Cognitive Development 27 (2012) 440–454

Contents lists available at SciVerse ScienceDirect

Cognitive Development

Using finite mixture of GLMs to explore variability in children's flexibility in a task-switching paradigm

Bruno Dauvier[a], Nicolas Chevalier[b,*], Agnès Blaye[c]

[a] Centre PSYCLE, Aix-Marseille Université, Aix-en-Provence, France
[b] Institute of Cognitive Science, University of Colorado Boulder, Boulder, CO 80309-0345, United States
[c] Laboratoire de Psychologie Cognitive, CNRS and Aix-Marseille Université, Marseille, France

ARTICLE INFO

Keywords:
Finite mixture model
Generalized linear mixed model
Latent class model
Task switching
Executive control
Cognitive strategies
Intra-individual variability
Inter-individual variability

ABSTRACT

The present study illustrates the usefulness of finite mixture of generalized linear models (GLMs) to examine variability in cognitive strategies during childhood. More precisely, it addresses this variability in set-shifting situations where task-goal updating is endogenously driven. In a task-switching paradigm 5–6-year-olds had to switch between color- and shape-matching rules as a function of a predetermined, predictable task sequence. A finite mixture of GLMs was fitted to explore individual differences in performance. The statistical model revealed five response profiles, defined by accuracy and response times. These response profiles likely correspond to different cognitive strategies with varying efficiency and differential relations to working memory capacity (assessed by backward digit span). These results illustrate the heuristic value of statistical modeling to reveal the behavioral and cognitive variability in the temporal dynamics of children's cognitive functioning.

Cognitive development has long been conceived as a stage-like progression toward increasing cognitive efficiency and maturity, as best illustrated by the large influence of Piaget's theory. According to stage theories, development consists of a universal progression through the same stages. At each stage, most or all children use the same processes and strategies. Yet, both within- and between-group variability in strategies is psychologically plausible. In most domains of development, recent research

* Corresponding author at: Institute of Cognitive Science, 345 UCB, University of Colorado Boulder, Boulder, CO 80309-0345, United States. Tel.: +1 303 492 6389; fax: +1 303 492 6389.

E-mail address: nicolas.chevalier@colorado.edu (N. Chevalier).

0885-2014/$ – see front matter
http://dx.doi.org/10.1016/j.cogdev.2012.07.004

has moved away from stage views and emphasized instead both intra- and inter-individual differences (Siegler, 1997). Exploring cognitive variability in any domain of cognitive development requires adequate methodological instruments. Here we argue that computational statistical modeling, more precisely finite mixture of autoregressive generalized linear models (GLMs; Lindsey, 1997), can provide new insights into cognitive variability, which we illustrate by examining the variability in the cognitive strategies that preschoolers use to update goal representations in a set-shifting task that assesses executive control.

Classical statistical frameworks (e.g., analysis of variance) generally are characterized by a substantial gap between the theoretical representations of the targeted psychological processes and the statistical hypotheses that are actually tested because these frameworks are almost always based on aggregated data at the group level while psychological processes occurred at the individual level. In contrast, statistical modeling bridges this gap by estimating parameters at the individual level; these parameters directly reflect cognitive processes and thus can be interpreted more straightforwardly in terms of cognitive functioning. GLMs offer a wide range of very flexible tools to investigate psychological processes. They also provide the opportunity to address several theoretical issues within a single analysis, reducing the risk of hidden effects due to data aggregation. For instance, in our study this type of modeling allowed us to test our main theoretical questions within a single model using one set of parameters for each question, whereas several distinct analyses would have been necessary using a classical statistical framework.

Statistical modeling offers the possibility of exploring individual differences in depth. Models based on latent classes identify groups based on their response profiles. When one examines variability in cognitive processes or strategies, expected individual differences are more qualitative than quantitative. In such cases, individual response profiles are not ordered along a continuum. Because latent class models handle a priori unknown qualitative differences among groups of individuals, they are especially well-suited to explore individual differences. Models that combine the flexibility of GLMs with the possibility of revealing latent classes are known in the statistical modeling literature as finite mixture of generalized linear models (Grün & Leisch, 2008) or variance components GLM (Aitkin, 1999). Their main principle is that the relations that exist among multiple variables in a dataset often are more accurately characterized by multiple regression models with different parameter values, fitted to different latent subgroups of individuals, relative to a single set of parameter values for the entire sample. In addition, these regression models can be built to reflect the temporal dynamics of cognitive processes (e.g., with age, across experimental sessions or even across trials within a session), which is done through an autoregressive term that uses the current state of the cognitive system as an explanatory variable to predict the next state of the system (Aitkin & Alfo, 2003). This methodology is a powerful tool to study behavioral and cognitive variability both between and within subjects. Here we illustrate the heuristic value of such statistical models in the context of children's executive control and, more specifically, goal updating strategies in set-shifting situations.

Executive control refers to the intentional and goal-directed regulation of one's own thoughts and actions. It allows one to orient attention toward goal-relevant information and appropriate behaviors. Executive control is required, and can be exerted, only if one has a specific goal to achieve. For instance, children can orient their attention to the information relevant to solve an arithmetic problem only if they intend to solve this particular problem. Forming a representation of the relevant task, that is, deciding about the relevant task goal, is challenging for preschoolers especially when tasks constantly change, as is the case in task-switching situations, such as the Advanced Dimensional Change Card Sort (Advanced DCCS), where children have to switch between matching a bidimensional stimulus with response options by color or shape as a function of task cues (e.g., a star beside the stimulus signals that color is relevant while a square means shape is relevant; Zelazo, 2006). Consistently, recent evidence suggests that set-shifting development is largely driven by improvement in goal representation (Chevalier & Blaye, 2009; Marcovitch, Boseovski, & Knapp, 2007; Morton & Munakata, 2002; Snyder & Munakata, 2010).

Thus far, research has begun to uncover the processes underlying goal representation in situations in which children are provided with some environmental information, such as task cues (Blaye & Chevalier, 2011; Chevalier & Blaye, 2009; Chevalier, Wiebe, Huber, & Espy, 2011), response feedback (Chevalier, Dauvier, & Blaye, 2009), or common stimulus features (Snyder & Munakata, 2010). How do preschoolers determine when and what goal to switch to when no external goal-related information is provided and they must rely exclusively on internal cues? Such situations are frequent in preschools, where children often receive series of instructions about tasks to perform successively. Such tasks are frequently used to assess set-shifting in older children and adults (Kiesel et al., 2010; Vandierendock, Liefooghe, & Verbruggen, 2010). For instance, in the alternating-runs version of the task-switching paradigm (Rogers & Monsell, 1995), participants are presented with no task cues and instructed to follow a predetermined and predictable task sequence (e.g., switch on every other trial: task A, task A, task B, task B, etc.). Without external cues, goal representation may especially tax endogenous processes. Indeed, in adults switch costs (i.e., the performance decrement between trials where the task repeats and trials requiring a task switch) observed in the alternating-runs paradigm are larger than those observed with task cues (Altmann, 2007).

When no external cues are available, the relevant task goal can be updated in working memory by keeping track of the current position in the task sequence, which likely requires memorizing and using the information related to the immediately preceding trials (e.g., if task A has just been performed twice, task B is now relevant). Further, the difficulty of updating the task goal may change across trials. As one moves across trials, the clear indication about the task to start with probably becomes less informative for the current trial (Altmann, 2004). Most importantly, memory traces related to prior trials accumulate progressively (although their activation likely decays with time). Therefore, interference increases as trials accumulate, making it harder to use the information of the most recent trials to keep track of the current position in the task sequence and update the task goal accordingly.

If performance relies on maintenance and use of previous trial information, it may vary as a function of working memory capacity. The role of working memory is all the more probable, since working memory is involved in task goal updating and maintenance in both adults (Baddeley, Chincotta, & Adlam, 2001; Emerson & Miyake, 2003; Saeki & Saito, 2004a, 2004b, 2009) and children (Marcovitch, Boseovski, Knapp, & Kane, 2010; Morton & Munakata, 2002). Children's ability to update task goals on switch trials and maintain the same goal on no-switch trials should thus relate to working memory. Moreover, individual differences in working memory may not only relate to quantitative differences in task goal updating (e.g., number of errors) but also to qualitative differences in updating strategy. In particular, children with high working-memory may have enough working memory resources both to store the information about the tasks performed on the previous two trials and to decide on the relevant task goal. They also may successfully resist increasing interference across trials. In contrast, children with low working memory capacity may not have enough resources to update task goals on the basis of previous trial information and hence be more likely to either perseverate on a single task or switch tasks in a random fashion.

The issue of strategy differences relates to the variability of flexible behaviors (or lack thereof) and underlying strategies across children. Indeed, substantial variability occurs both within and between age groups. For instance, preschool children commit multiple types of errors in task-switching situations that offer more than two response options, indicating response variability at the behavioral level (Chevalier & Blaye, 2008; Deák, 2000). Preschoolers also show variability at the cognitive level; some perseverate at the level of task dimensions while others do so at the level of specific stimulus features (Hanania, 2010). Finally, variability is observed between age groups. Age differences in fixation patterns (Chevalier, Blaye, Dufau, & Lucenet, 2010) and distinct neural networks observed in the Advanced DCCS (Morton, Bosma, & Ansari, 2009) are suggestive of age differences in cognitive strategies.

The heterogeneity of performance (and related cognitive strategies) often is inadequately captured by traditional scoring methods and statistical analysis. For instance, chance-level mean performance in the Advanced DCCS (50% correct with two responses options) may reflect random responding, random switching, or perseveration on the same task across all trials. Similarly, a switch-cost reduction may reflect different processes depending on whether it is due to a performance increase on switch trials or decrease on repetition trials. Therefore, it is important to go beyond mean performance and classify children in groups of homogeneous response profiles in order to better reveal behavioral and cognitive variability.

The present study illustrates the usefulness of computational statistical modeling for studying cognitive development, by exploring preschoolers' variability in goal updating strategies in a situation where goal representation is entirely endogenously driven. To this end, 5–6-year-olds completed an alternating-runs version of the Advanced DCCS (they had to match bidimensional stimuli with one of two possible response options by either shape or color on the basis of a task sequence that required switching on every second trial, e.g., color, color, shape, shape, etc.), and a backward digit span task. We addressed three main questions. (1) What strategies do children use to determine the relevant task goal and when to update or maintain it on the basis of the task sequence? Do these strategies vary across children? (2) Does performance change across trials? In particular, does strategy difficulty increase across trials because of the accumulative interference created by previous trials? (3) To what extent does goal updating depend on working memory, and is working memory associated with cognitive strategy use?

1. Method

1.1. Participants

Participants were 79 5–6-year-olds (M = 70.8 months, SD = 3.3 months, range = 65–77 months, 43 girls). One additional child was eliminated because she did not follow instructions. Children were recruited from two preschools in a small town in France. Most were Caucasian and came from middle- to upper-class backgrounds.

1.2. Materials and procedure

Participants were tested individually in a quiet room at their preschool. All children completed an alternating-runs version of the Advanced DCCS and a backward digit span task. Children completed these tasks, along with others not reported here, in two sessions of about 30 min each, one week apart. The order of the two tasks was counterbalanced.

1.2.1. Advanced DCCS

The Advanced DCCS (Zelazo, 2006) was administered on a laptop computer (15-in. monitor HP Compaq nx9000) and run with E-Prime® (Psychology Software Tools, Pittsburgh, PA, USA). Children had to respond by pressing one of two keys (corresponding to the ‘q’ and ‘p’ keys of a QWERTY keyboard). The remaining keys were masked. On every trial, children had to match a stimulus picture with one of two response pictures on either color (color game) or shape (shape game). Two pictures of different colors and shapes (e.g., a blue boat and a red rabbit), were displayed at the top of the screen one at a time. Each response picture matched each stimulus on either color or shape (e.g., a red boat and a blue rabbit). Both response pictures remained visible throughout the task and were displayed on the two bottom corners of the screen. Pictures were about 6 cm × 6 cm. Each participant completed a version of the task with one of 4 combinations of colors and shapes. Every trial started with a fixation cross followed by a stimulus that remained on screen until a response was entered. Then the stimulus was moved to the side of the response given for 500 ms, to emulate putting cards into boxes as in the card version of the Advanced DCCS, and the next trial initiated. Pace was controlled by the experimenter.

Children were told that they would see pictures and would play either the color game or the shape game. In the color game, they were instructed to press the key under the bottom picture of the same color as the top picture. In the shape game, they were instructed to press the key under the bottom picture of the same shape as the top picture. Children were asked to respond as quickly and accurately as possible. They started with two simple blocks in which they had to consistently play the color game or the shape game (order counterbalanced). The simple blocks were used for familiarization. Each simple block consisted of 5 training trials (repeated if children committed more than two errors) and 10 test trials. The experimenter helped children on training trials if necessary but not on test trials.

Children were then told they would now play the two games at the same time and proceeded to the mixed blocks, where they were expected to alternate between color and shape matching on every second trial. The alternating rule was explained as follows: “*Now we’re going to play both the Color and Shape Games. You’ll play the Color Game twice, and then you’ll change and play the Shape Game twice, and* *then twice again the Color Game, and twice again the Shape Game, and so on. So you'll play the same game twice and then change games. You'll do, for instance, color-color, shape-shape, color-color and so on.*" Pilot testing suggested that children understood these instructions.

Six demonstration trials were administered. The task sequence was repeated and children were instructed to start with one dimension (counterbalanced across blocks and children) and completed 20 test trials (first mixed block), after which they took a short break. Then they were repeated the task sequence and the game to start with and completed another 20 test trials (second mixed block). No feedback was provided on test trials.

One property of the alternating-runs version of the task-switching paradigm is that the correct task on a given trial depends on the two previous responses. Therefore, the outcome variable must adequately capture this temporal dynamic. To this end, we used a binary outcome variable (*Switch*) to code responses. The value 0 meant that the given response corresponded to a task repetition (relative to the previous trial), whereas the response was coded 1 when the response denoted a task switch. The correct response pattern was 1 after 0 (correct task switch after a repetition) and 0 after 1 (correct repetition after a switch). For example, in the response pattern CCSSCSSCC, S and C indicate the selection of the response that matches the stimulus shape and color. These responses would be coded _01011010 on the *Switch* variable, which corresponds to four switches and four repetitions. The first underscore denotes the fact that the first response cannot be coded (it is neither a switch nor a repetition). The second of the two consecutive "1s" in the middle reflects an unexpected switch.

The outcome variable *Switch* was used in the model to study the behavioral dynamics but it did not directly provide an index of accuracy. To validate the output of the model, an additional accuracy variable (*Acc*) was computed as follows: A switch following a non-switch and a non-switch following a switch were considered correct and coded 1, whereas a switch following another switch or a non-switch following another non-switch were considered inaccurate and coded 0. The first response of each block could not be coded. Our previous example (CCSSCSSCC) would be coded _11110111 on *Acc*, corresponding to seven correct responses and one error (color should have been repeated on the 6th trial).

Our scoring method assumes that the observed response directly reflects the task the child intended to perform on any given trial (i.e., the goal representation that was active when the response was entered). However, at times, children may fail to select the response corresponding to the task they intend to perform (i.e., the active goal). In such cases, responses do not match the intended task. Such response selection errors may reflect response inhibition errors because children may have selected either the last response associated with a stimulus (if perseveration occurred at the level of stimulus-response associations) or the motoric response that was performed on the previous trial (if perseveration occurred at the motor level). The latter seems improbable given that 5- and 6-year-olds show high performance levels in simple blocks that also require switching motoric responses (Chevalier & Blaye, 2009). It remains possible that selection errors at the level of stimulus-response associations occurred; however, there is no reason to expect the frequency of these errors to be influenced by the task sequence or to change as children progress across trials. Therefore, response selection errors should not affect differences in response profiles across trials.

Response times (RTs) were also assessed. RT outliers (<300 ms or >10,000 ms) were trimmed, resulting in1.7% of RT values being discarded.

1.2.2. Backward digit span task

Verbal working memory was measured with the backward digit span task. Children were told they would hear a series of digits that they would have to repeat in a backward fashion. Two demonstration series with two digits were administered, providing feedback and guidance as needed. Then children moved to the test series. Digits were pronounced at the pace of one digit per second. Children started with two test series, each containing two digits. If children correctly responded to at least one of these series, they were administered two series containing one additional digit. Series length progressively increased up to 8 digits or until the child incorrectly responded to both series of any given length, at which point the task was discontinued. Following Case, Kurland, and Goldberg (1982), scores consisted in the highest series length that children reached (correct responses to both trials) plus 0.5 point per additional correct response.

1.3. Model specification

A binomial model, specifically dedicated to binary outcome variables, was used in order to work at the response level, without preliminary aggregation, and to examine within-subject changes across trials. Given the need for a statistical instrument appropriate for a binary outcome variable that allows the value at time t to depend on the value at time $t-1$, generalized linear model (Lindsey, 1997) was used. It is a wide class of models that encompasses, among other techniques, analysis of variance, regression and, logistic regression. The temporal relation between a given response and the previous one was handled by an autoregressive term that computes the probabilities of a switch after a switch and after a repetition, by using responses at time $t-1$ as an explanatory variable to predict responses at time t. Of interest for the present study, GLM also offers the possibility to test main effects and interactions between categorical and quantitative explanatory variables.

The binary outcome variable (*Switch*) was used to statistically model the probability of a task switch on any given trial as a function of the following explanatory variables. The first explanatory variable (*PrevSwitch*) was the status of the previous response (i.e., task switch or repetition). Second, as interference was expected to grow as a function of trial number within a block, trial number was also taken into account as an explanatory variable (*Trial*). A third explanatory variable (*Block*) was the mixed block of trials (block 1 vs. block 2) that could have an influence because the break and instruction repetition between the two blocks may lower interference and/or help children further understand task instructions. Thus, in the model the probability of a switch was supposedly influenced by the previous response, trial number and the block, that is, a binary variable, a numerical variable and a categorical variable. Interactions between these three explanatory variables were also tested.

Qualitative differences across children, potentially reflecting different strategies, were investigated using models that combine the flexibility of GLMs with the possibility of revealing latent classes: finite mixture of generalized linear models (Aitkin, 1999; Grün & Leisch, 2008). In these models, several regression equations are fitted to the same dataset, yielding a different set of parameters for each group (or latent class) of subjects identified during the estimation procedure. In this model, we obtained eight parameters for each class of subject. Children with homogeneous responses profiles are grouped together. The final number of classes is defined by comparing the goodness of fit across a series of models with an increasing number of classes.

GLM with latent classes often yields a large set of parameters because the number of parameters corresponds to the number of explanatory variables and interaction multiplied by the number of latent classes. Such a large set of parameters can be challenging to interpret. To circumvent this problem, we took advantage of the model's capacity to compute predictions, as in any regression model, and to build graphical representations of the typical response pattern observed in each latent class. The latent classes were then compared on response times (RTs) which were analyzed using GLM with mixed effects, allowing us to plot the typical RT pattern for each latent class. Finally, a multinomial regression model (Faraway, 2005) was used to explore the link between working memory capacity and latent classes.

2. Results

2.1. Model fitting

The number of classes was determined by comparing the goodness of fit across models with an increasing number of classes. Models with more classes describe the data in further detail and show lower negative log-likelihood, but they are less parsimonious and have more degrees of freedom. When both precision and parsimony were taken into account with the Akaike Information Criterion (AIC, Akaike, 1974), the model with five classes (M5) was the best (Table 1). The results of a finite mixture model indicate, for each participant, the probability of belonging to each class. Six participants were not clearly classified in one of the five classes; their highest probabilities were below 0.80. They were nevertheless classified in their most likely class.

Table 1
Goodness of fit of the finite mixture of generalized linear models.

Model	Number of classes	DF	–2 log likelihood	AIC
M2	2	16	2912	2944
M3	3	24	2813	2861
M4	4	32	2748	2812
M5	5	40	2723	**2803**
M6	6	48	2715	2811

Note. The lowest AIC (Akaike, 1974) corresponds to the best-fitted model regarding both likelihood and parsimony. The degrees of freedom of the models correspond to their numbers of parameters: Eight times the number of classes.

Table 2
Probabilities of a switch and their slopes within each block in the model with five classes.

	N	Mean probability of a switch after a repetition/a switch	
		Block 1	Block 2
Class 1	8	**0.06/0.34***($p = 0.003$)	0.001/0
Class 2	12	**0.07/0.30*** ($p = 0.002$)	0.25/0.28
Class 3	38	**0.57/0.16*** ($p < .001$)	**0.54/0.08***($p < .001$)
Class 4	11	**0.92/0.05*** ($p < .001$)	**0.96/0.08*** ($p < .001$)
Class 5	10	0.36/0.23	**0.99/0.24*** ($p < .001$)

Note. Probabilities in bold were significantly different ($|t| > 2$) from each other in the block.

Table 3
Descriptive statistics by class.

	Switch %		RT median		Accuracy M	
	Block 1	Block 2	Block 1	Block 2	Block 1	Block 2
Class 1	10	3	1153	1149	0.13	0.03
Class 2	11	26	1143	1030	0.15	0.38
Class 3	41	37	1967	1799	0.68	0.67
Class 4	49	51	1859	2036	0.93	0.94
Class 5	33	57	1558	1438	0.52	0.87
Total	34	36	1647	1550	0.56	0.62

Note. The percentage of switch was computed on all trials. A perfect response pattern contains 50% of switches (but 50% switches do not necessarily denote perfect responding as switches and repetitions need to alternate regularly).

The five classes were significantly different from each other. The mean probabilities of a switch for each class in the model with five classes (M5) are given in Table 2.[1] Two of these probabilities significantly changed within a block. In class 3, the probability of a switch after a repetition decreased during the second block (slope = –0.1, $p < 0.05$), indicating that performance worsened across trials within block 2. The same phenomenon was observed in class 5, but within the first block (slope = –0.12, $p < 0.05$). The other slopes were not significantly different from 0. For each class of children, observed accuracy, percentage of switches, and RTs are provided in Table 3.

RTs were investigated using generalized linear mixed effect model (GLMM, package lme4 in R, Bates, 2011). Observed RTs were log transformed and a normal distribution with identical link function was used. Subject's class, revealed by the finite mixture model based on responses, was added in the dataset as a new categorical variable called *Class* to compute its interactions with the three explanatory variables: the block (*Block*), the previous response (switch or repetition, *PrevSwitch*), and

[1] The estimated parameters were initially expressed in the form of the linear term given by the regression equation. They are transposed in probability using the logistic function for an easier interpretation in Table 2. The logistic function or inverse-logit is: $\text{logit}^{-1}(\alpha) = \frac{\exp(\alpha)}{1+\exp(\alpha)}$.
The equation used in *R* was log*RT* ~ (*PrevSwitch* + *Trial* + *Block*): *Class* + *Class* + (1|Subject) – 1. This formulation facilitates comparisons across classes.

Table 4
Reaction times (in milliseconds) estimated by the GLMM for each class.

	Mean RT after a repetition/a switch		RT switch cost[a]	Slope of RT within a block	Between blocks variation in RT
	Block 1	Block 2			
Class 1	1199/1538	1229/1577	**−343***	0.002	34
Class 2	1232/1367	1036/1149	−124	0.004	**−207***
Class 3	2169/1816	1959/1640	**336***	0.0006	**−193***
Class 4	2166/1753	2303/1865	**425***	**0.007***	124
Class 5	1636/1351	1497/1236	**273***	0.0003	−127

Note. Reaction times were estimated at the middle of a block (trial 10.5). Values in bold (and flanked by a star) are significantly different from zero ($|t| > 2$).

[a] Due to the exponential transformation, the switch cost expressed in millisecond is slightly different in block 1 and in block 2 (differences lower than 30 ms), whereas there was no *Block* × *PrevSwitch* interaction in the model. The mean of the two values are provided in the column RT switch cost for convenience.

the trial number within the block (*Trial*). Other models including interaction effects among *Block*, *PrevSwitch*, and *Trial* were also tested, but no significant interactions were found. The model reported here estimated *RT* values after a switch (i.e., when a repetition is expected) and after a repetition (i.e., when a switch is expected) for each block, and the slope of *RT* as a function of trial number for each of the five classes given by M5.

Table 4 provides estimated RTs (in milliseconds) after an exponential transformation to facilitate interpretation and highlight significant within-class variation and slope values. Global between-class differences were also tested by comparing models with and without *Class* as explanatory variable. This comparison revealed a significant effect of class, likelihood ratio test, $\chi^2(4) = 33.7$, $p < 0.01$.

2.2. Profile interpretation

For each profile (class), Fig. 1 provides the estimated probability of a switch after a switch or after a repetition, as well as the slopes of these probabilities in each block. The most likely response pattern was computed by assuming that a switch occurred if its probability was higher than 0.5. It was used to plot estimated RTs in each block.

Class1 (“perseveration”) included 8 children who switched very rarely (6% of switches throughout the task) especially in block 2 (3%). In block 1, a switch occurred more often after a switch (0.34, see Table 2) than after a repetition (0.06), suggesting that those rare switches occurred in series. No significant trend in the probability of a switch within a block was observed and the most probable response pattern for this class was no switch at all. These children clearly showed a perseveration profile. Consequently, RTs were relatively fast and constant across the two blocks. Surprisingly, a significant repetition cost in RT (Table 4) was observed, due to a few observations that mostly occurred at the beginning of the first block, where RTs could have been a bit longer. Alternatively, children may have considered switches as errors, slowing down response on the next trial.

Class 2 (“random switch”) included 12 children. Like class 1 children, their level of accuracy was very low (0.15). In block 1, they switched more often when a repetition was expected than when a switch was required (0.30 and 0.07, respectively). In the second block, they switched more often (0.26) and more accurately (0.38), but still below chance level. In this block, switches occurred as frequently after a repetition as after a switch (0.25 vs. 0.28, Table 2), and did not yield any cost on RTs. Like class-1 children, their responses were fast and even significantly faster in block 2 than block 1 (−193 ms, see Table 4). The probability of a switch never reached 0.5 despite increasing within block 2. Although these children showed a perseveration profile in block 1, they seemed to use a random responding strategy in block 2, which artificially increased their accuracy level and speeded up their responses.

Class 3 (“mainstream class”) was by far the most inclusive with 38 children. Several indicators clearly showed that these children understood and followed the alternation rule. First, accuracy was above chance level in both blocks (0.68 and 0.67). Second, the probability of a switch was significantly higher after a repetition (i.e., when a switch was expected) than after a switch in the two blocks

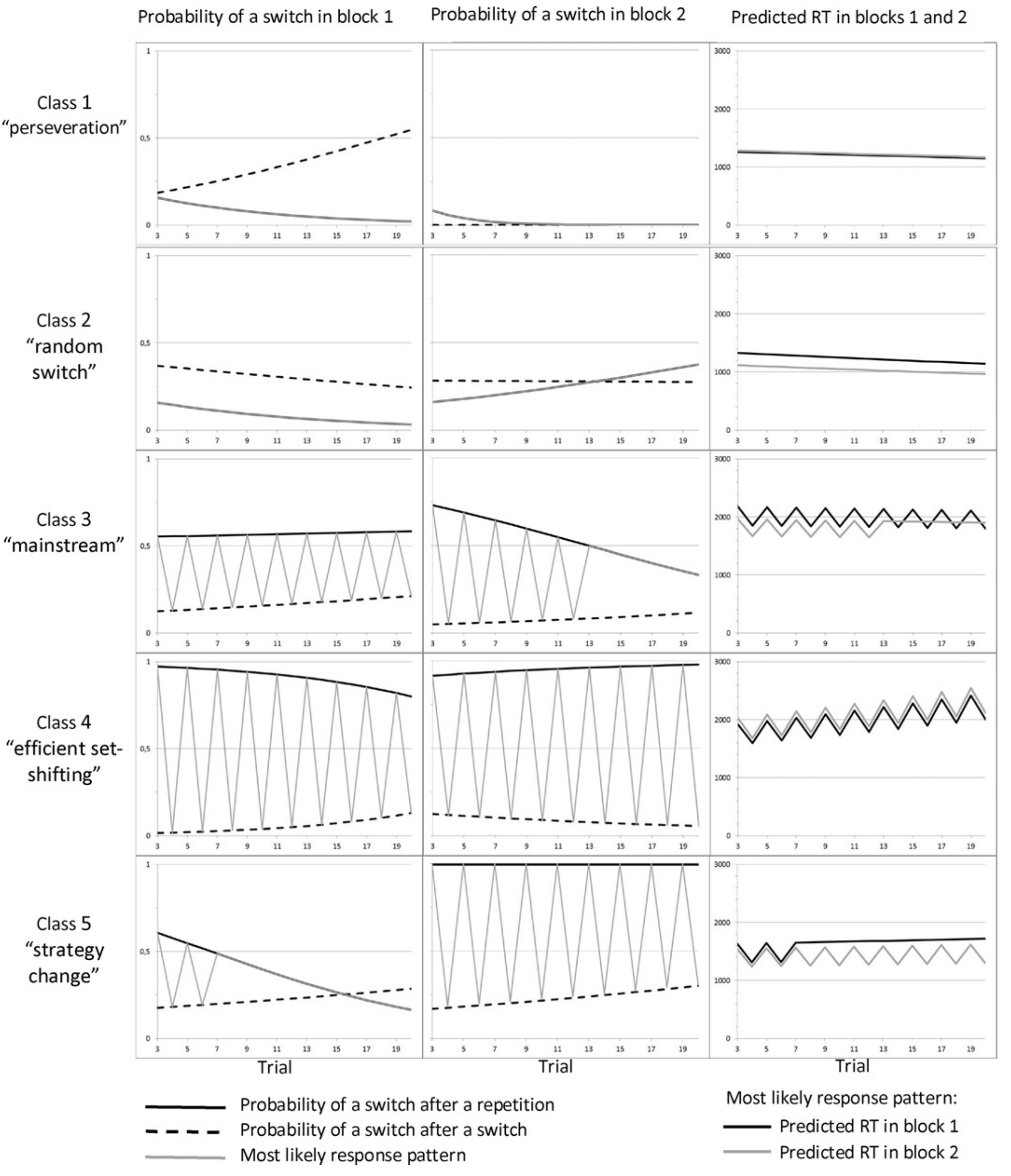

Fig. 1. Estimated probability of a switch and reaction time as a function of previous responses in blocks 1 and 2. The most likely response pattern was computed for each class, assuming that a switch occurred if its probability was higher than .5.

(0.57 and 0.54 vs. 0.16 and 0.08). Third, RTs were slower after a repetition than after a switch (switch cost in RT = 336 ms). Fourth, the most likely response pattern showed an alternation of switches and repetitions, except at the end of block 2. The performance in this class was nevertheless far from perfect. In block 1, the probability of a switch after a repetition was 0.57. This means that in 43% of cases, children did not switch when they were supposed to, whereas most expected repetitions did occur (0.84 in block 1 and 0.92 in block 2). Responses were faster in the second block (−207 ms), where the probability of a switch after a repetition decreased across trials (slope = −0.1, $p < 0.05$), hence suggesting a speed/accuracy tradeoff. This probability was around 0.75 at the beginning of the block, reaching a level below 0.5 at the end of the block (Fig. 1), hence reflecting a decrease in performance across trials. This trend is consistent with the hypothesis of increasing task-goal updating difficulty across trials due to growing interference from past trials.

The 11 children in class 4 ("efficient set-shifting") performed the best. Their response pattern was nearly perfect in both blocks. They showed a high accuracy level (0.93 and 0.94). Most task switches and repetitions occurred in accordance with the alternation rule (Table 2, 92/0.05 in block 1, 0.96/0.08 in block 2). None of the slopes of these probabilities was significant, suggesting that accuracy did not change across trials. Class 4 was the slowest (around 2 s), with the highest switch cost (425 ms). For this class only, RTs increased across trials within blocks. These children handled increasing interference across trials by slowing down responses to maintain accuracy. As in class 3 ("mainstream"), a speed-accuracy trade-off was observed in class 4. However, unlike class-3 children, class-4 children favored accuracy over response speed.

The fifth class ("strategy change") included 10 children and was characterized by highly contrasted profiles in blocks 1 and 2. Although the probability of a switch after a repetition was above 0.5 at the beginning of block 1 (Fig. 1), it rapidly dropped across trials (slope = −0.12, $p < 0.05$), and after a few trials, a perseveration pattern emerged (the probability of a switch was smaller than 0.25 at the end of block 1). Consequently, overall accuracy was at chance level for this block (0.52). In contrast, accuracy was much higher in block 2 (0.87). The probability of a switch after a repetition in block 2 was 0.99, meaning that these children switched when needed. Meanwhile, the probability of a switch after a switch (0.24) was significantly lower than after a repetition, but still relatively far from 0, reflecting a tendency to switch even when a repetition was expected (57% switches in block 2). Despite clearly distinct accuracy patterns across blocks, RT patterns were surprisingly similar in the two blocks: Responses were fast (around 1500 ms) with a significant switch cost (273 ms). No significant interaction appeared between *Block* and *PrevSwitch* or *Trial* on RTs. We nevertheless computed a separate GLM on the RT of class 5 to test for a specific interaction. This model showed no effect of *Trial* but a significant interaction between *PrevSwitch* and *Block*, reflecting the virtual absence of switch cost in block 1 (53 ms) but a significant switch cost in block 2 (578 ms, $p < 0.01$), which is consistent with a qualitative change between blocks.

2.3. Relation of response profiles to age, performances in simple blocks, and working memory

Table 5 provides descriptive statistics and correlations among performance on simple blocks, mixed blocks, age, and working memory performance. Mean accuracy was very high on simple blocks, with little individual variability. This ceiling effect probably explained the relatively low correlation observed between the two simple blocks ($r = 0.282$, $p < 0.05$). Only the second simple block was correlated with the two mixed blocks and Backward Digit Span performance (*BDS*; *r*s = 0.20, 0.24, and 0.23 respectively, $p < 0.05$ for the latter two). The two mixed blocks were highly correlated with each other ($r = 0.69$, $p < 0.01$) and showed a substantial correlation with *BDS* ($r = 0.42$, $p < 0.01$ and 0.30, $p < 0.01$). There was no correlation between *Age* and performance in the Advanced DCCS and BDS, probably because of the narrow age range (65–77 months).

Table 5
Descriptive statistics and bivariate Bravais–Pearson correlations of the performance in simple and mixed blocks, scores in the backward digit span (BDS) and age.

	M	SD	Simple 1	Simple 2	Mixed 1	Mixed 2	BDS
Simple 1	0.953	0.080	1				
Simple 2	0.946	0.090	0.282*	1			
Mixed 1	0.592	0.274	0.22	0.204	1		
Mixed 2	0.656	0.267	0.127	0.241*	0.691**	1	
BDS	2.139	0.747	0.089	0.228*	0.416**	0.298**	1
Age	70.91	3.324	−0.097	−0.121	−0.049	−0.031	0.032

* $p < 0.05$.
** $p < 0.01$.

Table 6
Mean score in the backward digit span task and simple blocks, age and parameters of the multinomial GLM by class.

	Age		Simple blocks		BDS		Model parameters	
	M	SD	Block 1	Block 2	*M*	SD	Intercept (*p*)	Coefficient (*p*)
Class 1	72	3.63	0.96	0.912	1.88	0.79	DF = 8	AIC = 223
Class 2	71.43	3.55	0.925	0.908	1.63	1	0.95 (0.18)	−0.31 (0.28)
Class 3	70.5	3.41	0.95	0.95	2.24	0.55	−0.05 (0.48)	0.77 (0.08)
Class 4	71.91	2.98	1	0.99	2.64	0.64	−5.09 (0.01)	2.33 (<0.01)
Class 5	70.22	3.23	0.94	0.95	2.05	0.8	−0.36 (0.39)	0.3 (0.32)

Note. No parameter is estimated for class 1 because it serves as the reference class.

Using classes as an explanatory variable to predict performance in the simple blocks (comparison of GLMs with a binomial distribution), we observed a significant effect of *Class*, χ^2 (4) = 21, p = 0.0003, but no significant differences between blocks and no significant *Class* by *Block* interaction. Children in class 2 ("random switch") performed the lowest in both simple blocks whereas children of class 1 ("perseveration") encountered difficulties especially in the second simple block. These results are compatible with the above profile interpretation. Classes did not differ in age.

The last model tested the relation between working memory and the probability of belonging to each class. The explanatory variable was quantitative (*BDS*) and the outcome categorical (*Class*). In this case, multinomial GLM is an elegant alternative to inverting the status of the variable in an ANOVA. By comparing two models, with and without *BDS* as an explanatory variable, it is possible to test the effect of *BDS* on the probability of belonging to each class. The likelihood ratio test of this comparison was significant, χ^2 (4) = 17.7, p = 0.005, indicating a significant relation between *BDS* and *Class*. The interpretation of the estimated parameters provided in Table 6 is not straightforward because the influence of each coefficient depends on the values of the others, but they can be used to plot the probability of belonging to each class as a function of *BDS* as shown in Fig. 2.

Most of the children with the highest working memory scores belonged to class 4 ("efficient set-shifting"), which corresponds to the best performers in the Advanced DCCS. Children whose performance was around the mean in BDS mostly belonged to class 3 ("mainstream"). At the bottom of the BDS score range, we found children of class 2 ("random switch"). Classes 1 and 5 were seemingly unrelated to *BDS* performance.

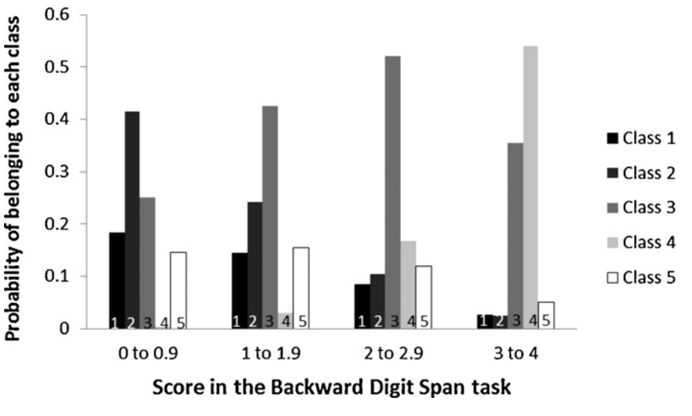

Fig. 2. Probability of belonging to each class as a function of the score in the Backward Digit Span task (BDS). For any given BDS score, the sum of the five probabilities is 1. Class 1 = "perseveration". Class 2 = "random switch". Class 3 = "mainstream". Class 4 = "efficient set-shifting". Class 5 = "strategy change".

3. Discussion

We used statistical modeling to examine children's strategies underlying task goal updating in an alternating-run version of the Advanced DCCS. The results revealed five distinct response profiles that differed in how performance (i.e., switch probabilities) changed across trials and blocks. These profiles were associated with distinct patterns of response times and working memory scores, but did not vary with age (within a small 12-month age range). In all, these results point to substantial variability in the strategies that support goal representation updating for flexible behaviors, although inference about the exact nature of those strategies is speculative and awaits confirmation by experimental manipulation.

Statistical modeling using a mixture of generalized linear models and generalized linear mixed model, provided in widely available packages, allowed a fine-grained study of behavioral dynamics both within subjects across trials and across blocks, and between subjects. Further, it highlighted qualitative differences between classes of children, making it a valuable instrument to study strategy variability. This modeling approach allows parallel analysis of response accuracy and RTs, as well as insightful graphical representation of behavioral variability, and combines goodness of fit and parsimony to determine the most appropriate number of classes to describe a sample of data. The number of classes obtained, however, can depend on the amount of information available and thus number of participants. An increase in the sample size can provide a finer classification into sub-classes by revealing more subtle but significant differences between children within a class. These very flexible statistical tools offer new ways to simultaneously identify individual differences, temporal dynamics and situational influences on behavior and cognition.

In the present study, task-goal updating entirely depended on endogenous processes. Such situations are probably among the most executively demanding and, not surprisingly, about 25% of 5-year-olds (those in the "perseveration" and "random switch" classes) failed to flexibly switch between tasks under such circumstances. Latent class analysis revealed that these children were not homogeneous in their response profiles. Children from the "perseveration" class tended to perseverate on a single task across all trials, hence switching very infrequently. By contrast, children from the "random switch" class seemed to respond randomly (in block 2), thus switching (whether intentionally or not) more frequently but independently of the expected task sequence. The heterogeneity of these response profiles is suggestive of distinct underlying cognitive processes. Perseveration may be especially likely when children fail to understand the necessity of updating task goals from the task sequence, whereas random responding/switching may relate to difficulty using the task sequence to correctly determine when to switch despite the understanding that tasks need to be switched occasionally. Indeed, recent evidence shows that children experience difficulty monitoring for the necessity to switch even in externally cued task-switching contexts (Chevalier et al., 2011). Consistent with this interpretation, children in the "random switch" class were likely to performed poorly on the backward digit span task, hinting at working memory capacity too low to update and maintain task goals consistently, whereas perseveration was not related to BDS, suggesting an unrelated and qualitatively different source of error.

Although some children struggled to switch tasks, most 5-year-olds (the 75% in the "mainstream", "efficient set-shifting", and "strategy change" classes) understood and at least partially implemented the task sequence, suggesting that most 5-year-olds can successfully update task goals and switch tasks accordingly, even when no external cues are available. What cognitive processes did children use to update and maintain task goals? We hypothesized that goal updating and maintenance rely on the use of information about the previous couple of trials. Children could compare the tasks (color or shape) performed on the previous trials and, if they match, switch tasks on the upcoming trial. Alternatively, children could keep track of both the last task performed and its position in the task sequence (i.e., whether it was performed for the first or second time). Either way, consistent goal updating and maintenance require keeping track of information on immediately preceding trials. We hypothesized increasing difficulty across trials due to growing interference from accumulating traces from past trials. Consistent with this claim, we observed that the probability of correct switching (i.e., switching after a repetition) progressively decreased across trials in the "mainstream" class (block 2) and "strategy change" class (block 1), while latency progressively increased across trials in the "efficient set-shifting" class (both blocks). Therefore, task-switching performance does change as trials unfold, further pointing out the interest of statistical modeling to capture temporal dynamics across trials (Chevalier et al., 2009).

Interestingly, increasing interference across trials differentially affected speed/accuracy tradeoff as a function of response profiles. In the "efficient set-shifting" class, children maintained accuracy at the expense of longer responses, whereas the opposite held true for the "mainstream" and "strategy change" classes. Our findings are consistent with a report of speed/accuracy tradeoffs in children's performance on task-switching paradigms (Davidson, Amso, Cruess Anderson, & Diamond, 2006), but they go a step further by showing that children seem to differentially handle such tradeoffs depending on their working memory capacity. Children from the "efficient set-shifting" class were likely to outperform those from the other classes on the backward digit span task. Their high working memory capacity may have allowed them to correctly maintain information in the face of large amounts of interference, but at the cost of longer responding. By contrast, large amounts of interference may have exceeded the more modest working memory capacity of children from the "mainstream" and "strategy change" classes, hence necessarily reducing accuracy but leaving children the possibility of maintaining response speed.

The response pattern of children from the "strategy change" class in the second block of trials was striking. These children showed very high accuracy, but with a relatively high probability of switching after a switch. Unlike children from the "efficient set-shifting" class, high accuracy was accompanied by relatively fast response times. What is more, high accuracy was maintained throughout the block without increasing latency, hence showing no speed/accuracy tradeoff or sensitivity to increasing interference across trials. These findings suggest that these children implemented a strategy of a completely different kind than that used by children from the "efficient set-shifting" class. In particular, this strategy probably did not rely on information about the previous trials (since performance was not affected by increasing interference across trials) and was relatively fast to implement, consistent with the absence of relation with BDS score for this class. These children may have estimated the time spent on the current task or used the rhythmicity of the task name sequence ("color, color, shape, shape") to update task goals, both of which would not require maintenance of information on previous trials, hence tentatively explaining why this response pattern was not strongly associated with working memory. The strategy employed on block 2 by children in the "strategy change" class" likely differs from that used by children in the "mainstream" and "efficient set-shifting" classes, highlighting the variability in task-goal updating strategies.

Not only did response profiles – and underlying strategies – vary across children; they also changed between blocks of trials within children. Such changes are especially conspicuous in the "mainstream" and "strategy change" classes. The hierarchical competing systems model (HCSM; Marcovitch & Zelazo, 2006, 2009) may help account for those within-children changes in strategy between blocks. This model suggests that executive control development relates to an increase in the level of conscious reflection on one's own mental representations and behaviors. Conscious reflection leads to novel representations and allows one to override former, less efficient or irrelevant responses. HCSM posits that the likelihood of conscious reflection increases with task experience as well as when new information can lead to a better appreciation of the task constraints and affordances. The main event occurring between the blocks in the present study was the repetition of task instructions, including the task sequence. The repetition of task instructions in conjunction with children's now substantial experience with the task (through block 1) probably helped children realize their initial strategy was inefficient and prompted conscious reflection that led to new strategies. Conscious reflection may have been especially likely for children in the "mainstream" and "strategy change" classes (3 and 5) because they performed at a medium level of accuracy (better than the "perseveration" and "random switch" classes but lower than "efficient-set shifting" class), leaving room for improvement. Children in the "strategy change" class may have found a new strategy with lower working memory demands. Children in the "mainstream" class may have used a new strategy that allowed more accurate responding on the first few trials of the block but was probably too demanding on working memory and thus susceptible to interference, as suggested by performance decrease across trials in the second block.

The variability in response profiles observed in the present study suggests that, within a given age group, multiple strategies coexist and some of these may be of similar efficiency. Although our data did not allow any conclusion on strategy variations with age, observing within age-group variability leaves open the possibility that developmental paths are more gradual, tortuous and diverse than a clear-cut transition from perseverative behavior to adaptive flexible switching (van Bers et al., 2011). Clearly, the field of executive control is now ripe for a shift from theories that describe its development as linear to views that emphasize both inter- and intra-individual variability and capture multiple developmental pathways (Chevalier et al., 2010; Hanania, 2010; Moriguchi & Hiraki, 2011; Morton et al., 2009). Making sense of the observed variability requires statistical tools that simultaneously allow latent classes analyses and the search for potential explanatory variables. We contend that the future of research on cognitive development lies in a closer consideration of variations both within- and between individuals, with statistical modeling becoming a critical tool for developmental scientists.

References

Aitkin, M. (1999). A general maximum likelihood analysis of variance components in generalized linear models. *Biometrics*, *55*(1), 117–128.

Aitkin, M., & Alfo, M. (2003). Longitudinal analysis of repeated binary data using autoregressive and random effect modelling. *Statistical Modelling*, *3*(4), 291–303.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 723–716.

Altmann, E. M. (2004). Advance preparation in task switching. *What work is being done? Psychological Science*, *15*, 616–622.

Altmann, E. M. (2007). Comparing switch costs: Alternating runs and explicit cuing. *Journal of Experimental Psychology: Learning Memory, and Cognition*, *33*, 475–483.

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task-switching. *Journal of Experimental Psychology: General*, *130*, 641–657.

Bates, D. (2011). *Lme4*: *Linear mixed-effects models using S4 classes*, R package version 0.999375-39, Retrieved from http://www.R-project.org.

Blaye, A., & Chevalier, N. (2011). The role of goal representation in preschoolers' flexibility and inhibition. *Journal of Experimental Child Psychology*, *108*, 469–483.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*, 386–404.

Chevalier, N., & Blaye, A. (2008). Cognitive flexibility in preschoolers: The role of representation activation and maintenance. *Developmental Science*, *11*, 339–353.

Chevalier, N., & Blaye, A. (2009). Setting goals to switch between tasks: Effect of cue transparency on children's cognitive flexibility. *Developmental Psychology*, *45*, 782–797.

Chevalier, N., Blaye, A., Dufau, S., & Lucenet, J. (2010). What visual information do children and adults consider while switching between tasks? Eye-tracking investigation of cognitive flexibility development. *Developmental Psychology*, *46*, 955–972.

Chevalier, N., Dauvier, B., & Blaye, A. (2009). Preschoolers' use of feedback for flexible behavior: Insights from a computational model. *Journal of Experimental Child Psychology*, *103*, 251–267.

Chevalier, N., Wiebe, S. A., Huber, K. L., & Espy, K. A. (2011). Switch detection in children's cognitive flexibility. *Journal of Experimental Child Psychology*, *109*, 353–370.

Davidson, M. C., Amso, D., Cruess Anderson, L., & 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.

Deák, G. O. (2000). The growth of flexible problem solving: Preschool children use changing verbal cues to infer multiple word meanings. *Journal of Cognition and Development*, *1*, 157–191.

Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language*, *48*, 148–168.

Faraway, J. J. (2005). *Extending the linear model with R: Generalized linear mixed effects and nonparametric regression models*. New York: Chapman & Hall/CRC.

Grün, B., & Leisch, F. (2008). Finite mixtures of generalized linear regression models. In C. Shalabh, & Heumann (Eds.), *Recent advances in linear models and related areas* (pp. 205–230). Berlin: Springer.

Hanania, R. (2010). Two types of perseveration in the Dimensional Change Card Sort task. *Journal of Experimental Child Psychology*, *107*, 325–336.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., et al. (2010). Control and interference in task switching – A review. *Psychological Bulletin*, *136*, 849–874.

Lindsey, J. K. (1997). *Applying generalized linear models*. New York: Springer.

Marcovitch, S., Boseovski, J. J., & Knapp, R. J. (2007). Use it or lose it: Examining preschoolers' difficulty in maintaining and executing a goal. *Developmental Science*, *10*, 559–564.

Marcovitch, S., Boseovski, J. J., Knapp, R. J., & Kane, M. J. (2010). Goal neglect and working memory capacity in 4- to 6-year-old children. *Child Development*, *81*, 1687–1695.

Marcovitch, S., & Zelazo, P. D. (2006). The influence of number of A trials on 2-year-olds' behavior in two A-not-B-type search tasks: A test of the hierarchical competing systems model. *Journal of Cognition and Development*, *7*, 477–501.

Marcovitch, S., & Zelazo, P. D. (2009). A hierarchical competing systems model of the emergence and early development of executive function. *Developmental Science*, *12*, 1–18.

Moriguchi, Y., & Hiraki, K. (2011). Longitudinal development of prefrontal function during early childhood. *Developmental Cognitive Neuroscience*, *1*, 153–162.

Morton, J. B., Bosma, R., & Ansari, D. (2009). Age-related changes in brain activation associated with dimensional shifts of attention: An fMRI study. *NeuroImage*, *46*, 249–256.
Morton, J. B., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental Psychobiology*, *40*, 255–265.
Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*, 207–231.
Saeki, E., & Saito, S. (2004a). Effect of articulatory suppression on task-switching performance: Implications for models of working memory. *Memory*, *12*, 257–271.
Saeki, E., & Saito, S. (2004b). The role of the phonological loop in task switching performance: The effect of articulatory suppression in the alternating runs paradigm. *Psychologia*, *47*, 35–43.
Saeki, E., & Saito, S. (2009). Verbal representation in task order control: An examination with transition and task cues in random task switching. *Memory & Cognition*, *37*, 1040–1050.
Siegler, R. S. (1997). Concepts and methods for studying cognitive change. In E. Amsel, & K. A. Renninger (Eds.), *Change and development: Issues of theory, method, and application* (pp. 77–97). Mahwah, NJ: Lawrence Erlbaum Associates.
Snyder, H. R., & Munakata, Y. (2010). Becoming self-directed: Abstract representations support endogenous flexibility in children. *Cognition*, *116*, 155–167.
van Bers, B. M. C. W., Visser, I., van Schijndel, T. J. P., Mandell, D. J., & Raijmakers, M. E. J. (2011). The dynamics of development on the Dimensional Change Card Sorting task. *Developmental Science*, *14*, 960–971.
Vandierendock, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*, *136*, 601–626.
Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, *1*, 297–301.