999; Kutas et al., 1977). These ERP results corroborate the RT finding that developmental congruency effects cannot be explained by differences in processing speed or attentional capture of the task-relevant stimulus dimension among the groups. Additionally, it is important to note that while several investigators have used the P3b as a measure of IC (e.g. Bruin et al., 2001; Davis et al., 2003; Smith et al., 2007), in this context we have no evidence that the P3b was related to IC.

The development of inhibitory control captured by the LRP

The current study provides further evidence that the LRP is an excellent tool with which to measure IC in children, as demonstrated previously (Ridderinkhof and van der Molen, 1995; Szucs et al., 2007, 2009a). There were four main developmental findings from the LRP analysis. First, the LRP showed, in all age groups, the expected initial correct/incorrect response preparation in congruent/incongruent trials (respectively), followed by correct response preparation in the incongruent condition. Second, replicating our previous results (Szucs et al., 2009a), we found that initial response preparation peak amplitudes became less extreme with increasing age. Specifically, the youngest children have stronger initial responses (i.e. more extreme amplitudes) in both conditions than adults, while the 8-year-olds only have stronger initial responses in the congruent condition as compared to adults. As the LRP is computed as an average of correct and incorrect response preparation, this most probably suggests that the youngest group have a stronger incorrect response to overcome in the incongruent condition. LRP latency data also provided a wealth of novel information. These latency data are summarized in Fig. 8. The third LRP finding was that in all groups, initial response preparation in the incongruent condition had an earlier onset than in the congruent condition. The congruency effect in initial response preparation onset latencies did not change with age, nor was there an interaction on response preparation duration measures. This suggests that all age groups were similar in their ability to effectively suppress incorrect motor response tendencies elicited by physical size differences which were evaluated quicker than real-life size differences. Fourth, in the incongruent condition the transition time from incorrect to correct response preparation was considerably longer in children than in adults (compare the time intervals marked by thick blue lines in Fig. 8). This finding suggests that the correction of incorrect response tendencies took more time in younger than older participants.

We suggest that the above findings can be understood to provide supporting evidence for Bunge's theory that there are at least two dissociable components of IC — one 'early' (interference suppression) and one 'late' (response inhibition), discussed in the Introduction. We have been able to identify these two contributory processes within one condition of a task using the LRP, and these data indicate that they have different developmental trajectories. The 'early' process, which is based upon the initial response preparation variables in this study, reflects how the conflict in the task is initially experienced. The data show that 5-year-olds, 8-year-olds and adults have equivalent congruency effects in the onset latencies of initial response preparations (referring to LRP finding three in the above paragraph). These latency findings suggest that the timing of initial responses to congruent and incongruent stimuli is fairly mature (although delayed in comparison to adults) even by 5 years of age. However, the larger incorrect response preparation amplitude in 5-year-olds suggests that the youngest group of children experience greater interference effects than the other groups at this early stage (referring to LRP finding two in the above paragraph). These results corroborate Jongen and Jonkman's (2008) finding that stimulus interference control was mature in their youngest participants (age 6 years); using a different experimental approach we have evidence it may even be mature by 5 years. Bunge et al. (2002) found that different brain regions were active during interference suppression in children (aged 8–12 years) and adults. However, these findings — that the early interference suppression process reaches maturity at different stages of development — are not necessarily contradictory, as fMRI and EEG measure different aspects of brain functioning.

The 'late' process, which is based upon the transition variables in this study, reflects the correction of incorrect motor tendencies in the incongruent condition. The data show that this increases in speed with age, suggesting that this later IC process is still developing after 8 years of age (referring to LRP finding five). This process reflects a combination of the incorrect motor tendencies being inhibited and the correct response being organized, also referred to in the literature (see Introduction) as response inhibition (Bunge et al., 2002) and response interference control (Jongen and Jonkman, 2008). Our results are in agreement with previous findings — that response inhibition is different in adults and in children aged 8–12 years

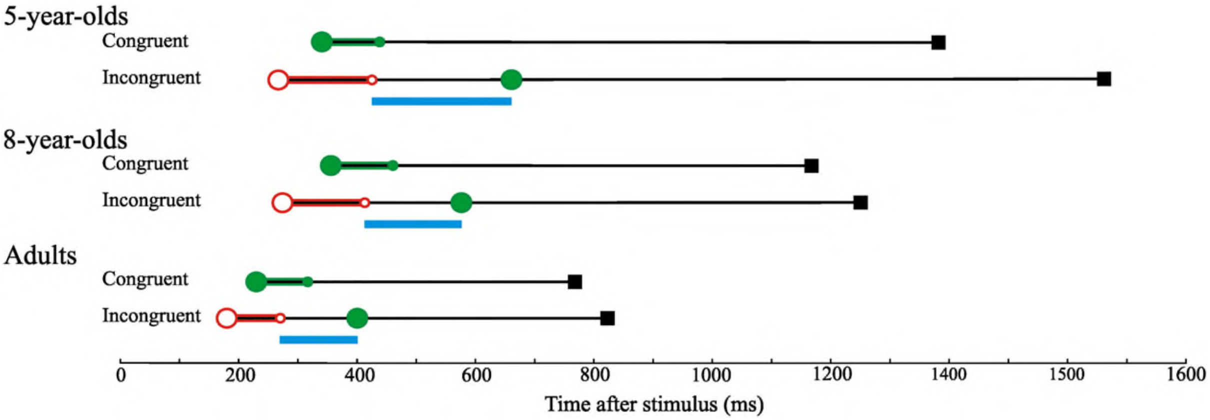

Fig. 8. Summary diagram of Lateralized Readiness Potential measures in each age group and each condition. The solid green dots relate to correct response preparation, the empty red dots relate to incorrect response preparation. The large dots represent response preparation onsets, small dots represent response preparation cessations, squares represent reaction times. Thick red and green lines represent initial response preparation durations. Thick blue lines under the incongruent condition represent the transition from incorrect to correct response preparation.

(Bunge et al., 2002), and response interference control only reaches maturity in a Stroop task by 12 years (Jongen and Jonkman, 2008).

The LRP analyses in this study show that only the late 'response inhibition' process shows developmental progression between children and adults. However, visual inspection of the initial response preparation duration measure (Fig. 6B) suggests that the youngest children do show a greater congruency effect in this measure, but this does not reach significance. This could be partly attributed to the strict nature of the F-value correction required when using jack-knifed variables (Ulrich and Miller, 2001), particularly when applied to EEG data produced by children which has much higher variability than adult data. Indeed, the peak amplitude data indicate that 5-year-olds experience a stronger incorrect initial response than older participants, suggesting that the early 'interference suppression' process does show development in this age range. These results perhaps do not reach significance because the 5-year-olds who survive the artifact rejection have more mature IC skills than other 5-year-olds, as they are able to sit still and blink on command. Perhaps with a larger sample or using younger participants we would also detect differences in the early 'interference suppression' process.

There is reason to believe that interference suppression does develop in this age range, as a number of key developments in self-regulation occur simultaneously. For instance, at 5 years of age children begin to internalize their private speech (Flavell et al., 1997), suggesting they are moving from being other- to self-regulated. Further, around 6 to 7 years of age, children become able to consciously monitor and control their behavior, and report learning strategies (Flavell et al., 1966; Keeney et al., 1967; Veenman et al., 2006). Young children's poor performance on memory tasks (Schneider and Pressley, 1997), and in a range of other domains (Hacker et al., 1998) is often considered to be caused by a breakdown in links between monitoring (possibly related to our observed 'early' interference suppression IC process) and control (possibly related to our observed 'late' response inhibition IC process), where the information made available through monitoring processes is underutilized in controlling subsequent performance. Clearly further research is required, but perhaps this 'early' process we have detected with the LRP, considered to reflect the ability to suppress interference, is related in a meaningful way to these important self-regulatory achievements. The fact that there is a protracted developmental trajectory of Eriksen flanker and Stroop tasks (Huizinga et al., 2006) and cognitive flexibility tasks (Davidson et al., 2006), could be attributed to the 'late' process evidenced by the LRP, the development of response inhibition.

Another interesting developmental finding emerges from the single-trial LRP analysis, where the representation of the children's data shows less consistency than the adult's data. This can partly be attributed to the data containing more noise and fewer trials. However, regardless of these reasons, it may not be surprising to see less consistency in the child data compared with the adult data, as the children's ages were specifically chosen because of rapid development in IC during this period. Therefore, some 5-year-olds would have IC skills similar to an 8-year-old, and vice versa — they are a less homogenous group than the adults, in terms of IC. In contrast, the adult group were all older than 23 years of age, after IC is thought to have finally reached maturity (Huizinga et al., 2006).

However, it should be noted that the LRP measures motor response preparation, and therefore only the later motor stages of IC. Therefore, one can only make indirect conclusions about early pre-motor IC processes based on LRP measures (in a similar manner to those studies that infer the locus of psychological phenomena from the LRP, e.g. Freitas et al., 2009; Hsieh and Yu, 2003; Jaskowski et al., 2007, 2009; Osman and Moore, 1993). While the existing literature can aid our interpretation of these LRP findings, we cannot fully determine the function of the two processes at this stage as any interpretation regarding pre-motor processes in IC using the LRP remains conjecture (for detailed discussion see Szucs and Soltesz, 2007; Szucs et al., 2009b).

Functional behavior of LRP across age groups

As yet, very few studies have used the LRP in children (Ridderinkhof and van der Molen, 1995; Szucs et al., 2007, 2009a) and even fewer have used it to study incorrect response preparation in children (Szucs et al., 2007, 2009a). Therefore, it was important to establish the robustness of findings and whether the LRP functions similarly in children and adults. To this end, a single-trial analysis was carried out. This analysis demonstrated that LRP deviations were very consistent across trials, showing a remarkable robustness of the LRP in these data even at the single-trial level. This indicated that the averaged LRPs were not dominated by a small subsection of trials. The analysis also revealed that the LRP behaved in the same way in all three groups: fast incongruent trials had weaker incorrect (i.e. less positive mean amplitudes), and stronger correct (i.e. more negative mean amplitudes) response preparations than slow trials. To our knowledge, only one paper has attempted to examine single-trials of the LRP to date — Stahl et al. (2010). However, in that case, single-trials were modelled from LRP averages as opposed to examining the raw single-trials. In our analysis a surprisingly coherent picture emerged from compiling many raw LRP trials. The incongruent condition of the animal Stroop task elicited very clear incorrect and subsequent correct response preparation which doubtless contributed to producing such a coherent picture. Considering the LRP from this perspective has contributed greatly to our understanding of LRP effects across the whole dataset, and the functional significance of the LRP.

Implications for education

Our findings have clear implications for education. The dissociation of two separate processes that contribute to successful IC could greatly inform educational research, in terms of the timing and nature of interventions that are developed. That is, it may not be particularly fruitful to provide interventions that address interference suppression or conflict monitoring for children over 5 years of age, whereas interventions that address response inhibition or controlling conflict may be beneficial for children up to 8 years of age. Indeed, these findings may help to explain the neural basis of why young children and those with poor working memory are easily distracted, for instance in a busy classroom with conflicting demands (Gathercole et al., 2006). Perhaps they need more time to notice conflict or suppress interference in the environment, which could be assessed using the methods identified here.

This dissociation between 'early' interference suppression and 'late' response inhibition IC processes could also shed light on a number of developmental disorders associated with IC such as ADHD (Corbett et al., 2009; Happe et al., 2006; Nigg, 2001; Ozonoff and Jensen, 1999), autism (Corbett et al., 2009; Solomon et al., 2008), and PKU (Diamond et al., 1997). For instance, it could be that children with autism may have more pronounced deficits in the 'early' interference suppression process, while children with ADHD may have more pronounced deficits on the later response inhibition process, based on their behavioral profiles. Educationally, the bilingual advantage in IC (Bialystok et al., 2004; Bialystok and Martin, 2004; Carlson and Meltzoff, 2008; Martin-Rhee and Bialystok, 2008) would be an interesting phenomenon to study using this LRP approach. In this case, we would expect that bilinguals would show an enhanced interference suppression process as compared to monolinguals, but no advantage on the later response inhibition process that reflects overcoming conflict. This is because the superior IC in bilinguals is hypothesized to be driven by practice in suppressing alternative representations of words, rather than motor response processes. This has been found in behavioral studies using two different tasks (e.g. Martin-Rhee and Bialystok, 2008), but to our knowledge has not yet been dissociated within one task that requires an amalgamation of both stages of IC.

Conclusion

We have studied the development of IC in 5-year-olds, 8-year-olds and adults in an animal Stroop task. We focused on latent motor response organization in correctly responded trials of the incongruent condition using various LRP parameters as dependent measures. We have controlled for the effect of increasing processing speed across the age groups using various ERP and RT measures. LRP data have dissociated the development of two processes that contribute to IC. First, all age groups initiated their initial response tendency later in the congruent than incongruent condition, and to a similar extent. Second, there was development between children and adults in the efficiency of correcting the incorrect response tendency and producing the correct response tendency. We suggest that these observations reveal the dissociated development of interference suppression and response inhibition. These results may inform the development of educational interventions and our understanding of developmental disorders. In this case, EEG has provided a unique insight to the developmental trajectory of IC — a skill that is clearly implicated in the early development of cognitive control processes which are crucial for educational success.

Supplementary materials related to this article can be found online at doi:10.1016/j.neuroimage.2010.12.006.

Acknowledgments

This work was funded by an ESRC PhD studentship grant awarded to D.B. and a Research and Development grant from the Faculty of Education to D.S.

References

Adleman, N.E., Menon, V., Blasey, C.M., White, C.D., Warsofsky, I.S., Glover, G.H., Reiss, A.L., 2002. A developmental fMRI study of the Stroop color-word task. Neuroimage 16, 61–75.

Alloway, T.P., Gathercole, S.E., Adams, A., Willis, C., Eaglen, R., Lamont, E., 2005. Working memory and phonological awareness as predictors of progress towards early learning goals at school entry. British Journal of Developmental Psychology 23, 417–426.

Bialystok, E., Martin, M.M., 2004. Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. Developmental Science 7, 325–339.

Bialystok, E., Craik, F.I.M., Klein, R., Viswanathan, M., 2004. Bilingualism, aging, and cognitive control: evidence from the Simon task. Psychology and Ageing 119, 290–303.

Blair, C., Razza, R.P., 2007. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. Child Development 78, 647–663.

Brass, M., Yves von Cramon, D., 2002. The role of the frontal cortex in task preparation. Cerebral Cortex 12, 908–914.

Brock, L.L., Rimm-Kaufman, S.E., Nathanson, L., Grimm, K.J., 2009. The contributions of 'hot' and 'cool' executive function to children's academic achievement, learning-related behaviors, and engagement in kindergarten. Early Childhood Research Quarterly 24, 337–349.

Bruin, K.J., Wijers, A.A., van Staveren, A.S.J., 2001. Response priming in a go/nogo task: do we have to explain the go/nogo N2 effect in terms of response activation instead of inhibition? Clinical Neuropsychology 112, 1660–1671.

Bunge, S.A., Wright, S.B., 2007. Neurodevelopmental changes in working memory and cognitive control. Current Opinion in Neurobiology 17, 243–250.

Bunge, S.A., Dudukovic, N.M., Thomason, M.E., Vaidya, C.J., Gabriela, J.D.E., 2002. Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. Neuron 33, 301–311.

Carlson, S.M., 2005. Developmentally sensitive measures of executive function in preschool children. Developmental Neuropsychology 28, 595–616.

Carlson, S.M., Meltzoff, A.N., 2008. Bilingual experience and executive functioning in young children. Developmental Science 11, 282–298.

Chikazoe, J., Jimura, K., Hirose, S., Yamashita, K., Miyashita, Y., Konishi, S., 2009. Preparation to inhibit a response complements response inhibition during performance of a Stop-Signal task. The Journal of Neuroscience 29, 15870–15877.

Christ, S.E., White, D.A., Mandernach, T., Keys, B.A., 2001. Inhibitory control across the life span. Developmental Neuropsychology 20, 653–669.

Coles, M.G.H., 1989. Modern mind–brain reading: psychophysiology, physiology and cognition. Psychophysiology 26, 251–269.

Corbett, B.A., Constantine, L.J., Hendren, R., Rocke, D., Ozonoff, S., 2009. Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder and typical development. Psychiatry Research 166, 210–222.

Davidson, M.C., Amso, D., Anderson, L.C., Diamond, A., 2006. Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. Neuropsychologia 44, 2037–2078.

Davis, E.P., Bruce, J., Snyder, K., Nelson, C.A., 2003. The X-trials: neural correlates of an inhibitory control task in children and adults. Journal of Cognitive Neuroscience 15, 432–443.

de Houwer, J., 2003. On the role of stimulus–response and stimulus–stimulus compatibility in the Stroop effect. Memory & Cognition 31, 353–359.

de Jong, R., Wierda, M., Mulder, G., Mulder, L.J.M., 1988. Use of partial stimulus information in response processing. Journal of Experimental Psychology: Human Perception and Performance 14, 682–692.

Deak, G.O., Narasimham, G., 2003. Is perseveration caused by inhibition failure? — Evidence from preschool children's inferences about word meanings. Journal of Experimental Child Psychology 86, 194–222.

Diamond, A., Prevor, M.B., Callender, G., Druin, D.P., 1997. Prefrontal cortex cognitive deficits in children treated early and continuously for PKU. Monographs of the Society for Research in Child Development 62, 1–207.

Dien, J., Spencer, K.M., Donchin, E., 2004. Parsing the late positive complex: mental chronometry and the ERP components that inhabit the neighborhood of the P300. Psychophysiology 41, 665–678.

Dimoska, A., Johnstone, S.J., Barry, R.J., 2006. The auditory-evoked N2 and P3 components in the stop-signal task: indices of inhibition, response conflict or error-detection? Brain and Cognition 62, 98–112.

Donkers, F.C.L., Van Boxtel, G.J.M., 2004. The N2 in go/no-go tasks reflects conflict monitoring not response inhibition. Brain and Cognition 56, 165–176.

Duncan-Johnson, C.C., Kopell, B.S., 1981. The Stroop effect: brain potentials localize the source of interference. Science 214, 938–940.

Durston, S., Casey, B.J., 2006. What have we learned about cognitive development from neuroimaging? Neuropsychologia 44, 2149–2157.

Epsy, K.A., 1997. The shape school: assessing executive function in preschool children. Developmental Neuropsychology 13, 495–499.

Eriksen, B.A., Eriksen, C.W., 1974. Effects of noise letters upon the identification of a target letter in a nonsearch task. Perception & Psychophysics 16, 143–149.

Espy, K.A., McDiarmid, M.M., Cwik, M.F., Stalets, M.M., Hamby, A., Senn, T.E., 2004. The contribution of executive functions to emergent mathematic skills in preschool children. Developmental Neuropsychology 26, 465–486.

Fagot, D., Dirk, J., Ghisletta, P., de Ribaupierre, A., 2009. Adults' versus children's performance on the Stroop task: insights from ex-Gaussian analysis. Swiss Journal of Psychology 68, 17–24.

Faust, M.E., Balota, D.A., Spieler, D.H., Ferraro, F.R., 1999. Individual differences in information-processing rate and amount: implications for group differences in response latency. Psychological Bulletin 125, 777–799.

Fernandez-Duque, D., Baird, J.A., Posner, M.I., 2000. Executive attention and metacognitive regulation. Consciousness and Cognition 9, 288–307.

Flavell, J.H., Beach, D.R., Chinsky, J.M., 1966. Spontaneous verbal rehearsal in a memory task as a function of age. Child Development 37, 283–299.

Flavell, J.H., Green, F.L., Flavell, E.R., Grossman, J.B., 1997. The development of children's knowledge about inner speech. Child Development 68, 39–47.

Freitas, A.L., Banai, R., Clark, S.L., 2009. When cognitive control is calibrated: event-related potential correlates of adapting to information-processing conflict despite erroneous response preparation. Psychophysiology 46, 1226–1233.

Garon, N., Bryson, S.E., Smith, I.M., 2008. Executive function in preschoolers: a review using an integrative framework. Psychological Bulletin 134, 31–60.

Gathercole, S.E., Lamont, E., Alloway, T.P., 2006. Working memory in the classroom. In: Pickering, S. (Ed.), Working Memory and Education. Elsevier Press, UK, pp. 219–240.

Gonzalez, C., Fuentes, L.J., Carranza, J.A., Estevez, A.F., 2001. Temperament and attention in the self-regulation of 7-year-old children. Personality and Individual Differences 30, 931–946.

Gratton, G., Coles, M.G.H., Sirevaag, E.J., Eriksen, C.W., Donchin, E., 1988. Pre- and post-stimulus activation of response channels: a psychophysiological analysis. Journal of Experimental Psychology: Human Perception and Performance 14, 331–344.

Hacker, D.J., Dunlosky, J., Graesser, A.C., 1998. Metacognition In Educational Theory and Practice. Lawrence Erlbaum, Mahwah, NJ.

Happe, F., Booth, R., Charlton, R., Hughes, C., 2006. Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: examining profiles across domains and ages. Brain and Cognition 61, 25–39.

Hsieh, S., Yu, Y.T., 2003. Switching between simple response-sets: inferences from the lateralized readiness potential. Cognitive Brain Research 17, 228–237.

Hughes, C., Ensor, R., Wilson, A., Graham, A., 2010. Tracking executive function across the transition to school: a latent variable approach. Developmental Neuropsychology 35, 20–36.

Huizinga, M., Dolan, C.V., van der Molen, M.W., 2006. Age-related change in executive function: developmental trends and a latent variable analysis. Neuropsychologia 44, 2017–2036.

Ilan, A.B., Polich, J., 1999. P300 and response time from a manual Stroop task. Clinical Neuropsychology 110, 367–373.

Jaskowski, P., Kurczewska, M., Nowik, A., van der Lubbe, R.H.J., Verleger, R., 2007. Locus of the intensity effect in simple reaction time tasks. Perception & Psychophysics 69, 1334–1343.

Jaskowski, P., Szumska, I., Sasin, E., 2009. Functional locus of intensity effects in choice reaction time tasks. Journal of Psychophysiology 23, 126–134.

Jensen, A.R., 1992. The importance of intraindividual variation in reaction time. Personality and Individual Differences 13, 869–881.

Johnson, J., Im-Bolter, N., Pascual-Leone, J., 2003. Development of mental attention in gifted and mainstream children: the role of mental capacity, inhibition, and speed of processing. Child Development 74, 1594–1614.

Jongen, E.M.M., Jonkman, L.M., 2008. The developmental pattern of stimulus and response interference in a color–object Stroop task: an ERP study. BMC Neuroscience 9, 1–24.

Kao, C.S., Dietrich, R., Sommer, W., 2009. Processing interrogative sentence mood at the semantic–syntactic interface: an electrophysiological research in Chinese, German, and Polish. PLoS ONE 5, 11.

Keeney, T.J., Cannizzo, S.R., Flavell, J.H., 1967. Spontaneous and induced verbal rehearsal in a recall task. Child Development 38, 953–966.

Kutas, M., McCarthy, G., Donchin, E., 1977. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation in time. Science 197, 792–795.

Lacouture, Y., Cousineau, D., 2008. How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times. Tutorials in Quantitative Methods for Psychology 4, 35–45.

Lamm, C., Zelazo, P.D., Lewis, M.D., 2006. Neural correlates of cognitive control in childhood and adolescence: disentangling the contributions of age and executive function. Neuropsychologia 44, 2139–2148.

Lansbergen, M., Kenemans, J.L., 2008. Stroop interference and the timing of selective response activation. Clinical Neuropsychology 119, 2247–2254.

Leth-Steensen, C., Elbaz, Z.K., Douglas, V.I., 2000. Mean response times, variability, and skew in the responding of ADHD children: a response time distributional approach. Acta Psychologica 104, 167–190.

Logan, G.D., 1994. On the ability to inhibit thought and action: a users' guide to the stop signal paradigm. In: Dagenbach, D., Carr, T.H. (Eds.), Inhibitory Processes in Attention, Memory, and Language. Academic Press, San Diego, pp. 189–239.

Luria, A.R., 1973. The Working Brain: An Introduction to Neuropsychology. Penguin Books Ltd., Middlesex.

Marsh, R., Zhu, H., Schultz, R.T., Quackenbush, G., Royal, J., Skudlarski, P., Peterson, B.S., 2006. A developmental fMRI study of self-regulatory control. Human Brain Mapping 27, 848–863.

Martens, U., Leuthold, H., Schweinberger, S.R., 2010. Parallel processing in face perception. Journal of Experimental Psychology: Human Perception and Performance 36, 103–121.

Martin-Rhee, M.M., Bialystok, E., 2008. The development of two types of inhibitory control in monolingual and bilingual children. Bilingualism: Language and Cognition 11, 81–93.

McAuley, T., Yap, M., Christ, S.E., White, D.A., 2006. Revisiting inhibitory control across the life span: insights from the ex-Gaussian distribution. Developmental Neuropsychology 29, 447–458.

Melcher, T., Gruber, O., 2009. Decomposing interference during Stroop performance into different conflict factors: an event-related fMRI study. Cortex 45, 189–200.

Meltzer, L., 2007. Executive Function in Education: From Theory to Practice. Guilford Press, London.

Mewhort, D.J.K., Braun, J.G., Heathcote, A., 1992. Response time distributions and the Stroop task: a test of the Cohen, Dunbar, and McClelland (1990) model. Journal of Experimental Psychology: Human Perception and Performance 18, 872–882.

Miller, J.O., Patterson, T., Ulrich, R., 1998. Jackknife-based method for measuring LRP onset latency differences. Psychophysiology 35, 99–115.

Miller, J., Ulrich, R., Schwarz, W., 2009. Why jackknifing yields good latency estimates. Psychophysiology 46, 300–312.

Mischel, W., Ebbesen, E.B., Zeiss, A., 1972. Cognitive and attentional mechanisms in delay of gratification. Journal of Personality and Social Psychology 21, 214–218.

Mordkoff, J.T., Gianoros, P.J., 2000. Detecting the onset of the lateralized readiness potential: a comparison of available methods and procedures. Psychophysiology 37, 347–360.

Myerson, J., Robertson, S., Hale, S., 2007. Aging and intraindividual variability in performance: analyses of response time distributions. Journal of the Experimental Analysis of Behavior 88, 319–337.

Nigg, J.T., 2001. Is ADHD a disinhibitory disorder? Psychological Bulletin 127, 571–598.

Osman, A., Moore, C.M., 1993. The locus of dual-task interference: psychological refractory effects on movement-related brain potentials. Journal of Experimental Psychology: Human Perception and Performance 19, 1292–1312.

Ozonoff, S., Jensen, J., 1999. Brief report: specific executive function profiles in three neurodevelopmental disorders. Journal of Autism and Developmental Disorders 29, 171–177.

Pfefferbaum, A., Ford, J.M., Weller, B.J., Kopell, B.S., 1985. ERPs to response production and inhibition. Electroencephalography and Clinical Neurophysiology 60, 423–434.

Piaget, J., 1954. Construction of Reality in the Child. Routledge & Kegan Paul, London.

Pritchard, V.E., Neumann, E., 2009. Avoiding the potential pitfalls of using negative priming tasks in developmental studies: assessing inhibitory control in children, adolescents, and adults. Developmental Psychology 45, 272–283.

Ramautar, J.R., Kok, A., Ridderinkhof, K.R., 2004. Effects of stop-signal probability in the stop-signal paradigm: the N2/P3 complex further validated. Brain and Cognition 56, 234–252.

Reed, M., Pien, D.L., Rothbart, M.K., 1984. Inhibitory self-control in preschool children. Merrill-Palmer Quarterly 30, 131–147.

Rhoades, B.L., Greenberg, M.T., Domitrovich, C.E., 2009. The contribution of inhibitory control to preschoolers' social–emotional competence. Journal of Applied Developmental Psychology 30, 310–320.

Ridderinkhof, K.R., van der Molen, M.W., 1995. A psychophysiological analysis of developmental differences in the ability to resist interference. Child Development 66, 1040–1056.

Roberts, R.J., Pennington, B.F., 1996. An interactive framework for examining prefrontal cognitive processes. Developmental Neuropsychology 12, 105–126.

Roncadin, C., Pascual-Leone, J., Rich, J.B., Dennis, M., 2007. Developmental relations between working memory and inhibitory control. Journal of the International Neuropsychological Society 13, 59–67.

Rubia, K., Smith, A.B., Taylor, E., Brammer, M., 2007. Linear age-correlated functional development of right inferior fronto-striato-cerebellar networks during response inhibition and anterior cingulate during error-related processes. Human Brain Mapping 28, 1163–1177.

Scheibe, C., Schubert, R., Sommer, W., Heekeren, H.R., 2009. Electrophysiological evidence for the effect of prior probability on response preparation. Psychophysiology 46, 758–770.

Schneider, W., Pressley, M., 1997. Memory Development Between Two and Twenty, 2nd Ed. Erlbaum, Mahwah, N. J.

Schroeter, H., Leuthold, H., 2009. Motor programming of rapid finger sequences: inferences from movement-related brain potentials. Psychophysiology 46, 388–401.

Schroeter, M.L., Zysset, S., Wahl, M., Yves von Cramon, D., 2004. Prefrontal activation due to Stroop interference increases during development—an event-related fNIRS study. Neuroimage 23, 1317–1325.

Shimamura, A.P., 2000. Toward a cognitive neuroscience of metacognition. Consciousness and Cognition 9, 313–323.

Smid, H.G.O.M., Mulder, G., Mulder, L.J.M., 1990. Selective response activation can begin before stimulus recognition is complete: a psycho-physiological and error analysis of continuous flow. Acta Psychologica 74, 169–201.

Smith, J.L., Johnstone, S.J., Barry, R.J., 2007. Response priming in the Go/NoGo task: the N2 reflects neither inhibition nor conflict. Clinical Neuropsychology 118, 343–355.

Solomon, M., Ozonoff, S., Cummings, N., Carter, C.S., 2008. Cognitive control in autism spectrum disorders. International Journal of Developmental Neuroscience 26, 239–247.

Span, M.M., Ridderinkhof, K.R., van der Molen, M.W., 2004. Age-related changes in the efficiency of cognitive processing across the life span. Acta Psychologica 117, 155–183.

St Clair-Thompson, H.L., Gathercole, S.E., 2006. Executive functions and achievements in school: shifting, updating, inhibition, and working memory. The Quarterly Journal of Experimental Psychology 59, 745–759.

Stahl, J., Gibbons, H., Miller, J., 2010. Modeling single-trial LRP waveforms using gamma functions. Psychophysiology 47, 43–56.

Steinhauser, M., Huebner, R., Druey, M., 2009. Adaptive control of response preparedness in task switching. Neuropsychologia 47, 1826–1835.

Stroop, J.R., 1935. Studies of interference in serial verbal reactions. Journal of Experimental Psychology 18, 643–662.

Swanson, H.L., 2006. Cognitive processes that underlie mathematical precociousness in young children. Journal of Experimental Child Psychology 93, 239–264.

Szucs, D., Soltesz, F., 2007. Event-related potentials dissociate facilitation and interference effects in the numerical Stroop paradigm. Neuropsychologia 45, 3190–3202.

Szucs, D., Soltesz, F., Jarmi, E., Csepe, V., 2007. The speed of magnitude processing and executive functions in controlled and automatic number comparison in children: an electro-encephalography study. Behavioral and Brain Function 3, 1–20.

Szucs, D., Soltesz, F., Bryce, D., Whitebread, D., 2009a. Real-time tracking of motor response activation and response competition in a Stroop task in young children: a lateralized readiness potential study. Journal of Cognitive Neuroscience 21, 2195–2206.

Szucs, D., Soltesz, F., White, S., 2009b. Motor conflict in Stroop tasks: direct evidence from single-trial electro-myography and electro-encephalography. Neuroimage 47, 1960–1973.

Tsujimoto, S., Kuwajima, M., Sawaguchi, T., 2007. Developmental fractionation of working memory and response inhibition during childhood. Experimental Psychology 54, 30–37.

Ulrich, R., Miller, J.O., 2001. Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. Psychophysiology 38, 816–827.

van der Molen, M.W., 2000. Developmental changes in inhibitory processing: evidence from psychophysiological measures. Biological Psychology 54, 207–239.

Veenman, M.V.J., Van Hout-Wolters, B.H.A.M., Afflerbach, P., 2006. Metacognition and learning: conceptual and methodological considerations. Metacognition and Learning 1, 3–14.

Wild-Wall, N., Falkenstein, M., Hohnsbein, J., 2008. Flanker interference in young and older participants as reflected in event-related potentials. Brain Research 1211, 72–84.

Wilson, S.P., Kipp, K., 1998. The development of efficient inhibition: evidence from directed-forgetting tasks. Developmental Review 18, 86–123.