1

General introduction
Flexibility is a significant ability in the present rapidly changing society. In a novel situation or when circumstances change, one’s automatic inclinations provide no guidance, and conscious cognitive control is required (Diamond, 2006a). Cognitive flexibility is one of the component abilities that constitute cognitive control (Diamond, 2013; Huizinga, Dolan, & Van der Molen, 2006; Zelazo, Müller, Frye, & Marcovitch, 2003). The main goal of this thesis is to gain insight into the cognitive flexibility of preschoolers. Previous research shows that important improvements in cognitive flexibility occur between the ages of three and five years (Carlson, 2005). We make novel contributions on three important issues in this field. First, we empirically test key questions about the mechanisms underlying performance and development on the Dimensional Change Card Sorting (DCCS) task, a widely used paradigm to study cognitive flexibility in preschoolers (Zelazo, 2006). Secondly, we investigate the dynamics of development of this important ability. Finally, we introduce an analysis method that is more developmentally appropriate than standard ways of analyzing DCCS task data.

This thesis starts with an introduction into the field of cognitive flexibility: What is cognitive flexibility? How do preschoolers perform on a cognitive flexibility task? What are the main theoretical frameworks and computational models proposed to explain preschoolers’ behavior on the DCCS task? It then continues with a discussion of the dynamics of development in cognitive flexibility. The question whether development does proceed in a stage-wise or in a continuous fashion, is a central and recurrent theme in developmental psychology (e.g. Brainerd, 1978; Fisher, Pipp, & Bullock, 1984; Jansen & Van der Maas, 2001). Three models for development on the DCCS task are presented: a continuous developmental model, a step-wise developmental model, and a discontinuous developmental model, as described by formal models of phase transitions (Thom, 1975; Van der Maas & Molenaar, 1992). Next, a new way of analyzing DCCS task data is introduced. Standard analyses of DCCS task data use sum scores. The alternative way of analyzing presented in this thesis is to model trial-by-trial behavioral data of the DCCS task with latent Markov models (Rabiner, 1989; Van de Pol & Langeheine, 1990, Visser, 2011). With these model-based analyses we can distinguish between different modes of behavior, study the stability of these behavioral modes and how these modes
depend on different factors in the environment. Finally, a short introduction of the studies in this thesis is given.

1.1 Cognitive flexibility

Every day you cycle home from work along the same route while thinking of all kinds of things except the route home. For example what happened that day at work, or what to cook for dinner that night. Then one day there is a roadblock and instead of cycling home absorbed in thought, you have to start thinking about an alternative route you can take. What you need in situations like this is conscious cognitive control, also called executive function. Executive function is required whenever going “on automatic” would be insufficient, and especially when it would lead one astray (Diamond, 2006a). Component cognitive abilities that constitute executive function are inhibition, working memory, and cognitive flexibility (Diamond, 2013; Huizinga et al., 2006; Zelazo et al., 2003). Inhibition is the ability to control one’s attention, behavior, thoughts, and/or emotions to suppress responses that are inappropriate in a particular context, and instead do what is more appropriate or needed. Working memory is the ability to hold information in mind while it is processed for further use. Cognitive flexibility is the ability to change plans in response to relevant changes in the environment, and, complementary, to maintain activities when changes in the environment are irrelevant. In order to flexibly adapt to novel situations one needs both working memory (to remember what is needed in the new situation), and inhibition (to suppress the tendency to act according to the old situation) (Diamond, 2013).

Cognitive flexibility is improving importantly between the ages of three and five years, especially the flexibility to change perspectives (Carlson, 2005). This is expressed in diverse cognitive tasks, such as Luria’s tapping task (Diamond & Taylor, 1996), or appearance reality tasks (Flavell, Flavell, & Green, 1986). In Luria’s tapping task, children have to tap once when the experimenter taps twice, and tap twice when the experimenter taps once, and at the same time inhibit the tendency to mimic what the experimenter does, making the opposite response instead. Performance on this task is improving significantly between 3.5 and 4 years of age (Diamond & Taylor, 1996). In appearance-
reality tasks children are presented a situation in which appearance and reality differ, for example a sponge that looks like a rock. Typically 3-year-olds report that it looks like a rock and really is a rock, or that it looks like a sponge and really is a sponge, while 4- and 5-year-olds correctly answer that it looks like a rock but really is a sponge (Flavell et al., 1986). Improvements on cognitive tasks in the preschool years go together with improvements in social cognition. Social cognition is commonly investigated with theory of mind tasks (Premack & Woodruff, 1978; Wimmer & Perner, 1983). Theory of mind implies inferring what another person might know, think, believe, or want, and use that to accurately predict what the other person might do. In a widely used theory of mind task, the Sally-Anne task, children are shown two dolls, Sally and Anne, together with a basket and a box. Sally has a marble, which she places in the basket before she leaves the room. While Sally is out of the room, Anne takes the marble from the basket and puts it in the box. Sally returns, and the child is asked where Sally will look for the marble. The child passes the task when he/she answers that Sally will look in the basket where she put the marble. The child fails the task when he/she answers that Sally will look in the box, where the child knows the marble is hidden. But Sally cannot know this since she did not see it being hidden there. Typically 3-year-olds fail theory of mind tests, whereas 4- and 5-year-olds succeed (Wimmer & Perner, 1983). Improvements on cognitive tasks in the preschool period also go together with marked changes in moral development. While 3-year-olds think that things are either wrong or right, and people are either good or bad, 4- and 5-year-olds know that even wise people can be wrong sometimes, and good people can do things they are ashamed of (Kohlberg, 1963).

A widely used paradigm to study cognitive flexibility in preschoolers is the Dimensional Change Card Sorting (DCCS) task (Zelazo, 2006). In this task, children are required to sort two different test cards according to shape or color on two stacks marked by target cards. Each test card matches one target card on color and the other target card on shape. The task consists of two phases: the pre-switch phase and the post-switch phase. The pre-switch phase starts with two demonstration trials to show children how to sort the test cards, whereupon children are required to sort the two test cards repeatedly according to one dimension (e.g. color). In the post-switch phase of the task
children are asked to sort the same two test cards according to the other dimension (e.g. shape). Children are told explicitly when to switch, but there are no demonstration trials in the post-switch phase. Children are not given feedback on any trial, but the relevant sorting rules are repeated before every trial, and the test cards are labeled with the relevant dimension only (e.g. “This is a red one, where does it go?”). See Figure 1.1 for an illustration of the target- and test cards used in the manual version of the task introduced by Zelazo (2006), and an example of the target- and test cards used in the computerized version of the task that was used in the studies in this thesis.

Figure 1.1 Illustration of the target- and test cards used in the manual version of the DCCS task introduced by Zelazo (2006), and an example of the target- and test cards used in the computerized version of the DCCS task that was used in the studies in this thesis.

Only the data of children who have sorted at least five of the six pre-switch trials correctly is used in the analyses of post-switch performance, because it is impossible to determine if a child can make a switch between rules when he or she can not consistently sort according to the first rule. In the general discussion we refer to the children that did not pass the pre-switch phase in the studies of this thesis. Generally, nearly all children sort correctly in
the pre-switch phase, regardless of which dimension is relevant first. However, 3-year-olds do have problems in the post-switch phase of the DCCS task when they have to switch sorting rules. Typically, 3-year-olds perseverate by sorting the test cards according to the dimension that was relevant in the pre-switch phase, whereas most 4- and 5-year-olds switch immediately to the new rules when asked to do so (Kirkham, Cruess, & Diamond, 2003; Perner & Lang, 2002; Zelazo, Frye, & Rapus, 1996). Still, the difficulty in switching sorting dimension on the DCCS task never disappears completely. Although older children and adults typically sort all cards correctly in the post-switch phase of the task, their response times increase when they have to switch sorting rules (Diamond & Kirkham, 2005).

A number of theoretical frameworks have been proposed to explain preschoolers’ perseverative behavior on the DCCS task. The Cognitive Complexity and Control theory-Revised (CCC-r) assumes that perseverators have difficulties in reflecting on their rule representations, that is formulating and using a higher order rule for selecting which pair of rules (color rules or shape rules) must be used on a particular trial (Zelazo et al., 2003). According to the attentional inertia theory inhibitory control plays a primary role in switching. Perseverators may know the new rules they should be following during the post-switch phase, but fail to inhibit attention to the pre-switch relevant information (Kirkham et al., 2003). Diamond and Kirkham (2005) interpret their results with adults on the DCCS task as attentional inertia that did not completely disappear with age. The activation-deficit or negative priming account assumes that in the post-switch phase perseverators fail to activate information that was automatically inhibited in the pre-switch phase, because it was irrelevant at that time (Chevalier & Blaye, 2008; Müller, Dick, Gela, Overton, & Zelazo, 2006). According to the re-description account perseverators do not understand that a single object can be considered from different viewpoints. They can conceptualize stimuli in a single way, i.e. according to the pre-switch rules, but fail to re-describe the stimuli in another way, i.e. according to the post-switch rules (Perner & Lang, 2002). Finally, the competing memory systems theory supposes that flexible behavior depends on the competition between active and latent memory traces. Perseveration occurs when an active memory trace of the post-switch relevant sorting rules is
not strong enough to compete against a latent memory trace of the pre-switch relevant sorting rules (Morton & Munakata, 2002; Munakata, 1998).

To explain the results of empirical studies with the DCCS task and to pinpoint the verbal theories in the form of a mechanism, a number of computational models of behavior and development on the task have been proposed as well (Buss & Spencer, 2008; Morton & Munakata, 2002; Marcovitch & Zelazo, 2000). These computational models specify mechanisms underlying the development of flexible sorting behavior. In these models, developmental change is formalized in terms of one or more parameters of the model, such as the strength of working memory relative to the strength of latent memory (Buss & Spencer, 2008; Morton & Munakata, 2002), or the impact of the length of training on performance (Marcovitch & Zelazo, 2000). Consequently, these models relate to cognitive and/or neural systems in the brain that play a crucial role in the development of flexible behavior. Such models have been evaluated by comparing the behavior of models with two different parameter settings (representing a younger and an older model) with empirical data from children of different ages. Because these models are defined as dynamical systems, they also have implications for the process of change between a younger and an older model (Van der Maas & Raijmakers, 2009). However, such implications have not been tested explicitly in these models.

The study described in Chapter 2 investigates the dynamics of development on the DCCS task and shows that there are important differences between the implications of the proposed computational models and the empirical data concerning the process of change. In the next section we introduce the dynamics of development more extensively.

1.2 Dynamics of development
In the DCCS task most children either sort none of the six post-switch cards correctly or sort all six post-switch cards correctly. Children's performance on the DDCS task seems to be bimodal. See Figure 1.2 for the frequency distribution of the number of correct post-switch trials in a standard version of the DCCS task. The bimodal distribution of the data would indicate two modes of behavior, and raises the question whether development on the DCCS task
does proceed continuously or discontinuously. The issue whether development proceeds in a stage-wise or in a continuous fashion is a central and recurrent theme in developmental psychology (e.g. Brainerd, 1978; Fisher et al., 1984; Jansen & Van der Maas, 2001).

Figure 1.2  
Frequency distribution of the number of correct post-switch trials in the standard version of the DCCS task (data from the study described in Chapter 2).

Jean Piaget was the major proponent of stage-wise psychological development in which discontinuous developmental change plays an important role (Piaget & Inhelder, 1969). He proposed a theory in which development includes four major stages; the sensorimotor stage, the pre-operational stage, the concrete operational stage, and the formal operational stage. Flavell (1971) noted that stage-to-stage development implies qualitative changes; periods of major reorganization in thought between the stages. The difference between stage-wise development and fast continuous development is hard to make
(Eckstein, 1999, 2000; Van Geert, 1991, 1998), and therefore a formal model of discontinuous development, and empirical criteria that can be applied in research are required.

Such a theoretical model can be found in a branch of mathematics called catastrophe theory (Thom, 1975; Van der Maas & Molenaar, 1992). Formal mathematical models of developmental dynamics, such as catastrophe models, can provide insight into the variables that control the developmental process, without making detailed assumptions about a possible mechanism, such as in computational models. In contrast to computational models, these formal models are descriptive models of the observed behavior. Nevertheless, these models may generate novel and testable predictions about performance on the task and could define additional demands on the computational models. At the same time, computational models typically have many parameters that cannot all be fitted to the data directly, unlike the parameters of a descriptive formal model. The cusp model, which is derived from catastrophe theory, is a formal model of discontinuity in development that provides a set of empirical criteria. In the cusp model, the change in a variable depends on the continuous change in two independent variables. Discontinuous change in many phenomena can be expressed in this model (Zeeman, 1976); the freezing of water being an example. In this example the condition of the water is the dependent variable, and the two independent variables are temperature and pressure. From the cusp model eight criteria, so called catastrophe flags can be derived (Gilmore, 1981). These catastrophe flags are typical properties of behavior that indicate, and sometimes predict, the occurrence of a discontinuous transition. Examples of catastrophe flags are bimodality, critical slowing down, sudden jumps, hysteresis, and divergence.

There is ample evidence for the existence of domain-specific transitions in cognitive development. For example Jansen, and Van der Maas (2001) studied a phase transition in the development of proportional reasoning with the balance scale task by testing catastrophe flags. In the balance scale task children are asked to predict the movement of a balance scale. Clearly this movement depends on the number of weights placed on each side of the scale and the distance of the weights to the fulcrum. Most children use one of four different rules to solve the items of the balance scale task. Children using Rule 1
only focus on the number of weights placed on the balance scale. Children using Rule 2 focus on the distance of the weights to the fulcrum, only when the number of weights placed on both sides of the balance scale is equal. When the number of weights on both sides of the scale is not equal, they focus on the number of weights. Children using Rule 3 and Rule 4 take both dimensions into account when predicting the movement of the scale. Children using Rule 3 add the number of weights and the distance of the weights to the fulcrum, and children using the correct Rule 4 multiply the number of weights and the distance of the weights to the fulcrum. Jansen, and Van der Maas (2001) studied children's transition from Rule 1 to Rule 2 to test whether it shows the properties of a phase transition. They hypothesized that by making the distance dimension more salient children that do not use Rule 2 can be encouraged to adopt this rule (catastrophe flag hysteresis). The salience of the distance dimension was increased by increasing the difference between the distances of the weights to the fulcrum of the two sides of the balance scale. Children received a series of five distance items (the number of weights on both sides of the scale are equal), in which the difference between the distances of the weights to the fulcrum on each side of the scale was gradually increased. A phase transition was observed: Some children who initially used Rule 1 switched to using Rule 2 when the difference between the distances on both sides increased.

Based on the bimodal distribution (one of the catastrophe flags) of accuracy data of the DCCS task in previous research, the possible dynamical models for DCCS developmental change can be restricted to a few general cases: A continuous developmental model with a rapid acceleration, a step-wise developmental model, and a discontinuous developmental model as described by formal models of phase transitions (Thom, 1975; Van der Maas & Molenaar, 1992). According to the continuous developmental model, children use one strategy across all ages, but gradually become better at generating correct answers with age. In a step-wise developmental model, children suddenly jump from one strategy to the other. Importantly, at each single moment during their development children exclusively use one strategy or the other, and once they have learned the correct strategy they stop using the incorrect strategy altogether (but see Scheibehenne, Rieskamp, & Wagenmakers, (2013) for an alternative interpretation). According to a discontinuous developmental model
children also make a sudden jump from one strategy to the other. However, there is a transitional period during which children have both strategies at their disposal and can use both of them within a short period of time. To reveal the dynamics of behavior from empirical data, statistical analyses should test whether the data can best be described in terms of multiple behavioral modes occurring between, but also within individuals. Furthermore, the dynamics of behavior should be expressed in terms of changes between behavioral modes.

In the study described in Chapter 2 we test the three different models for the dynamics of development on the DCCS task in empirical data. In the general discussion we introduce a cusp model for the transition from perseverating to switching on the DCCS task that incorporates the findings of all empirical studies in this thesis.

1.3 Statistical approach

Standard analyses of DCCS task data use sum-scores of the post-switch trials. For the continuous and step-wise developmental models, sum-scores contain all the information that is necessary to describe development. According to these models children use one strategy at a certain time point and, given this strategy, the probability of a correct response is the same on all six post-switch trials. However, according to the discontinuous developmental model, children in transition may use both strategies over a short period of time and, therefore, the probability of a correct response can change over the course of the post-switch trials. In order to distinguish between the three different developmental models of performance on the DCCS task we modeled trial-by-trial behavioral data of the post-switch phase of the task with latent Markov models (Rabiner, 1989; Van de Pol & Langeheine, 1990; Visser, 2011).

In latent Markov models one or more latent states are defined, which are associated with a prototypical pattern of responding. Moreover, with these models we can also quantify possible transitions between latent states over the course of the DCCS post-switch trials. The latent Markov models are defined by a number of parameters that allow us to quantify the nature of the latent states: response probabilities, initial probabilities, and transition probabilities. The response probability is the probability of a correct response, conditional
on being in a certain latent state. The initial probability is the probability of being in a certain latent state at the first trial. The transition probability is the probability of moving to another latent state, conditional on being in a certain latent state. By comparing the goodness-of-fit of different models that are consistent with different developmental trajectories, we can find out which model describes the dynamics of development on the DCCS task most optimal, that is, the most parsimonious but best fitting model. A posteriori, based on the optimal latent Markov model, for any given response pattern, the probability that that pattern was generated from only one state, or from transitioning between states can be computed. These, so called posterior probabilities are used to assign children to the latent subgroup that appears to be most likely. In the empirical studies with the DCCS task described in this thesis we determined for individual children whether they used one stable strategy, or whether they used more strategies within the course of the post-switch trials of the DCCS task.

The use of categorical latent variable techniques such as latent Markov models, for classifying subjects into subgroups, instead of classifying participants by eye-balling or a rule of thumb (e.g. by using sum-scores of post-switch trials and an arbitrary cutoff) has a number of important advantages. First, with these techniques, the criterion for classifying children into subgroups, such as perseverators and switchers is based on sound statistical inference. The optimal choices for cutoff criteria require statistical modeling, to avoid high numbers of false positives and false negatives. This means that one would wrongly conclude that a specific strategy is used by some of the children. Second, the use of latent variable techniques provides the possibility of detecting hitherto unknown groups of participants that show similar behavior. In particular, the parameters of latent Markov models provide an estimate of the consistency of strategy use in children. If the probability correct of a subgroup did not deviate from .5, for example, their scores would not deviate from guessing. Third, the application of latent Markov models also allows for detecting regularities in sequential behavior, which is essential in differentiating dynamics of change. See Van der Maas and Straatemeier (2008) for discussion on the necessity of applying categorical latent variable techniques in classifying strategy use. In the study described in Chapter 2 the
importance of these techniques is illustrated by means of a small simulation study.

1.4 Outline

The main goal of this thesis is to gain insight into the cognitive flexibility of preschoolers. We empirically test key questions about the mechanisms underlying performance and development on the DCCS task. We investigate the dynamics of development concerning this important ability. And we introduce a method that is more developmentally appropriate than standard ways of analyzing DCCS task data.

Chapter 2 describes a study that investigates the dynamics of development on the DCCS task in 3- to 5-year-olds using a computerized version of the standard DCCS task. A continuous, a step-wise, and a discontinuous developmental model are compared by fitting latent Markov models to the trial-by-trial accuracy data of the post-switch phase of the task. Consequences of the results of the study for the computational models proposed in literature that explain behavior and development on the DDCS task are discussed. Finally, a substantive interpretation of a dynamical model incorporating the findings of the study that could be used to test specific hypotheses about the variables that control development of DCCS performance is introduced.

Chapter 3 concerns an experiment that investigates the abstractness of children’s rule representations in the pre-switch phase of the DCCS task by letting 3- and 4-year-old children perform a generalization task and a separate standard DCCS task. The rule representations could theoretically have three different levels of abstraction. The least abstract level is a representation at the level of the specific stimuli (‘the red car goes with the red rabbit and the blue rabbit goes with the blue car’). The second level is a representation at the level of the values of dimensions (‘red goes with red and blue goes with blue’). And the most abstract level of abstraction is a representation at the level of dimensions (‘same colors go together’). A possible relationship between abstraction and flexibility is studied by relating performance on the generalization task and performance on a separate standard version of the DCCS task.

The studies described in Chapter 4 and 5 investigate the influence
of exogenous factors on the performance of preschoolers in the DCCS task. Exogenous factors are bottom-up, stimulus driven, and not under the voluntary control of the child. In the study described in Chapter 4 the influence of exogenous factors that are related to the sorting rules is studied by changing the values of one or both dimensions of the target cards and test cards in the post-switch phase of the task. The influence of exogenous factors that are not related to the sorting rules is studied by changing the outline shape or the position of the target cards and test cards on the screen in the post-switch phase of the task. In the study described in Chapter 5 the influence of changing the values of one or more dimensions is investigated with a DCCS task with three stimulus dimensions, color, shape, and size that vary between target- and test cards. Can changes in the value of a certain dimension direct the attention of the child to that dimension and as a result tempt the child to sort according to that dimension (even if that dimension was never relevant)?

Chapter 6 describes a study that investigates the direct and long-term effects of feedback on 3-year-olds switch behavior with a computerized version of the DCCS task. The task was designed such that feedback was connected to the stimulus and causally related to children’s behavior. Whether children learned from the received feedback was assessed with the administration of two subsequent standard DCCS tasks (without feedback) with different stimuli, one after five minutes and one after one week.

Finally, Chapter 7 summarizes the results of the empirical studies, introduces a conflict cusp model for the transition from perseverating to switching on the DCCS task with a first test of this model, and discusses directions for future research.