File ID	152779
Filename	4 Development of Reversal Shift Learning: An Individual Dif- ferences Analysis

SOURCE (OR PART OF THE FOLLOWING SOURCE):

Type	Dissertation
Title	A model-based analysis of the development of learning
Author	V.D. Schmittmann
Faculty	Faculty of Social and Behavioural Sciences
Year	2009
Pages	200

FULL BIBLIOGRAPHIC DETAILS:
http://dare.uva.nl/record/321389

4

Development of Reversal Shift Learning: An Individual Differences Analysis

The existence of distinct modes in category learning is posited in several theories (Ashby et al., 1998; Kéri, 2003). In the initial phase of discrimination shift learning tasks (DSL), which is a subtype of category learning tasks, different modes of learning have been found (e.g., Block et al., 1973; Raijmakers et al., 2001; Schmittmann et al., 2006). In this chapter, the existence and age-related changes of distinct modes in the reversal shift phase of a DSL task were examined, and the relation between the initial learning and the shift learning mode was analyzed. Hidden Markov models were employed to investigate the development in the cross-sectional sample of 4 to 20 year-old participants (Raijmakers et al., 2001; Schmittmann et al., 2006). The results revealed the existence of different shift learning modes in the sample, which were similar to the initial rational and slow learning modes. The probability of slow shift learning decreased with age. The response accuracy in the application of the reversed rule increased with age in slow and fast shift learners. The relation between slow and rational initial, and slow and fast shift learning modes showed age effects.

4.1 Introduction

Everyday life requires us to interpret information from the environment, to monitor our actions, and to adjust them to changes in the environment. This aspect of behavior, often referred to as 'executive function', undergoes significant changes during development, as evidenced by tasks , in which participants are required to adapt their behavior to rules that shift without announcement (Crone, Ridderinkhof, et al., 2004; Chelune & Baer, 1986; Esposito, 1975). In a discrimination shift learning (DSL) task, participants first learn by trial and error to discriminate stimuli according to features on one relevant dimension of multidimensional stimuli (e.g., color and brightness). When a learning criterion is reached, the reward contingencies are shifted, and participants learn a new discrimination. Different types of shifts have been presented.

In a reversal shift (R), all previous stimulus-reward contingencies are reversed, while the relevant dimension remains unchanged. In a non-reversal shift (NR), a previously irrelevant dimension is selected as the new relevant dimension, and stimulus-reward contingencies are changed for half of the stimuli. In R and NR shifts, exactly the same stimuli are used in the shift phase as in the initial learning phase. In intradimensional shifts (ID) and extradimensional shifts (ED), different stimuli are used in the shift phase than in the initial learning phase. In ID shifts, a discrimination is learned on the previously relevant dimension, while in ED shifts, the new discrimination is learned on a previously irrelevant dimension.[1]

While ID shifts are consistently learned faster than ED shifts in children, adults, and also in different species, the consistent superiority of R shifts over NR shifts emerges during development. Studies with young children report inconsistent results, some showing R superiority, some NR superiority (Esposito, 1975). The use of different paradigms has been put forward as reason for the inconsistent findings, but even when only studies of the same paradigm were taken into the comparison the results are still inconsistent (Esposito, 1975). A few connectionist models of shift learning have been proposed (e.g., Kruschke, 1996; O'Reilly, Noelle, Braver, & Cohen, 2002), and relevant aspects of the development of shift learning have been replicated in a cascade-correlation network (Sirois & Shultz, 1998).

However, research into the development of shift learning has been dominated by methods that require the assumption of age group homogeneity. Where age group heterogeneity was examined, this was done by classifying participants according to covariates, e.g., the dominance of the involved dimensions, and carrying out the analysis under the assumption that the age groups are homogeneous, conditional on the covariates. One exception is (Kendler & Kendler, H. H., 1959), who split up the sample into fast and slow learners according to the median number of trials to criterion on the initial learning phase, and compared the ease of NR to the ease of R shifts between these two groups. Reversal superiority was found in the fast learners while nonreversal superiority was found in the slow learners. However, the stimuli used in the shift phase only differed on the new relevant dimension on a given trial, which facilitates NR learning (Esposito, 1975).

[1] In optional shifts, the new discrimination is learned to criterion on a subset of stimuli, and generalization to the other subset of stimuli is tested subsequently. We will not discuss this paradigm.

The presence of subgroups that show qualitatively different behavior could be a reason for the inconsistent results found under the assumption of group homogeneity. In other words, different proportions of two latent groups, e.g., a R superior group and a NR superior group, would result in overall R superiority (given a larger proportion of R superior subgroup), overall NR superiority (given a larger proportion of NR superior subgroup), and overall no significant difference between R and NR shifts (given equal proportions of both subgroups).

As the initial phase of DSL tasks can be seen as a simple category learning task, several theories of category learning (e.g., Ashby et al., 1998; Kéri, 2003) predict the existence of different learning modes. The presence of two different learning modes in a sample can lead to the manifestation of different observable subgroups. Ashby et al. (1998) reviewed findings of different categorization learning paradigms, and the extensive evidence of neuropsychological (imaging and lesion) studies. They posited the existence of two neurologically separated and competing systems of category-learning: a verbal system and an implicit, i.e., procedural learning based system. The verbal system depends heavily on frontal and temporal language areas, whereas a key structure in the implicit system is the striatum. Both systems are assumed to operate in parallel within an individual. Depending on the individual, the task, and the environment, one system may dominate the other. For example, the verbal system is predicted to dominate the implicit system in normal adults on learning tasks, in which the optimal rule is easy to verbalize. Ashby et al. (1998) assume that the verbal system is less efficient in children, as its rule switching component relies on the prefrontal cortex which is not yet fully developed (e.g., Diamond & Goldman-Rakic, 1989). They therefore predict that children's performance is impaired on rule-based category-learning tasks, in which the rules are easily verbalized.

Different subgroups have been found in children and adults learning an initial discrimination (Kendler, 1979; Raijmakers et al., 2001; Schmittmann et al., 2006; Block et al., 1973). The learning processes of the adult participants and of a proportion of children, that increased with age, were well described by a model of fast learning by means of hypothesis testing (Schmittmann et al., 2006). This rational learning mode is expected to rely on the verbal system of category-learning within the theory of (Ashby et al., 1998). The verbal system is an explicit, hypothesis-testing system that involves working memory and executive attention, and relies primarily on the anterior cingulate, the prefrontal cortex, and the head of the caudate nucleus (Ashby & Maddox, 2005). More specifically, the selection of hypotheses is thought to rely on cortical structures, the anterior cingulate, and possibly on the prefrontal cortex, while a key structure in switching of hypotheses is the head of the caudate nucleus. In addition, the involvement of working memory seems to imply the activation of lateral prefrontal cortex and the head of the caudate nucleus (Ashby & O'Brien, 2005). The learning process in the remaining participants was best described by a model of slow, sudden learning, which can be interpreted as a an inefficient form of hypothesis testing, or, alternatively, as a specific type of implicit learning (Schmittmann et al., 2006).

The present study builds forth on the analysis of the initial learning phase described in (Schmittmann et al., 2006). Relevant results are briefly summarized in Section 4.4.1. In this chapter, the relation between the learning mode during the initial learning phase and the learning process during the shift phase is examined. Exploratory and confirmatory hidden Markov models were fitted to the trial-by-trial learning data to obtain a description of the number of subgroups, and the properties of the learning processes, and to statistically test hypotheses about these processes (Wickens, 1982a). The main questions are whether different modes of learning exist during the shift learning phase, in which way these modes are similar and different to the two learning modes in the initial learning phase, and what the conditional probabilities are of using either learning mode, given the fast or slow learning mode during the initial learning phase. In other words, we are interested, whether we find positive and/or negative transfer in the slow and fast learning component. Negative transfer in the fast learning component can not readily be explained by traditional theories of DSL, neither by the spontaneous overtraining hypothesis (Sirois & Shultz, 2006). In addition, age-related changes in the probability of using either mode are examined in a cross-sectional sample covering the age range, in which shift learning undergoes considerable changes.

4.2 Method

4.2.1 Participants

The final sample consisted of 230 participants (for detailed information on selection criteria see Schmittmann et al., 2006). Table 4.1 shows information about the four age groups, which were chosen in such a way that each age group contained a sufficient number of participants for parameter estimation.

Table 4.1: Number of participants in Age by Condition groups.

Age group	Age (yrs)	Shape	Brightness	Total
1	4-6	34 (0.62)	34 (0.76)	68 (0.69)
2	7-9	39 (0.56)	45 (0.96)	84 (0.77)
3	10-13	23 (0.86)	29 (0.93)	52 (0.90)
4	20	11 (1.00)	15 (1.00)	26 (1.00)
		107 (0.69)	123 (0.92)	230 (0.80)

Note. Between brackets the proportion that reached criterion.

4.2.2 Task and procedure

A two-choice discrimination learning task was employed, in which stimuli differed on two binary-valued dimensions: brightness (black or white) and shape (triangle or square). One of the four values was chosen randomly for each participant. Stimuli were presented in pairs of two on a computer screen. The participants were instructed to choose either the left or the right stimulus by pressing pre-assigned keys of a keyboard and to make as many correct choices as possible. Feedback was presented after each trial. As soon as a participant fulfilled the learning criterion of 9 correct out of 10 consecutive trials, the reinforcement contingencies underwent a reversal shift, i.e., the previously incorrect value of the relevant dimension was now the correct choice. For participants, who did not reach the criterion during 48 trials, the task was terminated (for detailed information see Raijmakers et al., 2001; Schmittmann et al., 2006).

4.3 Learning models

The learning processes were analysed by fitting exploratory and confirmatory hidden Markov models (Wickens, 1982a). The confirmatory models were based on different assumptions on the number of different learning modes, the properties of the underlying learning processes, and the relation between covert state and overt responses of different states in the learning processes. The Markov models that we used, are characterized by a finite set of states, and transitions between states. At each trial, a participant may be in one of these states of the learning process. From one trial to the next, the participant may either remain in the same state or switch to a different state. These transition probabilities are conditional on the present state, with lag-one memory. Each state is associated with a

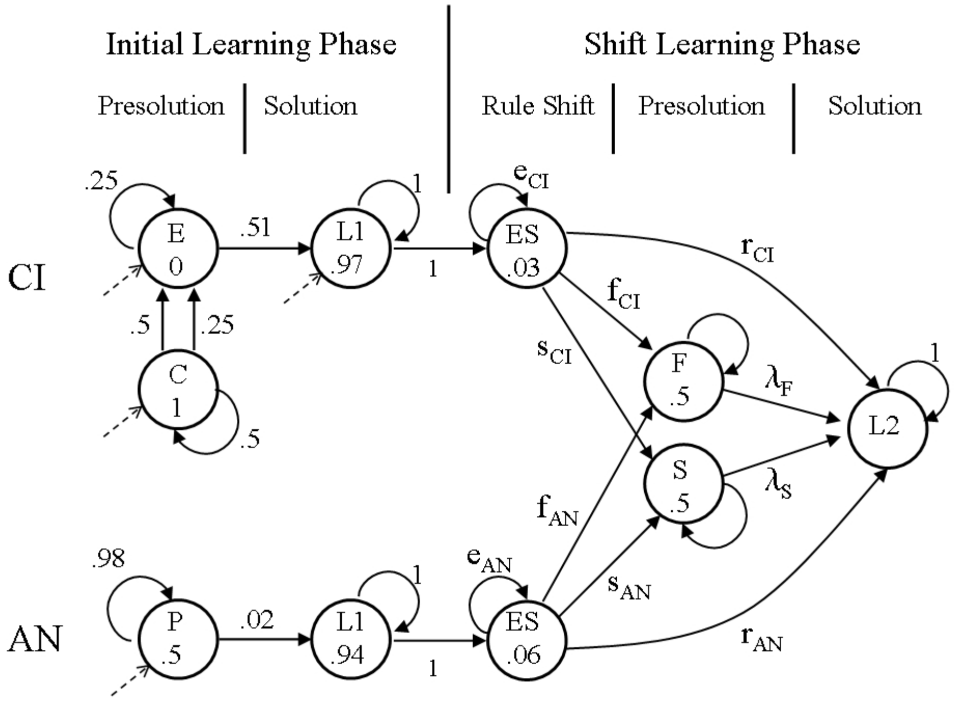

FIGURE 4.1: INITIAL LEARNING AND SHIFT RFS MODEL. CIRCLES DENOTE STATES; VALUES WITHIN CIRCLES INDICATE FIXED OR EXPECTED PROBABILITIES OF A CORRECT RESPONSE. DASHED ARROWS INDICATE POSSIBLE INITIAL STATES; SOLID ARROWS INDICATE POSSIBLE TRANSITIONS. VALUES ADJACENT TO ARROWS INDICATE FIXED OR EXPECTED TRANSITION PROBABILITIES. FOR DETAILS SEE SECTION 4.3.

probability π of a correct response. For instance, in a guessing state (G), π_G is expected to equal 1/2 in the present task. The model parameters, i.e., transition probabilities, observation probabilities, and initial probabilities were estimated in the R-package depmix (Visser, 2005). Multiple sets of starting values were used. The following sections contain descriptions of the models, which were used to formalize the initial learning processes, the shift learning processes, and the relation between initial and shift learning.

4.3.1 INITIAL LEARNING

RATIONAL LEARNING MODE

For the rational learning strategy, a latent version of the concept-identification (*CI*) model with a small modification was used (Bower & Trabasso, 1964;

Kendler, 1979). According to the CI model, a participant learns by testing simple hypotheses, e.g., "black is correct", or "triangle is correct", until the correct hypothesis is found. Following an error, the participant chooses randomly a hypothesis from the subset of all hypotheses that predict the occurrence of the last error. Once a participant has identified the correct hypothesis, he or she will continue to apply this hypothesis, as negative reinforcement fails to occur. The upper left panel of Figure 4.1 shows a graphical representation of the CI model. The CI model consists of the following three states, depicted as circles in the figure: (1) a learned state $L1$, in which a participant applies the correct hypothesis, and two pre-solution states, E and C, in which a participant applies an incorrect hypothesis and chooses either the incorrect stimulus (E), or the correct stimulus (C). It is assumed that a participant only switches hypotheses after committing an error. Participants switch from state E to the learned state $L1$ with learning probability λ_{CI}. Since two of the four (or six hypotheses, if a participant considers location as additional dimension) hypotheses remain after an error, the probability of choosing the correct hypothesis equals 1/2 (or 1/3) under the assumption that all hypotheses are equally likely. The remaining incorrect hypotheses lead to an error response with a probability of 1/2, as the wrong hypothesis in the relevant dimension (e.g., black, while white is correct) is always excluded. Participants start with the correct hypothesis with a probability of the learning parameter divided by two, and start in state E with probability 1/2, as half of the hypotheses result in an error response. The probability of committing an error in the learned state is estimated.

Slow learning mode

The slow learning mode was modeled with the all-or-none model of sudden learning (AN; (Wickens, 1982a)), which is depicted in the lower left panel of Figure 4.1. In this model it is assumed that all participants start in a pre-solution state (P), in which they respond at chance level (here the probability of an error equals 1/2), and remain in this state until they switch to a learned state ($L1$), in which they respond at perfect or near-perfect level. The learning probability of switching from P to $L1$ is stationary. Slow learning will result in a small learning parameter λ_{AN}, which is expected to be much smaller than in the CI model. As in the CI model, learned state $L1$ is absorbing, i.e., once it is entered, it cannot be left.

4.3.2 Shift learning

Participants were not aware of a rule shift, nor of its trial of occurrence. Therefore, all participants are assumed to start the shift learning phase by applying the initial rule. In the reversal shift design the application of the initial rule results in an error with the same probability with which a correct response is produced before the rule shift. Participants may not notice the rule change immediately, or maintain the previous rule in spite of the rule change. Upon disengaging from the previous rule, participants may follow three different trajectories from this initial error shift state (ES) to the second learned state, in which the reversal shift has been learned ($L2$). They may (a) choose the correct rule immediately, i.e., reverse the previous rule (R), (b) enter a fast learning presolution state (F), in which hypotheses of the irrelevant dimension are tested, or (3) enter a slow learning presolution state (S) in which the probability of a correct response may be biased (RFS model). In order to examine developmental effects, the probabilities of maintaining the previous rule, and of following the three different trajectories were additionally estimated in each age group. To examine the existence of different learning modes in the shift phase, the fit of the RFS model was compared to 1-mode models in which all participants enter the same presolution state ($P2$), from which they may switch to an absorbing learned state $L2$. The probabilities of switching from ES to $P2$ and from P2 to L2 were estimated separately in each age group (multigroup 1-mode models) or constrained to be equal across age groups (1-group 1-mode models).

4.3.3 Relation between initial and shift learning

The probabilities of maintaining the initial rule, and the probabilities of following the R F, and S trajectories are estimated conditional on the initial learning mode, i.e., e_{CI}, r_{CI}, f_{CI}, and s_{CI} given initial rational learning, and e_{AN}, r_{AN}, f_{AN}, and s_{AN}, given initial slow learning, in order to examine in which way the learning mode in the initial phase is related to the behavior during the shift learning phase. A graphical representation of the whole model is shown in Fig. 4.1

4.4 Results

4.4.1 Initial learning

Fourty-five participants did not satisfy the learning criterion. This result is predicted by learning models for slow learning. As the response accuracy in the learned state did not differ significantly between groups, the failure cannot be attributed to differences in the accuracy of applying a rule consistently. Models with two modes of learning, as compared to 1-mode models (not shown), provided a better fit to the data according to the fit indices AIC and BIC (e.g., Burnham & Anderson, 2002). A mixture model, in which two learning modes are accommodated by estimating mixture proportions of either mode in addition to the mode specific parameters, fitted the data relatively better than models with only one mode of learning, even when the learning parameter in the latter models varied with age. The fast, rational learning mode was best described by a CI model with an expected probability of learning given an initial set of four hypotheses, and with a small probability of a mistake in the learned state. Learning in this mode may be interpreted as a hypothesis-testing process. In the slow learners, sudden learning from one trial to the next was observed, as opposed to incremental learning. This mode was best described by an AN model with a small probability of learning, and a small probability of a mistake in the learned state. The proportion of the rational learning component increased with age (.2, .5, .7, 1, in age groups 1 to 4 resp.; remaining parameter estimates are shown in Fig. 4.1). In the younger age groups the proportion of rational learning was higher in brightness discrimination problems than in shape discrimination problems.

4.4.2 Shift learning

Of the 185 participants who entered the shift phase, 18 (8, 6, and 4 in age groups 1 to 3, resp.) did not reach the second learning criterion. The shift data of one outlying participant in age group 4, who failed to notice the rule change for 32 trials, was removed from the sample. In order to allow for state ES to be interpreted as a state in which the initial rule is applied, the probability of a correct response in this state was fixed to the value .038, which equals 1 minus the probability of a correct response in a single learned state in the initial models. Multi-group and one-group models with one learning mode fitted the data worse than the RFS models, as judged by AIC and BIC. In the one-mode models, the probability of a correct response in the presolution state was fixed to 1/2. Relaxing this constraint resulted in solutions which were essentially 2-mode models.

TABLE 4.2: PARAMETER ESTIMATES OF RFS MODEL

Age Group	r	f	s	π_{L2}
1	0	.214	.598	.893
2	.160	.437	.214	.854
3	.221	.400	.191	.930
4	.364	.447	-	.977

In the RFS models, relaxing the equality constraint on the observation probability of $L2$, π_{L2}, such that a separate response probability was estimated for the learned state in the S trajectory, did not result in a significant improvement in fit. So, in the following the different trajectories all lead to a single learned state $L2$, whose observation probability may differ between age groups. Estimating the observation probabilities π_F and π_S in the presolution states rather than fixing them to guessing probability 1/2 did not result in significant improvements of fit. However, imposing equality constraints on the r, f, and s transition parameters across age groups resulted in a significant decrease in fit, as did the addition of an equality constraint on π_{L2} across age groups. Relaxing the equality constraint on the learning parameters λ_F and λ_S across age groups did not result in a significant improvement of fit. The probability e of remaining in the ES state deviated significantly from zero. In the final model, three transition parameters r, f, s, and the observation probability in $L2$ were estimated in each age group. Additionally, two learning parameters λ_F and λ_S were estimated, but constrained to be equal across groups. The learning parameter in the S trajectory, λ_S, was estimated at .032, and the learning parameter in the F trajectory, λ_F, was estimated at .799. The probability of remaining in ES equals .188. The estimates of the remaining parameters of this model are shown in Table 4.2.

4.4.3 Initial and shift learning

Fig. 4.2 shows a scatterplot of the number of trials to criterion in the initial learning phase by the number of trials to criterion in the shift learning phase. The latter variable may overestimate the trial of learning in the younger age

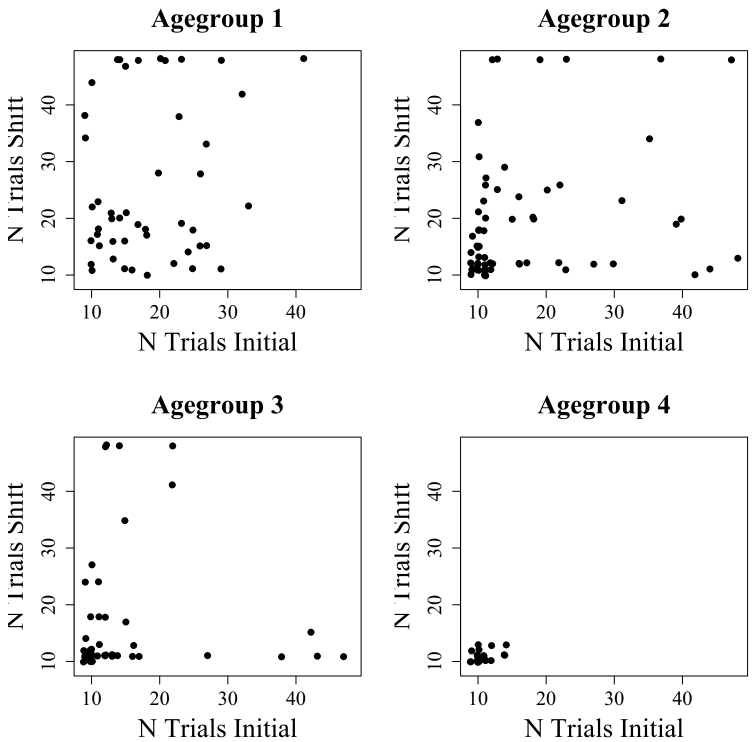

Figure 4.2: Number of Initial by Shift Learning Trials (Random jitter added)

groups, as their observation probabilities in the learned state are estimated at values below the criterion value of .9.

The conditional probabilities of following the R, F and S trajectory in the shift phase, given the learning mode in the initial phase were estimated in models, in which the parameters of the initial phase were fixed to the values estimated in ??.[2] One learning parameter, equal across age groups, was estimated for the R, the F, and the S trajectory, each. In addition, four age specific observation probabilities in the L2 state were estimated. The probability of a correct response in the S and F state were fixed to guessing probability 1/2. In age groups 1 to 3, the probability of following the R, F, and S trajectories were estimated conditional on the initial learning mode. In age group 4, the probabilities of the slow learning modes were fixed to zero. The probabilities e_{AN} and e_{CI} of remaining in the ES states were constrained to be equal to each other and equal across age groups, as relaxing these equality constraints did not result in significant improvements

[2]Preliminary model fits, in which all parameters were estimated simultaneously, resulted in comparable estimates. In order to present sound results of these models, an investigation of their statistical properties is required, as the complexity of these models is increased.

in fit. In the final model, the learning parameter in the F and S states were estimated at .822, and .032, resp., and e was estimated at .205. The remaining parameter estimates are shown in Table 4.3.

TABLE 4.3: PARAMETER ESTIMATES

Age Group	Initial Mode	r	f	s	π_{L2}
1	CI	0	.064	.731	.899
	AN	0	.246	.550	
2	CI	.364	.346	.084	.851
	AN	0	.406	.389	
3	CI	.238	.370	.186	.930
	AN	.205	.400	.190	
4	CI	.382	.414	-	.977
	AN	-	-	-	

4.5 DISCUSSION

The analysis of the sequential data of the shift learning phase revealed the existence of a slow and a fast learning mode, besides the immediate reversal of the previous rule. While learning parameters did not differ significantly across groups, the probabilities of immediately reversing the rule, and of fast and slow learning showed age effects. Age differences were also found in the accuracy with which the new rule was applied in the learned state. The estimates of the learning parameters were larger in the shift phase than in the initial learning phase. Note however that the sample in the shift phase is selected, as only participants who reached the first learning criterion within 48 trials entered the shift phase.

Given slow initial learning, the probability of slow shift learning decreases with age. So, the probability of positive transfer as a discrete improvement in learning efficiency increases with age in the participants, who learn the initial discrimination slowly. An immediate reversal of the previous rule given slow initial learning does barely happen, and it does so only in age group 3. Given fast initial learning, the results show a less clear pattern. The probability of slow shift learning is high in the youngest age group, and it is even higher than the probability given slow initial learning. As the relevant dimension remains the same, dimensional preferences alone, which have been proposed as a possible cause of differences in discrimination learning and shift behavior (Block et al., 1973; Esposito, 1975), cannot account for this finding. Set-maintenance problems do not seem to be a problem as the presolution accuracy did not deviate significantly from the guessing probability. A possible interpretation of this result might be that a group of young children has an initial preference for the correct value of the relevant dimension, which cannot be suppressed in the shift phase. An inefficiency in error processing, which has been shown to improve with age (Crone, Ridderinkhof, et al., 2004), might be a cause, and would be a single explanation accounting for slow learning as well in the initial, as in the shift learning phase.

While the probability of a correct response in the initial learned states did not differ significantly across age groups, these probabilities differed significantly between age groups in the shift learned state. The observation probabilities in the two younger age groups are below criterion, suggesting that the amount of overlearning is not the same for all participants, and that the commonly used number of trials to criterion overestimates the trial of learning in younger children. Using a more liberal learning criterion, e.g., 8/10 correct, comes at the cost of an increased probability of satisfying the criterion when guessing, and therefore does not solve the misclassification problem.[3] A possible explanation of the below-criterion accuracy would be more interference from the previously learned rule in the younger children, resulting in occasional responses according to the previous rule.

In the models that we employed, the presolution probability of a correct response in the slow shift learners (S trajectory) did not differ significantly from guessing probability, nor did the accuracy in the L2 state differ significantly from the fast shift learners' accuracy. In future research the nature of the learning process in the S trajectory should be subjected to a more detailed analysis. An important question is whether the presolution accuracy is stationary, and if not, in which way stationarity is violated. Slow learners may show a monotonic increase in accuracy, or they may show series of trials in which the old rule is applied in alternation with series of trials with an above chance level accuracy. In the initial learning phase, models of sudden learning provided a better model fit than models of incremental learning. However, this finding does not have to generalize to the shift learning phase.

[3]A guessing strategy on each trial results in probabilities of .02 (sd=.01),.15 (sd=.03), and .49 (sd=.03) of satisfying the three criteria 10/10, 9/10, and 8/10, respectively, within a maximum of 48 trials (Based on a simulation with 500 replications of 230 cases each).