



HAL
open science

Conflict processing in kindergarten children: New evidence from distribution analyses reveals the dynamics of incorrect response activation and suppression

Solène Ambrosi, Mathieu Servant, Agnès Blaye, Boris Burle

► To cite this version:

Solène Ambrosi, Mathieu Servant, Agnès Blaye, Boris Burle. Conflict processing in kindergarten children: New evidence from distribution analyses reveals the dynamics of incorrect response activation and suppression. *Journal of Experimental Child Psychology*, 2019, 177, pp.36 - 52. 10.1016/j.jecp.2018.06.006 . hal-01862984

HAL Id: hal-01862984

<https://hal.science/hal-01862984>

Submitted on 29 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Conflict processing in kindergarten children: New evidence from distribution analyses reveals the dynamics of incorrect response activation and suppression ☆

Solène Ambrosi^a, Mathieu Servant^{b,c}, Agnès Blaye^{a,1}, Boris Burle^{b,1,*}

^aLaboratoire de Psychologie Cognitive, Aix-Marseille Univ, CNRS, LPC, UMR 7290, 13331 Marseille, France

^bLaboratoire de Neurosciences Cognitives, Aix-Marseille Univ, CNRS, LNC, UMR 7291, 13331 Marseille, France

^cDepartment of Psychology, Vanderbilt University, Nashville, TN 37203, USA

Abstract

The development of cognitive control is known to follow a long and protracted development. However, whether interference effect in conflict tasks in children would entail the same core processes as in adults, namely an automatic activation of incorrect response and its subsequent suppression, remains an open question. We applied distributional analyses to reaction times and accuracy of 5- to 6-year-old children performing three conflict tasks (flanker, Simon and Stroop) in a within-participants design. This revealed both strong commonalities and differences between children and adults. As in adults, fast responses were more error-prone than slow ones on incompatible trials, indicating a fast “automatic” activation of the incorrect response. In addition, the strength of this activation differed across tasks, following a pattern similar to adults. Moreover, modeling the data with a Drift Diffusion Model adapted for Conflict tasks allowed to better assess the origin of the typical slowing down observed in children. Besides showing that advanced distribution analyses can be successfully applied to children, the present results support the notion that interference effects in 5- to 6-year-olds are driven by mechanisms very similar to the ones at play in adults but with different time courses.

Introduction

Cognitive control refers to a set of higher cognitive functions that regulate behavior to ensure goal attainment. Recent studies have revealed that the efficiency of cognitive control in childhood can predict individual differences in many domains of cognitive development such as early language ability or theory of mind, but also in academic achievement such as mathematics (Bull & Lee, 2014) and literacy (Colé, Duncan, & Blaye, 2014; see Diamond, 2013, 2014 for reviews). More generally, cognitive control proved to be more strongly correlated with school readiness than is IQ (e.g., Blair & Razza, 2007) and its efficiency in childhood revealed one of the best predictors of health and employment at adult age (Daly, Delaney, Egan, & Baumeister, 2015; Moffitt

☆This is the author’s (postprint) version of the article published in *Journal of Experimental Child Psychology*, 177 (2019) 36–52. DOI:10.1016/j.jecp.2018.06.006, Accepted 31 May 2018

*corresponding author: boris.burle@univ-amu.fr

¹Shared last authorship

et al., 2011). The broad relevance of cognitive control in children stresses the importance of understanding its development. Cognitive control in adults is often investigated through so-called “conflict tasks”, such as the Stroop (Stroop, 1935), flanker (Eriksen & Eriksen, 1974) or Simon (Simon, 1990) tasks. While those three tasks differ in some respects (see Kornblum, Hasbroucq, & Osman, 1990), they share a common structure: stimuli are composed of two dimensions; one being relevant for the task and determining the correct response, while the second one, irrelevant for the task, shares common features with the stimulus or response sets. In the Stroop task, participants are requested to name the color of a written word. The word can be compatible with its color (for example ‘red’ written in red) or incompatible (‘green’ written in red). In a standard version of the flanker task, participants must issue a right- or left-hand response as a function of the nature of a central letter (for example ‘H’ or ‘S’) flanked by distractors that can be a replication of the target (for example ‘HHH’, compatible trials) or a replication of the alternative target (‘SHS’, incompatible trials). In the Simon task, a lateralized response is required to a non-spatial dimension (for example the color) of stimuli that are presented either on the same side as the requested response (compatible trials) or on the opposite side (incompatible trials). In all those tasks, one usually assumes that the stimuli are processed along two parallel routes: a “fast” one, processing the irrelevant dimension in an automatic way, and a slower one, processing the relevant dimension in a more controlled way. On incompatible trials, the irrelevant dimension tends to activate the incorrect response which then needs to be suppressed to produce the correct response, whereas on compatible trials, both the relevant or irrelevant dimensions lead to the same correct response. Worse performance (longer RT and higher error rate) on incompatible trials as compared to compatible ones indexes the interference effect induced by the irrelevant dimension on the processing of the relevant one. Children-adapted versions of those tasks have been used to assess interference processing, even in very young children (around 3-4 years of age) (e.g., Davidson, Amso, Anderson, & Diamond, 2006; Gerstadt, Hong, & Diamond, 1994; Ikeda, Okuzumi, & Kokubun, 2013, 2014; Prevor & Diamond, 2005; Rueda et al., 2004; Wright, Waterman, Prescott, & Murdoch-Eaton, 2003). Developmental studies suggest a long and protracted development of cognitive control (e.g., Cao et al., 2013; Luna & Sweeney, 2004; Macdonald et al., 2014; Prevor & Diamond, 2005; Ridderinkhof, van der Molen, Band, & Bashore, 1997) likely sustained by a late maturation of neural networks engaged in conflict resolution (see e.g., Abundis-Gutiérrez, Checa, Castellanos, & Rosario Rueda, 2014; Durston & Casey, 2006; Rueda et al., 2004). These studies globally show decreased interference effects with age. However, interference effects have been evaluated with summary statistics, blind to the dynamics of the interference effects and hence to the underlying processes. Indeed, in the last few years, there has been growing evidences that the mean RT and error rate provides incomplete information. More specifically, RT distributions are not normally distributed but instead have a characteristic heavy right tail. Analyzing the shape of the RT distributions can provide essential information. This can be done in different ways, briefly summarized below. A first type of analysis consists in fitting statistical, non-gaussian, distributions to the acquired data. Different theoretical distributions have been used such as, for example, the log-normal (Ulrich & Miller, 1993) the Ex-Gaussian (Hohle, 1965, Burbeck & Luce, 1982), the gamma (McGill & Gibbon, 1965) or the Weibull distributions (Logan, 1988). After having fitted the chosen distribution, one can then average the parameters across participants and create an “average” distribution representative of all the participant’s one. Softwares to fit such distributions have been made available to the community (see e.g. QMPE, available at <http://www.newcl.org/node/8>). This descriptive approach proved useful in some circumstances. For example, Leth-Steensen, King Elbaz, and Douglas (2000) using ex-gaussian fitting in Attention Deficits with-or-without Hyperactivity Disorders

(AD/HD) children evidenced excessive long RTs in this population in warned reaction time task. However, although those distributions can provide useful statistical descriptions of the data, inferring the dynamics of the underlying processes from the parameters of the fitted distribution is generally not straightforward and must be done with caution (see Matzke and Wagenmakers, 2009 for a discussion).

Another, distribution free, approach (that will be used in the present study) is to construct an average distribution, representative of the individual ones, without assuming any a priori given shape. The general underlying idea is to take some common characteristic points describing each individual RT distributions and to average these characteristic points across participants to reconstruct a distribution whose properties are central tendencies of the individual distributions. To do so, data are classified into bins of same size. From there, different measures can be taken. One can take the boundaries of the created bins, referred to as quantiles. For some families of distribution (for example the so-called “location-scale” distributions), it can be shown that quantiles averaging provides an unbiased averaged distribution (Jiang, Rouder & Speckman, 2004). An alternative popular technique is the “Vincentization” (Vincent, 1912, Ratcliff, 1979) which bins the distribution into classes of equal sizes and takes the mean of each class (see methods for more details). These two measures provide very similar estimations of the distribution shape. This approach has proved extremely useful to reveal automatic response activations and their subsequent suppression in conflict tasks (Burle et al., 2002, Ridderinkhof, 2002, Pratte, Rouder, Morey & Feng, 2010, Burle, Spieser, Servant, & Hasbroucq, 2014)

Finally, one can use formal decision-making models to account for RT distributions: based on a set of parameters supposedly modeling the different processes underlying the reaction time, those models generate RT distributions that can be compared to the empirical ones. One class of models, the “accumulation to bound” type, revealed very successful in accounting for RT distribution shapes (Ratcliff, 1978, Ratcliff & McKoon, 2008). Although different types of models have been proposed, they all share the assumption that evidence from the environment is accumulated at a given rate (μ) until a predefined threshold (b) is reached and a response is given. An additional non-decisional time – Ter – is also usually added to account for the whole RT. Importantly, the best parameters accounting for a given data set can be inferred through fitting procedures, providing information about the processes at play in the reaction period. For example, differences in the accumulation rate parameter (μ) across experimental conditions is classically interpreted as reflecting a change in processing speed (see e.g. Palmer et al, 2005), a change in the response threshold (b) will index a variation of response caution (see e.g. Bogacz et al., 2009), and a variation of Ter indexes an effect on the non-decision component (e.g. Ratcliff & McKoon, 2015). This approach will also be used in the present study.

In the context of conflict tasks, distribution analyses in adults have revealed that the interference effect (in terms of RT and error rate) is not constant across the RT distribution (see below for more details). Figure 1 shows different toy examples to illustrate different dynamics of interference effect despite identical mean effects. The left panel depicts three examples where the dynamics of the interference effect varies as a function of RT length, despite equal mean interference effects (equal to 30 ms on these examples). These dynamics, often depicted as “delta-plots”, are derived from the vincentized cumulative density functions of RT in incompatible and compatible trials, by plotting the interference (incompatible – compatible RTs) as function of the response latency (see below for more details). The right panel of Figure 1 shows three different repartitions of errors as a function of RT, for the same mean error rate (10% errors). This analysis is often referred to as “Conditional Accuracy Functions” (CAF, Gratton, Coles, Sirevaag, Eriksen, & Donchin, 1988, Lappin & Disch, 1972). As can be seen from this figure, the mean

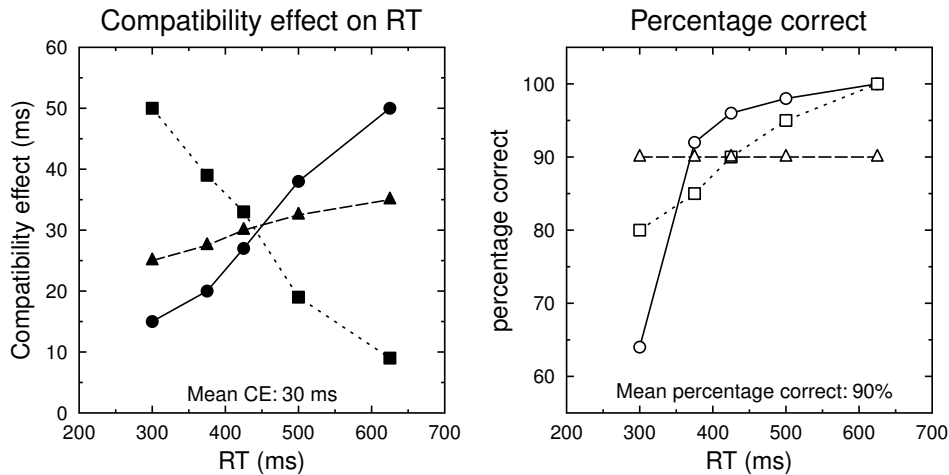


Figure 1: Schematic representation of different dynamics leading to identical mean effects. The left panel show three examples in which the mean compatibility effect on RT is equal to 30 ms, but the evolution of this effect across RTs is different. The right panel illustrates three examples of evolution of proportion of correct response as a function of response speed in which the global percentage of correct response is identical (90% correct).

error rate can mask important differences in the dynamics of error production.

As evidenced in the adult literature, such differences in dynamics provide essential information about the underlying processes (Ridderinkhof, 2002, Pratte, et al., 2010, Ulrich, Schröter, Leuthold, & Birngruber, 2015). Indeed, the three main conflict tasks (Stroop, Flanker and Simon) differ in the dynamics of the interference effects across RTs (Burle, et al., 2014; Pratte et al., 2010; Servant, Montagnini, & Burle, 2014). The interference effect on RT increases as RTs get longer in the flanker and Stroop tasks but decreases and even often disappears for long RTs in the Simon task. From a functional point of view, these between-tasks differences have been linked to changes in the strength of the incorrect-response suppression (Burle et al., 2002, 2014, Ridderinkhof, 2002, Ridderinkhof, Van den Wildenberg, Wijnen, & Burle, 2004). Between-task commonalities and differences in CAF have been less systematically studied. A common feature is that on incompatible trials, errors tend to largely concentrate on the fastest RTs, and error rate decreases as RTs lengthen (Gratton, et al., 1988; van den Wildenberg et al., 2010). CAF on compatible trials tend to be much higher and much flatter than on incompatible ones. The difference in error rate between incompatible and compatible trials is hence essentially due to the fastest responses. This initial drop in accuracy for fast incompatible responses has been proposed to be due to the automatic response activation by the irrelevant dimension of the stimulus (Gratton, Cole, & Donchin, 1992; van den Wildenberg et al., 2010).

Very few studies have looked at RT distributions in children. First evidence of the usefulness of delta plot analyses in comparing cognitive control between different children populations comes from Ridderinkhof, Scheres, Oosterlaan, & Sergeant (2005). They compared the distributions of RTs on an arrow version of the flanker task in AD/HD children and in typically-developing children (mean age: 9.6, age range 6-12). Such analysis revealed steeper delta-plots in unmedicated AD/HD but medication led to delta-values much more similar to normally developing children. This was interpreted as reflecting suppression deficits that can be overcome by

medication. Bub, Masson, & Lalonde (2006) compared two groups of elementary school children below and above 9 years of age (age range 7-11) on a Stroop task. The delta plot analyses for older children (9- to 11-year-olds) were consistent with the generally observed adult pattern: the interference effect increased with RT. For younger children (7- to 9-year-olds), however, the Stroop effect diminished for the longest response latencies. According to the authors, this unexpected pattern suggests that younger children applied more suppression of word reading than did the older children. Nevertheless, this pattern of results might be accounted for by more variability in reading fluency in the younger age group that would have particularly affected the last segments of the delta function time (that have been shown to be very sensitive to noise; e.g., Schwarz & Miller, 2012). Concerning the Simon task, Iani, Stella, & Rubichi (2014) showed that the Simon effect was present for fast responses and decreased as RT increased in 1st and 2nd graders (6 and 7 years old), as for adults. This decrease appeared weaker in 1st graders where the Simon effect remained significant for all bins whereas it was no more significant for the last bin in 2nd graders, leaving open the question of the shapes of the distributions at an earlier age. To the best of our knowledge, distribution analyses have never been applied under the age of 6, and no systematic comparisons of the three main conflict tasks (flanker, Stroop, and Simon) have been done with these tools.

As for RT distributions, very few studies have used CAF in children (Bub et al., 2006; Ridderinkhof et al., 2005; Stins, Polderman, Boomsma, & de Geus, 2007). Stins et al. (2007) analyzed CAF's in 12-year-old children on an arrow-flanker and a Simon tasks (see also Ridderinkhof et al., 2005 for similar results in 6- to 12-year-olds on an arrow-flanker task). They found a pattern of results similar to the one observed in adults: congruence effects were more pronounced for fast responses in both tasks. The CAFs obtained by Bub et al. (2006) in 7- to 11-year-olds performing a Stroop task also showed the typical decreased accuracy for faster RTs on incompatible trials. However, the distributions also revealed decreased performance for the longest RTs (for both compatible and incompatible trials), raising further doubt on Bub et al.'s (2006) interpretation of enhanced suppression in young children based on the delta plots. Altogether, these studies revealed that children present a bias toward responding to the irrelevant dimension of the stimulus that diminishes as RT increased, as if they were more likely to avoid errors when they take more time to respond. This parallels data obtained in preschoolers (range from 3 to 6, depending on studies) showing that incompatible settings leading to longer RTs are generally associated with better accuracy (Gerstadt et al., 1994, Diamond, Kirkham, & Amso, 2002; Ling, Wong, & Diamond, 2016; Simpson et al., 2012; Simpson & Riggs, 2007). The first goal of the present study is hence to characterize the dynamics of interference effect, on RT and error rate, in kindergarten children (5-6 years old), in the three main conflict tasks described above, namely the flanker, Simon and Stroop tasks.

Besides this distribution-free approach, we fit a model of conflict tasks recently developed in adults (Ulrich et al., 2015) to go one step further in characterizing the dynamics of the processes engaged by children. This model hereafter referred to as Diffusion Model for Conflict tasks, DMC), allows to quantify, among other psychological components, response caution and the quality of perceptual processing and to parametrically estimate the time course of the automatic response activation. This model, which has recently received support thanks to neurophysiological measurements (Servant, White, Montagnini, & Burle, 2016), is currently the only one being able to account for the differences in the shape of RT distributions across conflict tasks. Fitting this model to kindergarten children's performance will allow to assess whether the same underlying processes are at play in young children and adults, and to unveil the dynamics of processes underpinning the overall mean interference effects.

In the present study, we re-analyzed data collected and reported in Ambrosi, Lemaire, & Blaye (2016). These authors investigated interference processing in 5- and 6-year-old children using a Stroop, a Simon and a flanker task in a within-subject design. Results showed interference effects both on latencies and error rates in the three tasks albeit with a weaker accuracy interference effect in the flanker- than in the two other tasks. The opportunity to combine considerations from delta-plots, conditional accuracy functions, and fits of a formal model in such a young age range across three conflict tasks offers a new approach to assess the processes underlying the engagement of control processes in kindergartners, an age at which cognitive control is still partially immature.

Method

Participants

Fifty-three children completed the three tasks in the study reported by Ambrosi et al. (2016): 29 five-year-olds (14 girls; mean age= 67 months, SD=3 months) and 32 six-year-olds (16 girls; Mean = 79 months, SD=2 months).

Materials and Procedure

Material and procedures are described in detail in Ambrosi et al. (2016) A brief overview of critical details is recounted below. The three conflict tasks, that are the focus of the present study, were initially included in a larger battery of cognitive tasks. The fatigability of young children and their overall limited attentional resources required to split the experiment in two sessions. Children were individually tested in two sessions, taking first the flanker task and, one week later, the Simon followed by the Stroop tasks. Based on a pilot experiment, this fixed order was selected as a plausible order of ascending difficulty of the tasks². The tasks were administered on a laptop computer (15-inch monitor Dell Latitude E6410ATG) running E-Prime® software. Each task began after a short training phase. It involved six blocks of 17 trials, with a break in-between each block. Each task contained four different stimuli, two compatible and two incompatible. In the flanker task (adapted from Rueda et al., 2004), stimuli consisted of a row of five yellow fish. Participants must issue a right or left response as a function of the direction of the central fish flanked by distractors that swim in the same direction (compatible trials) or in the opposite direction (incompatible trials). In the Simon task (adapted from Simon & Berbaum, 1988) stimuli were a blue square and a red circle. A lateralized response was required to the shape/color dimension (i.e., left for blue square and right for red circle) of stimuli that were presented on the same side as the requested response (compatible trials) or on the opposite side (incompatible trials). In the Color-Object Stroop task (adapted from Archibald & Kerns, 1999), stimuli consisted of a colored line drawing of a carrot or a salad presented in the center of the screen. Children were requested to press the response key corresponding to the stimulus canonical color (i.e., orange for carrot and green for salad). The color displayed on the screen and the canonical color

²Although motivated by practical considerations, the use of the fixed order of tasks did not seem to have produced any carry-over effects. Indeed, such carry-over effects, when they exist, are normally expressed by a reduced interference effect after first performing another conflict task. The analyses run by Ambrosi et al. (2016) on the magnitude of the interference effect did reveal between-tasks differences only when computed on accuracy but the weaker effect was observed on the flanker task presented on session 1 and there were no differences between the Simon and the Stroop effects ran in session 2.

matched in compatible trials but differed in incompatible trials. A total of 48 trials per trial type and per task were hence created. Each trial began with a fixation cross (+) displayed for 850 ms in the center of the screen. Then, the stimulus was displayed and remained on the screen until a response key was pressed. Children had to respond on the bases of the values of the relevant dimensions by pressing either the left (s) or the right (l) response key on an AZERTY keyboard.

Data processing

Children's RT data are known to be variable and very prone to lapses of attention, leading to excessively long RTs on some trials. Data were hence pruned by removing all RTs larger than 3 standard deviations from participants' mean RT for each trial type (0.8%). All reported results are based on those cleaned data. Mean RT and accuracy have already been described in a previous report (Ambrosi et al., 2016). We will here focus on distribution analyses.

RT Distributions analysis

To perform distributions analysis, the "Vincent averaging" or "vincentization" technique was used (Ratcliff, 1979; Vincent, 1912). Basically, the RTs were sorted in ascending order, binned in classes of equal size (same number of trials), and the mean of each bin was computed. This was done for each participant, task (flanker, Simon and Stroop), and trial type (compatible and incompatible) separately, including only correct trials preceded by correct trials. For each task and trial type, the mean of each bin was averaged across participants to obtain an average distribution of RT representative of the individual ones. "Delta-plots" were then computed from these averaged distributions: This was done by plotting the difference between incompatible and compatible RTs in each bin against the mean RT of the same two values.

Conditional accuracy functions

Conditional accuracy functions are constructed with the same logic, but differ on two main points. First, vincentization is performed on correct and erroneous trials put together. Second, for each bin, the probability of correct responses is computed, and plotted against the mean RT of the bin. Plotting all bins provides the probability of a correct response as a function of response speed. This was also done for each participant, each compatibility condition and each task separately, and then averaged across participants.

Statistical analysis

A first set of ANOVAs were performed with Task (flanker, Simon, Stroop), Trial type (compatible vs. incompatible trials) and Quantile (from 1 to 5) as within-participant factors and Age (5 vs. 6 years of age) as a between-participants factor (although, based on Ambrosi et al. 2016 results, we do not expect any effect of age). The analyses of variance were conducted with the R software (R Core Team, 2015, version 3.2.3 for linux) using the ezANOVA function of the 'ez' package (version 4.4-0, Lawrence, 2016). Mauchly's test (Mauchly, 1940) indicated that the assumption of sphericity had been violated for all the comparisons that included the factors Quantile, and Task. Degrees of freedom were therefore corrected using Greenhouse-Geisser estimates of sphericity (ϵ) (Greenhouse & Geisser, 1959); ϵ and p -value after correction are reported for significant effect along with Partial Eta squared (η_p^2) values that index effect size measures (J. Cohen, 1973).

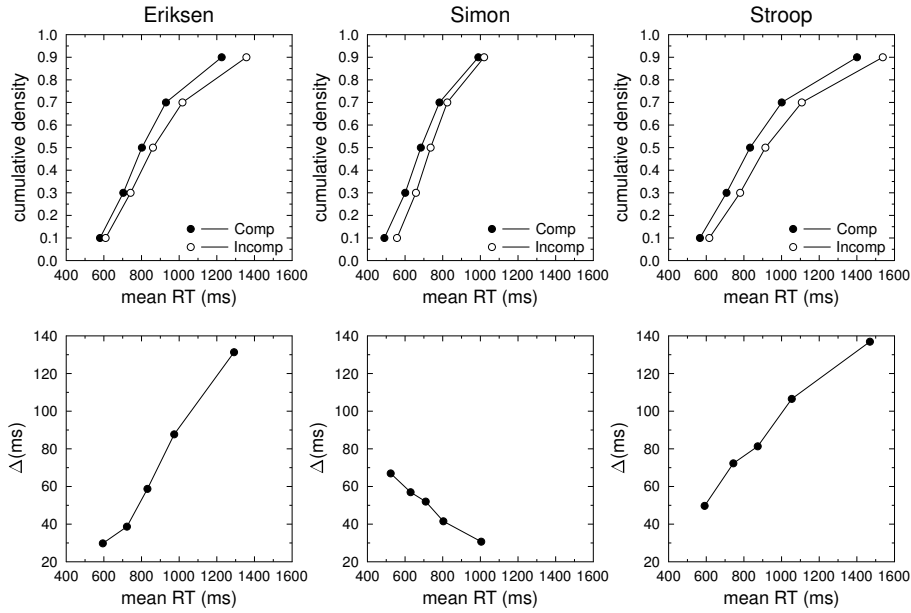


Figure 2: Cumulative density function and delta plots. The upper panels shows distribution for compatible (black circle) and incompatible (white circle) condition in the flanker task (left), in the Simon task (middle), and in Stroop task (right) as a function of response times. The lower panels show the corresponding delta plots.

Results

RT Distribution analysis

Analyses on RTs revealed no effect of Age, neither as a main effect ($F(1, 51) = 1.09, p = .301, MSE = 701402$), nor as an interaction term (all F s < 1 , except Age \times Task \times Compatibility, $F(2, 102) = 1.67, MSE = 20392, p = .193$). Figure 2 shows the cumulative density functions (upper row) for each task and each trial type, along with the corresponding delta function (lower row)³. Main effects of Task ($F(2, 102) = 26.06, \epsilon = .941, MSE = 226490, p < .001, \eta_p^2 = .34$), Trial type ($F(1, 51) = 76.92, MSE = 24854, p < .001, \eta_p^2 = .60$), and a trivial effect of Quantile ($F(4, 204) = 482.86, \epsilon = .267, MSE = 45499, p < .001, \eta_p^2 = .90$) were observed. The factors Task and Quantile interact significantly ($F(8, 408) = 30.47, \epsilon = .267, MSE = 20392, p < .001, \eta_p^2 = .37$) indicating that the shapes of the RT distributions differ between tasks. Trial type and Quantile also interact ($F(4, 204) = 482.86, \epsilon = .414, MSE = 3400, p < .001, \eta_p^2 = .16$) showing that compatibility also affects the shape of RT distributions. More interestingly, the interaction between Task, Trial type and Quantile was also significant ($F(8, 408) = 4.22, \epsilon = .267, MSE = 5368, p = .001, \eta_p^2 = .08$) revealing that the effect of compatibility on the shapes of the RT distributions is different between tasks. Indeed, as Figure 2 shows, whereas the difference between compatible and incompatible trials distributions increases as RTs lengthen for flanker and Stroop task, it decreases for the Simon task.

³The values of the “delta plots”, or even their slopes, are often used as dependent variables. However, this amounts to reducing an interaction to a main effect. As a matter of fact, an effect on delta plots corresponds to an interaction between quantiles and trial type. We prefer to keep the interaction rather than reducing it to a main effect.

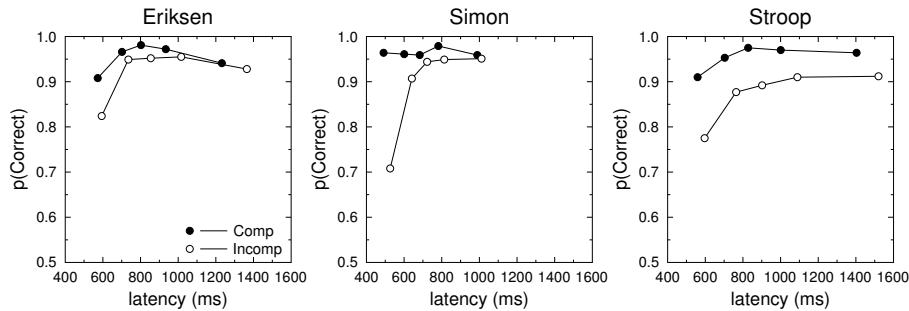


Figure 3: Conditional accuracy function. Probability of correct responses for compatible (black circle) and incompatible (white circle) condition in the flanker task (left), in the Simon task (middle), and in Stroop task (right) as a function of response times.

CAF analysis

As for RTs, no effect of age was observed on accuracy, nor as a main effect, nor as an interaction term (all $F_s < 1$). Figure 3 plots the CAF for each of the three tasks and each Trial type. Main effects of Task ($F(2, 102) = 4.26, \epsilon = .815, MSE = .019, p = .024, \eta_p^2 = .60$), Trial type ($F(1, 51) = 75.84, MSE = .02, p < .001, \eta_p^2 = .08$), and Quantile ($F(4, 204) = 56.46, \epsilon = .617, MSE = .011, p < .001, \eta_p^2 = .53$) were observed. The interaction between Trial type and Quantile was significant ($F(4, 204) = 34.44, \epsilon = .808, MSE = .006, p < .001, \eta_p^2 = .40$) indicating a general decrease in accuracy for fast incompatible responses. The Task by Trial type interaction also proved significant ($F(2, 102) = 5.79, \epsilon = .837, MSE = .015, p = .007, \eta_p^2 = .10$). More importantly, the two way interaction between Task, Trial type and Quantile was also significant ($F(8, 408) = 6.68, \epsilon = .747, MSE = .007, p < .001, \eta_p^2 = .12$) and revealed that the dynamics of the interference effect differ between the three tasks. Two main aspects are driving this interaction. First, accuracy for fast responses on incompatible trials is lower for the Simon than for the other tasks (see Figure 3). The increase in accuracy with RT (i.e., the left part of the curve) is also steeper for the Simon task. Second, while the probability of correct response tends to converge towards 1 for flanker and Simon task (for both compatible and incompatible trials), it stays lower for incompatible trials in the Stroop task.

Model fitting

To go one step further in assessing the similarity of processes between children and adults, we fit a formal model of conflict tasks (DMC) recently proposed by Ulrich et al. (2015), which implements the dual route architecture introduced above. The DMC builds upon the well-established “Drift Diffusion Model” of perceptual decision making (DDM, Ratcliff, 1978; Ratcliff & McKoon, 2008). The general logic underlying the DDM is that, during a decision, noisy samples of task-relevant sensory information are accumulated at a given speed (μ) until a pre-defined level of evidence (b) is reached, at which point the decision terminates in a choice and the response is executed. The DMC extends this framework by incorporating components of automatic processing (Cohen, Dunbar, & McClelland, 1990; Logan, 1980). The automatic and controlled processing routes (see above) converge at the decision level, where both task-relevant and task-irrelevant sensory information are accumulated. The contribution of the automatic processing route is short-living, presumably due to passive decay, active suppression, or both. This

architecture has proven successful to capture data from flanker and Simon tasks (Servant et al., 2016; Ulrich et al., 2015). The model explains RT distributional differences between the two tasks by the relative speed of automatic and controlled processes. Specifically, an automatic activation that develops relatively early (or late) tends to generate an interference effect that decreases (or increases) as decision time increases. We will first briefly present the parameters of the model and then the results of the fits.

As mentioned above, the DDM relies on the idea that noisy samples of task-relevant sensory evidence from the environment are accumulated from a starting point to a criterion called decision threshold (b or $-b$, for correct and incorrect responses). We assume that the starting point is at 0, halfway between the two thresholds, since the two responses are equiprobable⁴. Changing the decision threshold allows to regulate the speed-accuracy trade-off: Higher thresholds produce slower but more accurate responses. The mean rate of evidence accumulation is called the drift rate (μ), and is determined by the quality of sensory evidence and the efficiency of attentional processes. The total RT is equal to the sum of the time to reach a decision threshold, called decision latency, and residual sensory encoding and motor execution latencies (grouped into one parameter Ter). In the DMC, the controlled route (processing the relevant dimension,) is modeled as a standard DDM. In addition, the DMC assumes an automatic process that operates on the task-irrelevant sensory information. The automatic process is transient and modeled as a pulse-like function that represents its short-living contribution to decision making. Formally speaking, it is modeled as a (scaled) gamma function (see Ulrich et al., 2015 for more details). Without going into technical details, two relevant parameters concerning this automatic activation are its peak amplitude (parameter A) and latency (τ)⁵. The automatic process briefly spills over to the decision-making process. Decision making is thus determined by superimposed activations of controlled and automatic processes. The model generates a correct (incorrect) response when this superimposed process hits the correct (incorrect) decision threshold b ($-b$). On compatible trials, automatic and controlled processes converge on activation of the correct response, thereby facilitating RT and accuracy. On incompatible trials, the early incorrect automatic activation hampers the decision process and increases the likelihood of an incorrect choice.

To summarize, by fitting the model to data, one can extract 5 main parameters, namely: μ = drift rate (i.e. speed of information accumulation) of the controlled process, b = height of the decision threshold, Ter = mean duration of non-decisional processes, A = peak amplitude of the automatic activation, and τ = peak latency of the automatic activation. Following Ulrich et al. (2015), we added two assumptions inherited from the standard DDM: variability in non-decision time (normally distributed with mean Ter and standard deviation σ_r) and starting point (uniformly distributed with range σ_z). Unlike most applications of the standard DDM, we did not incorporate between-trial variability in the drift rate of the controlled process. The diffusion noise was fixed at 4 to satisfy a mathematical scaling property of the model (see Ulrich, Schröter, Leuthold & Birngruber, 2016, corrigendum for a theoretical justification). The only parameter allowed to vary between compatible and incompatible trials was the sign of the amplitude of the gamma function modeling the automatic activation, leading to a positive gamma impulse

⁴We verified empirically that participants were not biased: the ratio between left and right response is equal to .9974, which was considered as being close enough to 1 to assume unbiased response

⁵The parameter τ is actually the characteristic time of the function, and the peak latency of the gamma automatic activation is located at $\tau(\alpha - 1)$, where α is the shape parameter of the gamma function. To simplify parameter estimation, we fixed α at 2, consistent with Ulrich et al. (2015). The parameter τ thus directly corresponds to the peak latency of the automatic activation.

Table 1: Best-fitting model parameters for the current data.

	Task	Parameter						
		b	μ	Ter	A	τ	σ_R	σ_Z
Children	Flanker	99.8	0.276	555.7	15.0	441.8	88.3	123.9
	Simon	80.6	0.358	532.4	28.8	59.4	89.6	85.7
	Stroop	100.4	0.199	525.6	12.6	504.2	88.8	102.1
Ulrich et al. (2015)	Flanker	51.3	0.69	331.8	19.2	118.3	36.6	nf
	Simon	54.6	0.69	332.8	16.0	34.94	38.6	nf
Servant et al. (2016)	Simon	60.5	0.47	334.9	15.1	29.4	46.4	nf

Note. For sake of comparison, the parameters obtained in adults in previous studies are reported. b , correct response decision threshold; μ , mean drift of the controlled process; Ter , nondecision time; A , amplitude of the automatic activation; τ , characteristic time of the gamma function (in the current case the latency of its peak) representing the automatic activation; σ_R , standard deviation of nondecision time; σ_Z , starting point range. The τ given for Servant et al. corresponds to the highest color saturation of this study. nf, not fitted.

in compatible trials (i.e., favoring the correct response) and a negative impulse in incompatible trials (i.e., favoring the incorrect response).

For each conflict task, the DMC was fit simultaneously to the cumulative RT data and to the conditional accuracy functions by minimizing the weighted root-mean-squared error statistic (see Ulrich et al., 2015, Appendix E). Data from 5- and 6-year-old children were merged to improve the signal-to-noise ratio, leading to a sample of 53 children. We fit data aggregated across children, because individual fits would have required computational power beyond our current resources.

Inferences from RT models are valid if their parameters are adequately recovered. White, Servant, & Logan (2017) recently conducted a parameter-recovery study for the DMC model. While this study found relatively high correlations between simulated and recovered parameters, these correlations were computed from individual fits, and not based on aggregated data, and thus cannot be used to support the validity of our model parameters. Assessing the validity of our parameters would require a similar parameter recovery study, but on data averaged across subjects. However, the computational burden makes such studies impractical for each individual dataset. This is a limitation of our approach, which requires great caution in interpreting model fits. We will focus on large parametric modulations to reduce the likelihood of drawing erroneous conclusions.

The best fitting parameters are summarized in Table 1, and the quality of the fit can be appreciated in Figure 4 that plots the predicted (lines) and observed data (symbols) for the flanker, Simon and Stroop tasks. Although designed to account for data obtained in adults, the DMC nicely captures children data in the three tasks⁶. This provides further support for the similarity of the underlying processes in both populations. Best fitting parameters presented in Table 1 follow a pattern in agreement with previous standard DDM fits, but also reveal interesting new features. Concerning general DDM parameters (non-decision time - Ter , boundary separation - b and drift rate of the controlled process - μ), the present results extend to conflicts tasks the findings of previous fits of DDM to children data (Weeda et al., 2014, Janczyk et al., 2017, Schuch & Konrad, 2017). First, the non-decision time (Ter) is much longer (around 530 ms) than what was usually reported in adults (300-350 ms, Servant et al., 2016; Ulrich et al., 2015). Second, the

⁶It is interesting to note that the model also accounts for performance on the Stroop task, for which it has not yet been tested on adults

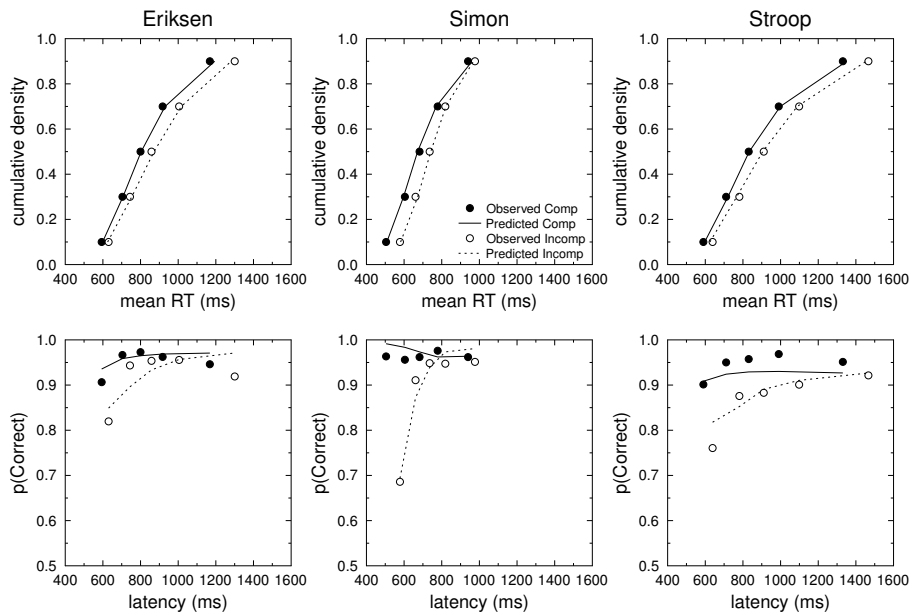


Figure 4: Model fitting. The experimental data are depicted as circles, and the predicted ones as lines. The upper row shows the cumulative density functions, the lower one presents the CAFs.

decision threshold (b) is higher (almost twice in the present dataset) in children than in adults, suggesting that children adopt a more conservative decision criterion. Third, the drift rate for the controlled process (μ) is lower (almost twice) in children (between .2 and .36) than in adults (between .40 and .70).

Concerning the parameters specific to conflict tasks, several aspects are worth mentioning. First, the main between-tasks difference is observed in the dynamics of the automatic activation. This can be inferred from the τ parameter. To better visualize this effect, Figure 5 plots the reconstructed time course of the automatic activation. It reveals much more transient in the Simon than in the Flanker task, as reported on adults. Although no direct comparison with adult data is possible with the Stroop, dynamics in this task are rather similar to the flanker in the children fits, which parallels the fact that the RT distributions are also similar between flanker and Stroop but largely differ from Simon in this population. Although the between-tasks differences in the dynamics of the automatic activation reveal very similar to the ones in adults, within each task, the latencies of the peak of activation are delayed in children compared to adults (parameter τ , 442, 59 and 504 ms respectively for flanker, Simon and Stroop in children vs. 118 and 29-35ms for flanker and Simon, as reported by Servant et al., 2016 and Ulrich et al., 2015 in adults).

Discussion

Conflict processing is one of the core aspects of cognitive control. In adults, it has largely been studied through the so-called “conflict tasks”, among which the flanker, Simon and Stroop tasks are the most often used. Adapted versions of these tasks have revealed useful to investigate conflict processing in children (see Best & Miller, 2010; Diamond, 2013; Garon, Bryson, &

Smith, 2008; Rueda, Posner, & Rothbart, 2005 for reviews). Although longer RTs and higher error rates are reported on incompatible trials compared to compatible ones, in both children and adults, such mean effects mask the underlying dynamics of interference control. The main goal of the present study was to assess whether distributional tools used in adults to track the dynamics of correct and incorrect response activations, can also be applied in children as early as 5-6 years old. Besides showing that it is indeed the case, the present study revealed both strong commonalities but also some differences between children and adults that will be discussed in turn.

Similar dynamics in children and adults

Distributional analyses of RTs support the idea that interference effects at 5-6 years of age, an age at which inhibitory control is still considered as immature, have similar relative dynamics as in adults, both within and between tasks. In adults, it has repeatedly been shown that while in both Stroop and flanker tasks chronometric interference effect increase with RTs, this effect decreases in the Simon task (Burle et al., 2014, Pratte et al., 2010, Ulrich et al., 2015). Interestingly, despite much slower responses in children than in adults, the same between-tasks differences were observed. CAFs were also comparable in children and adults: as in adults, the fastest RTs on incompatible trials were associated to a dramatic drop in accuracy. This initial drop in accuracy, largely participating to the mean interference effect, was present in all three tasks, although with different dynamics. Between-task differences have been less systematically studied on CAF than on RT distributions, but the differences between Simon and flanker tasks resemble the ones reported by Servant et al. (2014) in adults: the initial drop appears steeper and shorter in the Simon than in flanker tasks. The comparable dynamics evidenced in both RT and accuracy distributions between children and adults strongly support the hypothesis of a similarity of the processes at play in the two populations.

The CAF and the shape of the delta-plots (that are often used to summarize effects across RT distributions) have been proposed to, respectively, index automatic response activation and suppression (Ridderinkhof, 2002, Wylie et al, 2010, van den Wildenberg et al., 2010 for an overview). Although the link might be less straightforward than initially proposed (see Fluchère et al., 2015, Spieser et al., 2015, Burle et al., 2014, Ulrich et al., 2015 for discussions), these markers remain very good proxys of the underlying processes. As for adults, it hence appears that interference stems from an “automatic” and (rather) transient response activation by the irrelevant dimension. Interestingly, in the Stroop task, even if the interference effect on error rates also reduces as RTs lengthen, it remains present for long RTs. A similar pattern was observed by Bub et al. (2006) in 7- to 9-year-old children. The origin of this sustained interference requires further investigations.

The reduction of the Simon effect for long RTs, observable in the delta-plots, has been proposed to reflect the “suppression” of the response activated by the stimulus position (De Jong et al., 1994, Ridderinkhof, 2002). This reduction was very strong in the Simon task, much weaker or even absent in the flanker and the Stroop task. Based on both empirical data (Burle et al., 2014) and on theoretical considerations (Ulrich et al., 2015), this difference likely stems from a quantitative difference in the magnitude and/or timing of the suppression, rather than a qualitative difference. According to this view, in young children as in adults, the suppression is stronger and/or faster in the Simon task than in the two other tasks, in which it is likely not strong enough to take over and reduce the interference effect. Although another study already reported negative going delta-plots in the Simon task in 6-year-old children (Iani et al., 2014), we here show that this pattern is specific to this task, as the same children present steadily increasing delta-plots in

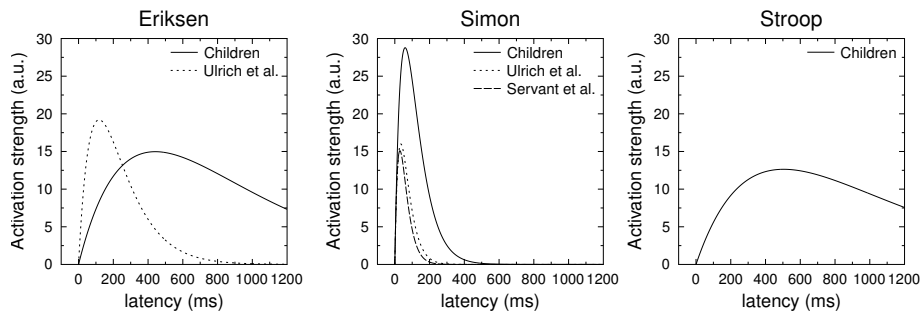


Figure 5: Representation of the dynamics of the automatic response activation for the three tasks. The solid lines correspond to the current dataset, and the dotted lines correspond to Ulrich et al. (2015) and Servant et al. (2016) data in adults. a.u., arbitrary units.

the other tasks. This has direct functional consequences. Indeed, if one accepts that the negative-going delta-plots is linked to the active suppression of the incorrect response activation (Burle, et al., 2002; Ridderinkhof, 2002), this would suggest that this suppression mechanism is already present at 5 years of age, at least in the Simon task. Whether it is already mature, or whether its efficiency keeps on improving with age would require investigations on a wider age range.

The hypothesis of common processes at play in adults and in young children is further strengthened by fits of the DMC model to data. Although best fitting parameters have to be taken with caution (see model fitting section), several aspects are worth mentioning. First, the DMC model, initially developed to account for adults' data, was shown to also nicely fit the children's ones. This suggests that children's and adults' data have the same underlying structure and can be described as parametric variations of each other. Second, the same between-tasks difference between children and adults further supports this common architecture. Indeed, as in adults, the "automatic" activation lasts longer for the flanker than for the Simon task (see Figure 5 for adults – dotted lines, and children – solid lines, no adult fits are available for the Stroop task).

Altogether, the above results illustrate the feasibility and relevance of applying non-parametric distribution analyses to children. Along with the modeling results, they support the notion that interference effects in 5- to 6-year-old children are driven by mechanisms very similar to the ones at play in adults.

... *But with a different within-trial time course*

Besides the similarities in the processes involved between children and adults, the present data also provide evidence for important differences in their time course (see also Rueda et al. 2004). Indeed, RTs are much longer in children than in adults. Fitting the DMC to young children data provide some clues on the origin of their slowed responses. As already mentioned, the DMC inherits some parameters from the standard DDM, and incorporates new ones related to the "automatic" response activation. Concerning the parameters of the "standard" DDM, our results are consistent with the few previous studies that fitted the standard DDM (in simple 2-choice tasks) to children data (Weeda et al., 2014, Janczyk et al., 2017, Schuch & Konrad, 2017). First, the non-decision time (T_{er}), is much longer than in adults, suggesting that a substantial part of the RTs increase in children is not related to decision processes. The respective contributions of sensory encoding and motor processes in this lengthening require further investigations. Second,

children seem to adopt a more conservative decision criterion than adults, since their decision thresholds (b) are almost twice higher than in adults. In contrast to what traditional impulsivity measures suggest (e.g., Kagan, Rosman, Day, Albert, & Phillips, 1964; Messer, 1976; Salkind & Wright, 1977, see Ling, Wong, & Diamond, 2016 for more recent discussion), young children tend to be careful before engaging a response. Note that this higher threshold also contributes to the longer RTs in children. A third cause of RT lengthening is the low drift rate for the controlled process suggesting that the quality of evidence being accumulated is much lower in children than in adults. Although the current data do not allow to trace the reasons for this low drift rate, one can speculate on its potential sources. In general, developmental studies suggest a protracted development of sustained attention from preschool years to young adulthood (Fortenbaugh et al., 2015). Based on a recent study by Lewis, Reeve, Kelly, and Johnson (2017), more lapses of attention in 6-year-olds than in older age groups can be envisaged, that would degrade the sampling of information conveyed by the stimulus, hence leading to a slower and/or more variable (McVay & Kane, 2012) accumulation.

Besides these general aspects, the time course of the supposedly “automatic” response activation also differs largely. The activation of the incorrect response on incompatible trials is often considered “fast and automatic” in adults (Kornblum et al., 1990). In agreement with this view, the drop of accuracy on the CAF is usually observed around 200-300 ms in adults. Although the same drop is observed in children, it occurs much later, around 550-600 ms. As mentioned above, DMC fits revealed that non-decisional latencies (T_{er} including perceptual and motor processes) are much longer in children than in adults. Could a long sensory encoding time potentially explain this delay? This is very unlikely since early sensory processes, as assessed by electroencephalography visual evoked potentials, are almost mature at the ages considered here and occur in the same time range in adults and children (Courchesne, 1978). Furthermore, in an EEG study of the flanker task in 4-year-olds, Rueda et al. (2004) reported latencies of the early (N1 and N2) event-related potentials (ERP) that were very close to the ones measured in adults (between 20 and 30 ms differences maximum) suggesting that a visual representation of the irrelevant dimension does not emerge much later in children than in adults. The activation of the incorrect response hence appears delayed compared to the visual processing, casting doubts on the direct, automatic link between sensory processing and motor activation (see Valle-Inclán & Redondo, 1998, for a similar discussion in adults). Response activation by the irrelevant dimension does not only start later in children, it also lasts longer (about twice longer, see Figure 5). This is true for the three tasks. Whether this duration is due to a passive decay or an active suppression is still an open issue. But, whatever the reason, the incorrect response activation seems to reach its peak much later and vanishes much more slowly in children than in adults. This indexes difficulties in coping with the irrelevant dimension. Further work will be needed to clarify the origin of these dynamics. Beyond the specificities of time courses, the considerable longer delays of responses in children compared to adults are worth considering further. Interestingly, a key finding of research on the development of inhibitory control is that longer delays before responding help. For example, several studies using the day-night version of the Stroop task (Gerstadt et al., 1994,) or Go-Nogo tasks (e.g., Barker & Munakata, 2015, Wiebe, Sheffield, & Espy, 2012) have shown that when young children take more time to respond, they tend to be more accurate. This led to the notion that the interfering effect of the prepotent response activation (i.e., naming the picture in the day-night task, and producing a go response in the Go-Nogo task) is only transient. Following Diamond et al. (2002), several studies revealed that imposing a delay before responding both in Stroop and Go-Nogo tasks was indeed beneficial (see also Ling, et al., 2016; Montgomery & Fosco, 2012; Simpson et al., 2012; Simpson & Riggs, 2007,

see however Barker & Munakata, 2015 for a lack of benefit). A recent, still open, debate is why does delays help (Barker & Munakata, 2015; Ling et al., 2016). Whereas, Diamond and colleagues conclude the underpinning mechanism consists in “allowing the prepotent response to subside and the more-considered answer to reach response threshold”, Barker and Munakata suggest that most studies include a factor confounded with the longer imposed delays namely, hints that could play a role of goal reminders. Their findings actually support this hypothesis. Although the present study was not designed to specifically address this debate, the time course analysis of activation and suppression processes used here confirmed that response accuracy on incompatible trials improves with longer RTs, even without resorting to any external intervention that might have contributed to reactivate the goal. This does not imply that the extra time taken on slow trials was not associated to a spontaneous goal rehearsal. But CAFs allow to analyze the impact of delay within participants and without artificially introducing any external delay, by simply relying on the spontaneous variability in response time (see Barker & Munakata, 2015 for more arguments on the difference between spontaneous and externally imposed delays). We suggest that investigating the influence of goal reminders proposed immediately before stimulus onset on CAFs shapes, could shed new light on the debate. If goal reminding is the key factor, it should support the activation of the correct response that should exceed the response production threshold earlier on and, as a consequence, should allow performance on incompatible trials to become closer to ceiling. More generally, CAF analysis, which relevance in characterizing the time course of automatic response activation in children as young as 5 has been established here, opens new avenues to re-examine developmental issues related to control development.

Conclusion

The present study has both methodological and theoretical implications for the study of conflict processing in children. First, it establishes the feasibility of dynamic analyses of performance through distribution analyses, in children as young as 5-6 years, thereby evidencing the possibility of examining interference control processes separately. Furthermore, it reveals that the two main components evidenced in adults, namely automatic activation and its subsequent suppression, are already present. The time course of these processes differs from the one observed in adults. It, however, presents similar inter-task characteristics. Altogether, the present findings provide a framework to investigate the developmental paths of the different control processes from early childhood on.

Acknowledgements

This study was supported by a French Agence Nationale de la Recherche grant (ANR-15-CE28-0008, DOPCONTROL) and by the European Research Council under the European Community’s Seventh Framework Program (FP7/2007-2013 Grant Agreement 241077).

References

- Abundis-Gutiérrez, A., Checa, P., Castellanos, C., & Rosario Rueda, M. (2014). Electrophysiological correlates of attention networks in childhood and early adulthood. *Neuropsychologia*, *57*, 78–92. <https://doi.org/10.1016/j.neuropsychologia.2014.02.013>
- Ambrosi, S., Lemaire, P., & Blaye, A. (2016). Do young children modulate their cognitive control? *Experimental Psychology*, *63*(2), 117–126. <https://doi.org/10.1027/1618-3169/a000320>

- Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology (Neuropsychology, Development and Cognition: Section C)*, 5(2), 115–129. <https://doi.org/10.1076/chin.5.2.115.3167>
- Barker, J. E., & Munakata, Y. (2015). Time isn't of the essence. *Psychological Science*, 26(12), 1898–1908. <https://doi.org/10.1177/0956797615604625>
- Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, 81(6), 1641–1660. <https://doi.org/10.1111/j.1467-8624.2010.01499.x>
- 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(2), 647–663. <https://doi.org/10.1111/j.1467-8624.2007.01019.x>
- Bub, D. N., Masson, M. E. J., & Lalonde, C. E. (2006). Cognitive control in children. *Psychological Science*, 17(4), 351–357. <https://doi.org/10.1111/j.1467-9280.2006.01710.x>
- Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives*, 8(1), 36–41. <https://doi.org/10.1111/cdep.12059>
- Bogacz, R., Wagenmakers, E. J., Forstmann, B. U., & Nieuwenhuis, S. (2009). The neural basis of the speed–accuracy tradeoff. *Trends in Neurosciences*, 33, 10–16. <https://doi.org/10.1016/j.tins.2009.09.002>
- Burbeck, S. L., & Luce, R. D. (1982). Evidence from auditory simple reaction times for both change and level detectors. *Perception & Psychophysics*, 32, 117–133. <https://doi.org/10.3758/BF03204271>
- Burle, B., Possamaï, C.-A., Vidal, F., Bonnet, M., & Hasbroucq, T. (2002). Executive control in the Simon effect: an electromyographic and distributional analysis. *Psychological Research*, 66(4), 324–336. <https://doi.org/10.1007/s00426-002-0105-6>
- Burle, B., Spieser, L., Servant, M., & Hasbroucq, T. (2014). Distributional reaction time properties in the Eriksen task: marked differences or hidden similarities with the Simon task? *Psychonomic Bulletin & Review*, 21(4), 1003–1010. <https://doi.org/10.3758/s13423-013-0561-6>
- Cao, J., Wang, S., Ren, Y., Zhang, Y., Cai, J., Tu, W., ... Xia, Y. (2013). Interference control in 6–11 year-old children with and without ADHD: behavioral and ERP study. *International Journal of Developmental Neuroscience*, 31(5), 342–349. <https://doi.org/10.1016/j.ijdevneu.2013.04.005>
- Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*, 33(1), 107–112. <https://doi.org/10.1177/001316447303300111>
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, 97(3), 332–361. <https://doi.org/10.1037/0033-295X.97.3.332>
- Colé, P., Duncan, L. G., & Blaye, A. (2014). Cognitive flexibility predicts early reading skills. *Frontiers in Psychology*, 5(June), 565. <https://doi.org/10.3389/fpsyg.2014.00565>
- Courchesne, E. (1978) Neurophysiological correlates of cognitive development: Changes in long-latency event-related potentials from childhood to adulthood. *Electroencephalography and Clinical Neurophysiology*, 45, 468–482. [https://doi.org/10.1016/0013-4694\(78\)90291-2](https://doi.org/10.1016/0013-4694(78)90291-2)
- Daly, M., Delaney, L., Egan, M., & Baumeister, R. F. (2015). Childhood Self-Control and Unemployment Throughout the Life Span. *Psychological Science*, 26(6), 709–723. <https://doi.org/10.1177/0956797615569001>
- 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. *Advances in Developmental Cognitive Neuroscience*, 44(11), 2037–2078. <https://doi.org/10.1016/j.neuropsychologia.2006.02.006>
- De Jong, R., Liang, C.-C., & Lauber, E. (1994). Conditional and unconditional automaticity: A Dual-Process model of effects of spatial Stimulus-Response correspondence. *Journal of Experimental Psychology: Human Perception and Performance*, 20(4), 731–750. <http://doi.org/10.1037/0096-1523.20.4.731>
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Diamond, A. (2014). Want to optimize executive functions and academic outcomes?: Simple, Just Nourish the Human Spirit. *Minnesota Symposia on Child Psychology*, 37, 205–232.

- <https://doi.org/http://dx.doi.org/10.1002/9781118732373.ch7>
- Diamond, A., Kirkham, N., & Amso, D. (2002). Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology, 38*. <https://doi.org/10.1037/0012-1649.38.3.352>
- Durston, S., & Casey, B. J. (2006). What have we learned about cognitive development from neuroimaging? *Neuropsychologia, 44*, 2149–2157. <https://doi.org/10.1016/j.neuropsychologia.2005.10.010>
- 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*(1), 143–149. <https://doi.org/10.3758/BF03203267>
- Fluchère, F., Deveaux, M., Burle, B., Vidal, F., Wildenberg, W., Witjas, T., ... Hasbroucq, T. (2015). Dopa therapy and action impulsivity: subthreshold error activation and suppression in Parkinson's disease. *Psychopharmacology, 232*(10), 1735–1746. <http://10.0.3.239/s00213-014-3805-x>
- Fortenbaugh, F. C., Degutis, J., Germine, L., Wilmer, J. B., Grosso, M., Russo, K., & Esterman, M. (2015). Sustained attention across the life span in a sample of 10,000: Dissociating ability and strategy. *Psychological Science, 26*, 1497–1510. <https://doi.org/10.1177/0956797615594896>
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60. <https://doi.org/10.1037/0033-2909.134.1.31>
- Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2-7 years old on a Stroop-like day-night test. *Cognition, 53*(2), 129–153. [https://doi.org/10.1016/0010-0277\(94\)90068-X](https://doi.org/10.1016/0010-0277(94)90068-X)
- Gratton, G., Coles, M. G., Sirevaag, E. J., Eriksen, C. W., & Donchin, E. (1988). Pre- and poststimulus activation of response channels: a psychophysiological analysis. *Journal of Experimental Psychology: Human Perception and Performance, 14*(3), 331–344. <https://doi.org/10.1037/0096-1523.14.3.331>
- Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*(4), 480–506. <http://doi.org/10.1037/0096-3445.121.4.480>
- Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*(2), 95–112. <https://doi.org/10.1007/BF02289823>
- Heathcote, A., Brown, S. D., & Wagenmakers, E.-J. (2015). An introduction to good practices in cognitive modeling. In B. U. Forstmann, & E.-J. Wagenmakers (Eds). *An Introduction to Model-Based Cognitive Neuroscience* (pp. 25-48). Springer: New York. https://doi.org/10.1007/978-1-4939-2236-9_2
- Hohle, R. H. (1965). Inferred components of reaction times as functions of foreperiod duration. *Journal of Experimental Psychology, 69*, 382-386. <http://dx.doi.org/10.1037/h0021740>
- Hübner, R., Steinhauser, M., & Lehle, C. (2010). A dual-stage two-phase model of selective attention. *Psychological Review, 117*, 759–784. <http://dx.doi.org/10.1037/a0019471>
- Iani, C., Stella, G., & Rubichi, S. (2014). Response inhibition and adaptations to response conflict in 6- to 8-year-old children: Evidence from the Simon effect. *Attention, Perception, & Psychophysics, 76*(4), 1234–1241. <http://dx.doi.org/10.3758/s13414-014-0656-9>
- Ikeda, Y., Okuzumi, H., & Kokubun, M. (2013). Age-related trends of stroop-like interference in animal size tests in 5-to 12-year-old children and young adults. *Child Neuropsychology, 19*(3), 276–291. <http://dx.doi.org/10.1080/09297049.2012.658364>
- Ikeda, Y., Okuzumi, H., & Kokubun, M. (2014). Age-related trends of inhibitory control in Stroop-like big-small task in 3- to 12-year-old children and young adults. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2014.00227>
- Janczyk, M., Büschelberger, J., & Herbort, O. (2017). Larger between-task crosstalk in children than in adults: Behavioral results from the backward crosstalk paradigm and a diffusion model analysis. *Journal of Experimental Child Psychology, 155*, 95–112. <https://doi.org/10.1016/j.jecp.2016.11.007>
- Jiang, Y., Rouder, J. N., & Speckman, P. L. (2004). A note on the sampling properties of the Vincentizing (quantile averaging) procedure. *Journal of Mathematical Psychology, 48*, 186-195.

- <http://doi.org/10.1016/j.jmp.2004.01.002>
- Kagan, J., Rosman, B. L., Day, D., Albert, J., & Phillips, W. (1964). Information processing in the child: Significance of analytic and reflective attitudes. *Psychological Monographs: General and Applied*, 78(1), 1–37. <https://doi.org/10.1037/h0093830>
- Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility – a model and taxonomy. *Psychological Review*, 97(2), 253–270. <https://doi.org/http://dx.doi.org/10.1037/0033-295X.97.2.253>
- Lappin, J. S., & Disch, K. (1972). The latency operating characteristic: I. Effects of stimulus probability on choice reaction time. *Journal of Experimental Psychology*, 92(3), 419–427. <https://doi.org/10.1037/h0032360>
- Lawrence, M. A. (2016). ez: Easy analysis and visualization of factorial experiments. Retrieved from <http://CRAN.R-project.org/package=ez>.
- Leth-Steensen, C., King Elbaz, Z., & Douglas, V. I. (2000). Mean response time, variability, and skew in the responding of ADHD children: A response time distributional approach. *Acta Psychologica*, 104, 167–190. [https://doi.org/10.1016/S0001-6918\(00\)00019-6](https://doi.org/10.1016/S0001-6918(00)00019-6)
- Lewis, F. C., Reeve, R. A., Kelly, S. P., & Johnson, K. A. (2017). Evidence of substantial development of inhibitory control and sustained attention between 6 and 8 years of age on an unpredictable Go/No-Go task. *Journal of Experimental Child Psychology*, 157, 66–80. <http://doi.org/https://doi.org/10.1016/j.jecp.2016.12.008>
- Ling, D. S., Wong, C. D., & Diamond, A. (2016). Do children need reminders on the day–night task, or simply some way to prevent them from responding too quickly? *Cognitive Development*, 37, 67–72. <https://doi.org/10.1016/j.cogdev.2015.10.003>
- Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, 12(4), 523–553. [https://doi.org/10.1016/0010-0285\(80\)90019-5](https://doi.org/10.1016/0010-0285(80)90019-5)
- Logan, G. (1988). Toward an Instance Theory of Automatization. *Psychological Review*, 95, 492–527. <http://dx.doi.org/10.1037/0033-295X.95.4.492>
- Luna, B., & Sweeney, J. A. . (2004). The Emergence of Collaborative Brain Function: fMRI Studies of the Development of Response Inhibition. *Annals of the New York Academy of Sciences*, 1021(1), 296–309. <https://doi.org/10.1196/annals.1308.035>
- Macdonald, J. A., Beauchamp, M. H., Crigan, J. A., & Anderson, P. J. (2014). Age-related differences in inhibitory control in the early school years. *Child Neuropsychology*, 20(5), 509–526. <https://doi.org/10.1080/09297049.2013.822060>
- Matzke, D. & Wagenmakers, E.-J. (2009). Psychological interpretation of ex–Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, 16, 798–817. <https://doi.org/10.3758/PBR.16.5.798>
- Mauchly, J. W. (1940). Significance Test for Sphericity of a Normal n-Variate Distribution. *The Annals of Mathematical Statistics*, 11(2), 204–209. <http://www.jstor.org/stable/2235878>
- McGill, W. J. & Gibbon, J. (1965). The General-Gamma Distribution and Reaction Times. *Journal of Mathematical Psychology*, 2, 1–18. [https://doi.org/10.1016/0022-2496\(65\)90014-3](https://doi.org/10.1016/0022-2496(65)90014-3)
- McVay, J. C., & Kane, M. J. (2012). Drifting from slow to “d’oh!”: Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 525–549. <http://dx.doi.org/10.1037/a0025896>
- Messer, S. B. (1976). Reflection-impulsivity: A review. *Psychological Bulletin*, 83(6), 1026–1052. <https://doi.org/10.1037/0033-2909.83.6.1026>
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., ... Ross, S. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, 2693–2698. <https://doi.org/10.1073/pnas.1010076108>
- Montgomery, D. E., & Fosco, W. (2012). The Effect of Delayed Responding on Stroop-Like Task Performance Among Preschoolers. *The Journal of Genetic Psychology: Research and Theory on Human Development*, 173(2), 142–157. <http://doi.org/http://dx.doi.org/10.1080/00221325.2011.583699>
- Palmer, J., Huk, A. C. & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision, *Journal of Vision*, 5(5):1. doi:10.1167/5.5.1

- Pratte, M. S., Rouder, J. N., Morey, R. D., & Feng, C. (2010). Exploring the differences in distributional properties between Stroop and Simon effects using delta plots. *Attention, Perception, & Psychophysics*, 72(7), 2013–2025. <https://doi.org/10.3758/APP.72.7.2013>
- Prevor, M. B., & Diamond, A. (2005). Color–object interference in young children: A Stroop effect in children 3–6 years old. *Cognitive Development*, 20(2), 256–278. <https://doi.org/10.1016/j.cogdev.2005.04.001>
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108. <https://doi.org/10.1037/0033-295X.85.2.59>
- Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, 86(3), 446–461. <https://doi.org/10.1037/0033-2909.86.3.446>
- Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, 20(4), 873–922. <https://doi.org/10.1162/neco.2008.12-06-420>
- Ratcliff, R. & McKoon, G. (2015). Aging effects in item and associative recognition memory for pictures and words. *Psychology and Aging*, 30, 669-674. <http://dx.doi.org/10.1037/pag0000030>
- Ratcliff, R., Thapar, A., Gomez, P. & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging*, 19, 278-289. <http://dx.doi.org/10.1037/0882-7974.19.2.278>
- Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, 16, 323-341. <http://dx.doi.org/10.1037/0882-7974.16.2.323>
- Ratcliff, R., Thapar, A. & McKoon, G. (2003). A diffusion model analysis of the effects of aging on brightness discrimination. *Perception & Psychophysics*, 65, 523-535. <https://doi.org/10.3758/BF03194580>
- Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language*, 50, 408-424. <https://doi.org/10.1016/j.jml.2003.11.002>
- Ridderinkhof, K. R. (2002). Micro- and macro-adjustments of task set: activation and suppression in conflict tasks. *Psychological Research*, 66(4), 312–323. <https://doi.org/10.1007/s00426-002-0104-7>
- Ridderinkhof, K. R., Scheres, A., Oosterlaan, J., & Sergeant, J. A. (2005). Delta plots in the study of individual differences: new tools reveal response inhibition deficits in AD/Hd that are eliminated by methylphenidate treatment. *Journal of Abnormal Psychology*, 114(2), 197–215. <https://doi.org/10.1037/0021-843X.114.2.197>
- Ridderinkhof, K. R., van der Molen, M. W., Band, G. P., & Bashore, T. R. (1997). Sources of interference from irrelevant information: A developmental study. *Journal of Experimental Child Psychology*, 65(3), 315–341. <https://doi.org/10.1037/0021-843X.114.2.197>
- Ridderinkhof, K. R., van den Wildenberg, W. P., Wijnen, J., & Burle, B. (2004). Response inhibition in conflict tasks is revealed in delta plots. In M. Posner (Ed.), *Cognitive Neuroscience of Attention*. New York: Guilford Press, pp. 369-377.
- Rouder, J. N. & Speckman, P.L. (2004). An evaluation of the Vincentizing method of forming group-level response time distributions. *Psychonomic Bulletin & Review*, 11, 419-427. <https://doi.org/10.3758/BF03196589>
- Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, 42(8), 1029–1040. <https://doi.org/10.1016/j.neuropsychologia.2003.12.012>
- Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The Development of Executive Attention: Contributions to the Emergence of Self-Regulation. *Developmental Neuropsychology*, 28(2), 573–594. https://doi.org/10.1207/s15326942dn2802_2
- Salkind, N. J., & Wright, J. C. (1977). The development of reflection-impulsivity and cognitive efficiency. *Human Development*, 20(6), 377–387. <https://doi.org/10.1159/000271569>
- Schuch, S., & Konrad, K. (2017). Investigating task inhibition in children versus adults: A diffusion model analysis. *Journal of Experimental Child Psychology*, 156, 143–167. <http://doi.org/https://doi.org/10.1016/j.jecp.2016.11.012>
- Schwarz, W., & Miller, J. (2012). Response time models of delta plots with negative-going slopes. *Psychonomic Bulletin & Review*, 19(4), 555–574. <https://doi.org/10.3758/s13423-012-0254-6>

- Servant, M., Montagnini, A., & Burle, B. (2014). Conflict tasks and the diffusion framework: Insight in model constraints based on psychological laws. *Cognitive Psychology*, *72*, 162–195. <https://doi.org/10.1016/j.cogpsych.2014.03.002>
- Servant, M., White, C., Montagnini, A., & Burle, B. (2016). Linking theoretical decision-making mechanisms in the Simon task with electrophysiological data: A Model-based Neuroscience Study in Humans. *Journal of Cognitive Neuroscience*, *28*(10), 1501–1521. <https://doi.org/10.1162/jocn.a.00989>
- Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In Robert W. Proctor and T. Gilmour Reeve (Ed.), *Advances in Psychology* (Vol. Volume 65, pp. 31–86). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)61218-2](https://doi.org/10.1016/S0166-4115(08)61218-2)
- Simon, J. R., & Berbaum, K. (1988). Effect of irrelevant information on retrieval time for relevant information. *Acta Psychologica*, *67*(1), 33–57. [https://doi.org/10.1016/0001-6918\(88\)90023-6](https://doi.org/10.1016/0001-6918(88)90023-6)
- Simpson, A., & Riggs, K. J. (2007). Under what conditions do young children have difficulty inhibiting manual actions? *Developmental Psychology*, *43*(2), 417–428. <https://doi.org/10.1037/0012-1649.43.2.417>
- Simpson, A., Riggs, K. J., Beck, S. R., Gorniak, S. L., Wu, Y., Abbott, D., & Diamond, A. (2012). Refining the understanding of inhibitory processes: how response prepotency is created and overcome. *Developmental Science*, *15*(1), 62–73. <https://doi.org/10.1111/j.1467-7687.2011.01105.x>
- Spieser, L., van den Wildenberg, W. P., Hasbroucq, T., Ridderinkhof, K. R., & Burle, B. (2015). Controlling Your Impulses: Electrical Stimulation of the Human Supplementary Motor Complex Prevents Impulsive Errors. *The Journal of Neuroscience*, *35*(7), 3010–15. <https://doi.org/10.1523/JNEUROSCI.1642-14.2015>
- Stins, J. F., Polderman, J. C. T., Boomsma, D. I., & de Geus, E. J. C. (2007). Conditional accuracy in response interference tasks: Evidence from the Eriksen flanker task and the spatial conflict task. *Advances in Cognitive Psychology*, *3*(3), 409–417. <https://doi.org/10.2478/v10053-008-0005-4>
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. <https://doi.org/10.1037/h0054651>
- Teodorescu, A.R. & Usher, M. (2013). Disentangling Decision Models: From Independence to Competition. *Psychological Review*, *120*, 1–38. <http://dx.doi.org/10.1037/a0030776>
- Thapar, A., Ratcliff, R., & McKoon, G. (2003). A diffusion model analysis of the effects of aging on letter discrimination. *Psychology and Aging*, *18*, 415–429. <http://dx.doi.org/10.1037/0882-7974.18.3.415>
- Ulrich, R. & Miller, J. (1993). Information processing models generating lognormally distributed reaction times. *Journal of Mathematical Psychology*, *37*, 513–525. <https://doi.org/10.1006/jmps.1993.1032>
- Ulrich, R., Schröter, H., Leuthold, H., & Birngruber, T. (2015). Automatic and controlled stimulus processing in conflict tasks: Superimposed diffusion processes and delta functions. *Cognitive Psychology*, *78*, 148–174. <https://doi.org/10.1016/j.cogpsych.2015.02.005>
- Ulrich, R., Schröter, H., Leuthold, H., & Birngruber, T. (2016). Corrigendum to Automatic and controlled stimulus processing in conflict tasks: Superimposed diffusion processes and delta functions. *Cognitive Psychology*, *91*, 150. doi:10.1016/j.cogpsych.2016.11.003
- Valle-Inclán, F., & Redondo, M. (1998). On the automaticity of ipsilateral response activation in the Simon effect. *Psychophysiology*, *35*(4), 366–371. <https://doi.org/10.1111/1469-8986.3540366>
- van den Wildenberg, W. P., Wylie, S. A., Forstmann, B. U., Burle, B., Hasbroucq, T., & Ridderinkhof, K. R. (2010). To head or to heed? Beyond the surface of selective action inhibition: a review. *Frontiers in Human Neuroscience*. <https://doi.org/10.3389/fnhum.2010.00222>
- Vincent. (1912). The function of the vibrissae in the behavior of the white rat. *Animal Behavior Monographs*, *1*(5), 1–82.
- Weeda, W. D., van der Molen, M. W., Barceló, F., & Huizinga, M. (2014). A diffusion model analysis of developmental changes in children’s task switching. *Journal of Experimental Child Psychology*, *126*(0), 178–197. <http://doi.org/http://dx.doi.org/10.1016/j.jecp.2014.05.001>
- White, C. N., Ratcliff, R., & Starns, J. J. (2011). Diffusion models of the flanker task: Discrete versus gradual attentional selection. *Cognitive Psychology*, *63*, 210–238.

- <https://doi.org/10.1016/j.cogpsych.2011.08.001>
- White, C. N., Servant, M., & Logan, G. D. (2017). Practical considerations for using conflict-based diffusion models to interpret choice RT data. *Psychonomic Bulletin & Review*.
<https://doi.org/10.3758/s13423-017-1271-2>
- Wiebe, S. A., Sheffield, T. D., & Espy, K. A. (2012). Separating the fish from the sharks: A longitudinal study of preschool response inhibition. *Child Development*, *83*(4), 1245–1261.
<https://doi.org/10.1111/j.1467-8624.2012.01765.x>
- Wright, I., Waterman, M., Prescott, H., & Murdoch-Eaton, D. (2003). A new Stroop-like measure of inhibitory function development: typical developmental trends. *The Journal of Child Psychology and Psychiatry*, *44*, 561–575. <https://doi.org/10.1111/1469-7610.00145>
- Wylie, S. A., Ridderinkhof, K. R., Elias, W. J., Frysinger, R. C., Bashore, T. R., Downs, K. E., . . . van den Wildenberg, W. P. M. (2010). Subthalamic nucleus stimulation influences expression and suppression of impulsive behaviour in Parkinson's disease. *Brain*, *133*(12), 3611–3624.
<http://doi.org/10.1093/brain/awq239>