Transfer of Cognitive Skills in Developmental Tasks

Sarah Rupp (srupp@uos.de)
Niels A. Taatgen (n.a.taatgen@rug.nl)
Institute of Artificial Intelligence, University of Groningen, Nijenborgh 9,
9747 AG Groningen, Netherlands

Abstract

The main question we try to answer in this paper is whether stage-like progression in cognitive development can be explained by transfer of cognitive skill among tasks. We focus on the following question: To what extent does training on one task improve the performance on another task? The tasks are Piaget's (1959) Balance Scale Task and Number Conservation Task, and a task that we will call the Une-Sentence Task, which is taken from Karmiloff-Smith's (1979) experiment on the acquisition of determiners in French. We re-implemented already existing models within the framework of the PRIMs cognitive architecture (Taatgen, 2013). Each task was subdivided in certain stages related to the complexity of the problem-solving strategies. We show that mastery of a certain stage of a problem becomes easier if a higher stage of another task is mastered first.

Keywords: Transfer, PRIMs, Cognitive Architecture, Developmental Tasks, Cognitive Modeling

Introduction

A central topic of debate in developmental psychology is whether children develop in stages, or if development is a gradual monotonic process. The idea of stages was originally introduced by Piaget in his *theory of cognitive development* (1959). The claim about stages of development from early childhood up to adulthood is still influential, even though the idea of across-the-board stage transitions is unlikely to be true. Nevertheless, it is quite plausible that progress in a certain domain can also support or enable progress in another domain. In this paper we will explore this idea in the context of concrete cognitive models that learn developmental tasks that need similar strategies.

The key idea in the modeling effort is that progress on a certain task leads to new cognitive skills that might provide the key missing piece that is necessary for progress on another task. If the new skill is applicable in multiple domains, this may give the impression of a stage-like transition, even though specific learning has to be done within each of the individual tasks.

To identify phases or stages in performance, Siegler introduced the *rule-assessment approach* (Siegler, Strauss, & Levin, 1981; Siegler, 1976). The assumption of that approach is that a certain set of rules produces a particular strategy that can later be identified in behavior. For example, in the balance scale task that we will discuss in more detail later on, both weight and distance have to be taken into account to determine the correct answer. However, early in development children only take weight into account. This produces a characteristic pattern of errors that uniquely identifies the strategy. Siegler's assumption is that progress from one stage or phase to the next was produced by general problem solving strategies. In other words, progress is independent of progress on other tasks, but only dependent on knowledge that doesn't change.

Differently but also similarly, Alison Gopnik advocates the *theory-theory* (Gopnik, 2003). This theory assumes children use reasoning methods similar to what scientists use to generate and refine their knowledge about the world. The theory-theory integrates the process of scientific research and children's ability to construct new knowledge.

An alternative mechanism to explain developmental progress is transfer from other tasks. The idea of transfer was first introduced by Thorndike and Woodworth (1901). Their conclusion, however, was that there is in fact very little evidence for transfer, and that people have to discover a strategy for a new problem all over again even if it is very similar to a problem they just solved. Only when the knowledge needed for the task is identical to knowledge from another task will there be transfer: the *identical elements* theory, which is still assumed to be the dominant explanation for transfer (or, lack thereof).

The PRIMs Theory

The PRIMs theory (Taatgen, 2013) offers an alternative to both the identical elements theory, and approaches that assume that new skills are produced by general immutable strategies. The assumption of PRIMs is that general cognitive skills are learned as a byproduct of learning specific tasks. PRIMs builds on the successful ACT-R (Adaptive Control of Thought - Rational) architecture (Anderson, 2007), and inherits most of its principles. However, the key difference is the basic building block of skill. In PRIMs these are Primitive Information Processing Elements (hence PRIMs). PRIMs are clustered together into cognitive operators that carry out the skill.

This method allows you to construct nearly any task from a fixed set of building blocks. PRIMs compare information between different cognitive modules (vision, memory, etc.) or move that information from one module to another. If a particular combination of PRIMs is used more often, they will be compiled into a single unit that can be used in any situation that that combination can be used. This means that if a particular combination of PRIMs is useful and is trained in a particular task, but that combination is also useful for another task, that other task can be learned faster because it already has some of the necessary building blocks. This mechanism can explain effects of transfer, even though the process does not actually transfer knowledge from one task to another. Instead general knowledge (the useful combinations of PRIMs) is generated by learning one task, and reused by another task. For example, a task in which a letter is displayed on the screen, and a judgment has to be made whether the letter is a vowel requires transforming a perceptual input into a memory query with a task-specific component (vowelness). This step takes four PRIMs, which can be reused for a task in which a number is displayed and the subsequent number has to be named.

Goal of this paper

This paper builds on Gittelson & Taatgen (2014), who modeled stages in the balance scale task and a decision making task. The limitation of that study was that the decision making task was not a typical developmental task. In this paper we will examine three developmental tasks in which multiple stages have been identified. The three tasks have a common strategic element: handling multiple aspects or dimensions of a particular problem. Typical stage 1 behavior is to only take one dimension into account. In a later stage, two dimensions are recognized, but are only handled one at a time. In the final stages information from multiple dimensions is integrated to reach a decision.

The approach is to build separate PRIMs models of all three tasks for each of their stages, so in total 11 models. The main question we want to answer is how easy it is to learn a new stage of a task given different types of prior knowledge. We will do this by determining the amount of transfer between combinations of models. First, we will examine how a lower stage can support a higher (typically more complex) stage. For example, to what extent is mastery of stage 1 helpful in learning stage 2. We will examine a more interesting question next: to what extent is knowledge of stage 2 of another task helpful in learning stage 2. So, is mastery of stage 2 in task A helpful in getting to stage 2 of task B. Finally, we will look at the "stage-transition" question: if you have mastered stage 1 of a task A, and stage 2 of task B, is this combination a better support to learn stage 2 of task A than just mastery of stage 1? If this is the case, progress in stage 2 in any task can trigger progress in many other tasks that are still in stage 1, producing a more general stage-like transition.

Developmental Task and Model

In the following section we will explain the three tasks we used: the Balance Scale Task (further on also Balance Task called), the Number Task and the Une-Sentence Task. We will also explain the structure of the models for each of the tasks. The first two tasks are mainly based on Siegler's (Siegler et al., 1981; Siegler, 1976) work and is re-implemented for our purposes in PRIMs. The third task was already implemented by Zondervan and Taatgen (2003), which is the basis for the PRIMs implementation.

Balance Scale Task

The Balance Task was originally described by Inhelder and Piaget (1959) and used by several other researchers. As shown in Figure 1, the children had to predict to which side the scale would tilt, or if it would stay balanced, taking into account the distance to the fulcrum and the weight of the stacks.

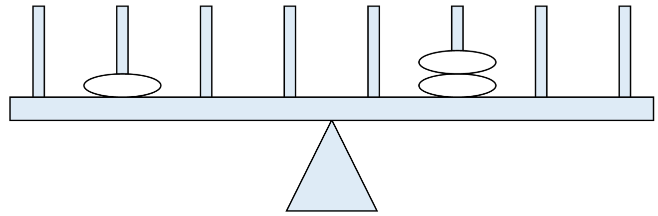

Figure 1: Balance Scale Task

Siegler labeled weight as the dominant dimension and distance as the subordinate (Siegler et al., 1981). He implemented each of the stages by a set of decision rules. For our model we re-implemented these rules directly into PRIMs. In correlation to the real world experience and according to the Piagetian stages, children need to learn to integrate multiple dimensions of the problem. In the beginning they typically only consider a single dimension. In the second stage they will consider more dimensions – but only one at a time. In the third stage they take multiple dimensions into account, and learn to weigh the dimensions properly. But there are still cases where there is a need of the proper combination of both dimensions and will always lead to the correct solution. This is what we call stage four. There is a debate going whether children even ever reach that stage without being taught explicitly. As it is proposed by Inhelder and Piaget (1959), children will reach that stage, but it has been even rejected by the findings from Siegler (1976). As we will see this is only an issue in the Balance Task, so it will not play a major role in our discussion.

In the model the first stage of the Balance Scale Task will decide just on the basis of the dominant dimension. In the second stage, the model will work just as the first stage, but in cases of equal values on the dominant dimension the model will take the subordinate dimension into account and decide on that one. In the third stage the model will take the subordinate dimension into account either way. So even in the case of different dominant dimension, it will now also take the subordination into account. If the subordinate dimension is equal, the decision is based on the dominant dimension. If both the dominant and the subordinate dimension differ, the model will guess (Siegler called this case 'Muddle through'). The fourth stage will solve the 'Muddle through' problem, with taking the dominant and the subordinate dimension comparable into account, which arises with an algebraically combination of the distance and the weight on each side (the torque law).

Number Conservation Task

Another developmental task we used is the Number Conversation Task. Like the Balance Task, it has its roots in the Piagetian account, and was also discussed by Siegler. In the number conservation task children have to decide which of two rows has more coins. The task itself concerns two phases. The first phase consists of the initial position of the coins, which are two rows of equal length and quantity. In phase two the experimenter performs a transformation to the lower row and the participant has to determine which row contains more coins. A transformation can consist of lengthening or shortening the row, and/or adding or subtraction coins from the row. This makes the task very similar to the Balance Task: both the length and numerosity are dimensions to be considered, and length is initially dominant. The difference is, of course, that the solution only involves numerosity, the subordinate dimension.

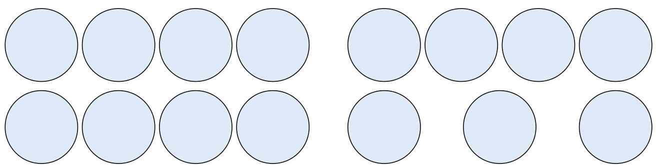

Figure 2: Number Conservation Task (Phase one is the left cluster and phase two the right cluster.)

Because of the similarity, the models for the task are very similar to the Balance Task models. Children start to learn to make the decision based on a single dimension, length. After that they will take a second dimension into account, but again only in case where the first dimension leads to a non-sufficient solution, which is the case if the values for the dominant dimension are equal, therefore considering the second dimension will help to decide. Next step is to consider both stages at any time. But in cases of two ambiguous outcomes for each dimension, children will guess. The final stage 4 rule is of course different, and similar again to stage 1, except with a different dimension.

Une - Sentence Task

The third task we modeled is based on Karmiloff-Smith's experiment on the acquisition of the function of determiners in French (Karmiloff-Smith, 1979). The French feminine word for "a", "une", has different meanings. It can be used to introduce a new item ("a balloon"), but also to indicate a specific number ("one balloon"). To assess the reasoning process that children to determine the meaning of "une", Karmiloff-Smith showed them pictures of two playgrounds, one with a girl and the other with a boy (Figure 3).

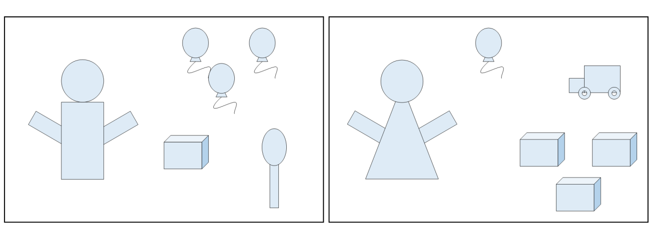

Figure 3: Playgrounds

Those playgrounds were constructed precisely; e.g. if the boy had three balloons the girl would have one. The experimenter always 'talked' to either the boy or the girl and the child had to decide to which one the experimenter was talking. Questions were of the form: 'Please lend me the X.' or 'Please lend me a X'. The word 'la' is the French word for 'the'. The French word 'une' is translated in English 'one' (number) and 'a' (indefinite article). Therefore by hearing 'Please lend me a box' (consider a playground situation as in Figure 3), the correct response of the child would be: 'You are talking to the girl'. However, if the child would answer 'boy', it would indicate they interpret 'une' as 'one'.

The task was constructed to test whether children are able to distinguish between 'une' and 'la' sentences. Results revealed that there did almost perfectly fine on 'la' sentences, but there appeared a U-shaped performance curve for 'une' sentences. At younger ages (3-4 year olds) they do well on 'une' sentences. Then there a low point at the age of 5. And then at rising ages (the testing was up to the age of 10) the performance on 'une' sentences was nearly perfect again.

Taatgen and Zondervan presented a model of the performance curve of the indefinite French article 'une' (2003) based on the representational redescription (RR) theory (Karmiloff-Smith, 1992). Generally the stages describe, as well as the stages of the other two tasks, the cognitive development. The main idea of the performance drop and the rise of performance is, that children first consider only the one dimension they know, which would be 'une' as an indefinite article – which indeed leads to the correct answer if they have to make a choice between two playgrounds: they pick the playground with multiple items. The subsequent drop is explained by a switch to the other dimension. The theory is that they now consider 'une' as a number, which leads to an incorrect response. In the final stage the children learn how to differentiate between the two interpretations by taking into account another dimension of the problem, in this case whether or not the object in question is already the focus of the discussion. If boxes are already to topic of conversation, the 'one' interpretation is correct, but if they are not, the 'a' interpretation is correct. This means that in stage 3 the decision process consists of two steps: first settle the topic question, and based on that make the right choice between the two playgrounds.

Similar to the other task and their stages, for the first stage of the Une-Sentence Task, the decision is based on one dimension (you might call it the dominant dimension). The first stage decides correctly with the information of the dominant dimension. The second stage also only take dominant dimension into account, but decides strictly incorrectly (in the context of this test, normally interpreting 'a' as 'one' does not lead to many problems). In the third stage in the task two dimensions have to be taken into account in a way that is similar, but not identical to the stages 3 in the other two tasks.

Method

To assess the amount of transfer between (stages of) tasks, we used the methodology developed by Katona (1940). This involves a comparison of two training situations. The first step is to train the model just on the task in question (let us call it task T1), and see how much progress is made after a certain amount of training (200 trials in our case). In the second step (after cleaning the environment and staring of with no prior knowledge) training of a different task will be performed. There are two cases: In the first case the model is trained on only one different task T2, which I called *one task training*. In the other case it will be trained on two different tasks (in the figure noted as T2 and T3), this will be mentioned as (two task) *shared training*. Afterwards the model is always tested on T1. We can express the benefit of T2 (or the combination T2 and T3) on T1 by comparing performance after training (Figure 4).

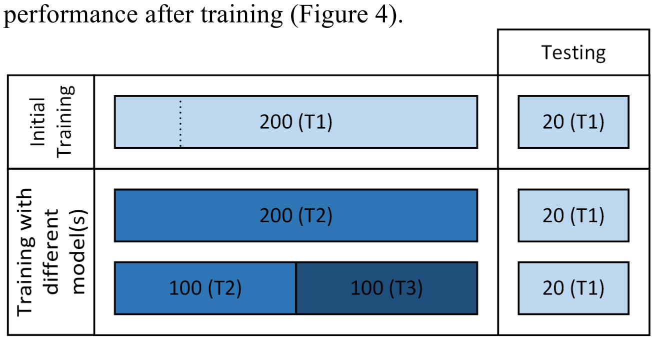

Figure 4: Illustrates the different training procedures necessary to assess transfer from task T2 to task T1, or from the combination of T2 and T3 to T1.

The transfer calculation is as follows:

$$\text{transfer} = \frac{T2_{\text{training}}}{T1_{\text{training}}} * 100\%$$

$T2_{\text{training}}$ Difference of average time of the first model (without any training) and average time after the training session on a second model (in case of shared training: the training session includes two different models)
$(T1_{\text{without\_training}} - T1_{\text{with\_training\_T2}})$

$T1_{\text{training}}$ Difference of average time of the first model (without any training) and average time after the training session on itself.
$(T1_{\text{without\_training}} - T1_{\text{with\_training\_T1}})$

Results

The results are separated into three parts, as indicated in the introduction.

Within-task transfer

We will first examine the transfer of a lower stage to a higher stage within a task. Figure 5 shows percentages for each of the three tasks. The arrow indicates the testing direction. For example: Balance Task, Stage 1 to Stage 2 has the value 46 – which indicates the transfer from stage 1 to stage 2. What we can see is that lower stages typically support the next stage fairly well. The most difficult transition seems to be from stage 1 to 2 (or 2 to 3 in the Sentence Task). Noticeable for the Number Task is that the transfer of the third stage to the fourth is not as high as the other values. The reason is that stage 4 is in fact a lot simpler than stage 3, whereas is most of the other cases a higher stage is an extension of the lower stage. Another point of notice for the Sentence Task is the transfer value for the first to the second stage, which is around 100% (noise in the simulation explains the deviance from 100%). This indicates that the structure of the solution is identical between the two stages. This problem arises again in the next part. The 100% does not mean that mastery of stage 1 automatically leads to mastery of stage 2 in the sentence task. The model still has to discover that it has to use a different dimension in stage 2 than in stage 1, and aspect we have not modeled here (but see Zondervan & Taatgen, 2003).

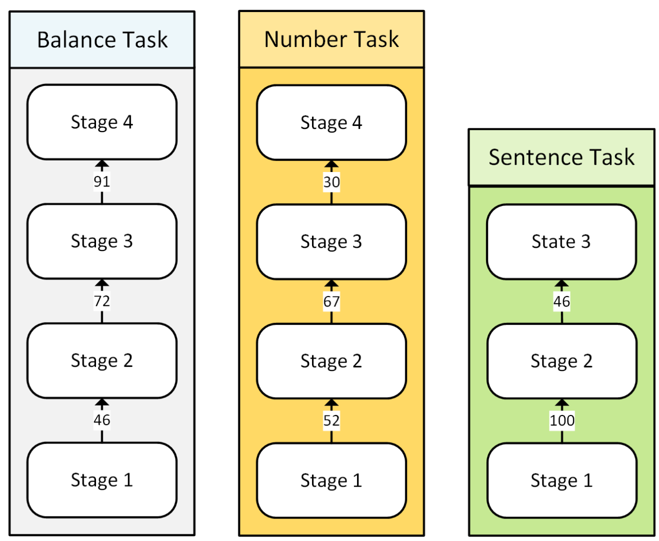

Figure 5: Within-task transfer

Transfer between almost identical tasks

The Balance Task and the Number Conservation Task are almost identical in their decision structure for each of the different stages. We therefore would expect that if a particular stage is mastered in one of the tasks, this would transfer perfectly or almost perfectly to the other task in the same stage.

Figure 6 shows that this indeed the case: transfer between the first three stages is in all cases very high. This means that after a stage is mastered in one of the stage, the cognitive skills are available to also learn that stage in a different task.

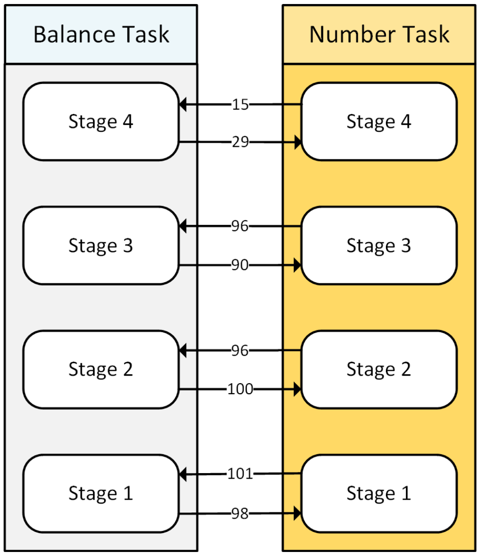

Figure 6: Transfer between the Balance and Number Task

The only exception is stage 4, but there the two tasks differ: whereas the Balance Task requires a complex multistep decision, the solution for the Number Task is quite simple, and only requires the skills learned for stage 1.

Shared Training

The most interesting case of transfer is when the model has mastered a certain stage in task A, but has already progressed to a higher stage at task B. Will skill in task B help progress on task A? To examine this we tested progress on the Balance Task given different levels of training on the Sentence task, two tasks that have similarities but are not identical in strategy.

Figure 7 shows the results. They reveal that there can be more transfer with shared training. As shown in the middle panel of the figure, training on both sentence stage 1 and balance stage 1 leads to the same transfer to balance stage 2 as just training on balance stage 1 (even though the total number of training trials is the same). Even better is the performance after shared training with the stage 3 of the sentence task: this leads to much better (62% vs. 46%) transfer than training on just stage 1 of the Balance Task. The shared training is extra effective because stage 3 of the sentence task includes skills that can handle multiple dimensions in a problem, something that is required for stage 2 of the Balance Task but that is not part of stage 1 of the Balance Task.

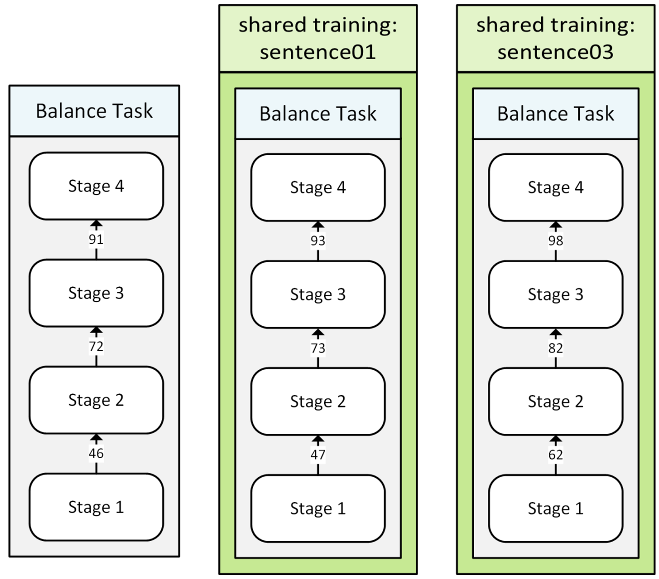

Figure 7: Shared training session. Left: repeats the within-task task transfer chart; Middle: indicates shared training with the first model of the sentence task and each model of the Balance task respectively; Right: similar to the middle column, but this time a shared training with the third model of the sentence task and each balance model.

The additional support for the stage 1 to 2 transition is particularly encouraging, because that is the hardest step to make (Figure 5). Training on stage 3 of the Sentence task is also helpful for the later transitions in de Balance task, although the effect is smaller.

Discussion

The main question of this paper is whether stage-like progress in development can be explained by the discovery of general cognitive skills. In the example we discussed here, the general skill is how different dimensions of a decision tasks are integrated. All three tasks we have looked at start with a strategy that only takes a single dimension into account, and then progress to more advanced strategies that integrate multiple dimensions. As we have seen in the simulations, different tasks can support each other. In the case of the Balance Task and the Number Conservation Task the underlying structure was almost identical. This means that progressing on one of the two tasks should make it really easy to progress to the next stage at the other task. Even when two tasks only overlap in structure, like is the case in the Balance Task and the Sentence Task, can progress on the other task help the other task. For example: We are testing the model of the second stage of the Balance Task, and train on the second stage of the Number Task, the transfer achievement is around 96 %. So we could conclude that mastery of a certain stage of task A, basically makes you are able to master another task at the same stage. If we call the mastery of a stage of a task, a skill that we achieved, we could talk about an emergent property. This holds at least for tasks that share many similar structural elements. Therefore new strategies allow you to be better at all tasks (so all tasks within that 'strategy-level'), which would be then something similar to a so-called 'stage'. These findings do not support the idea of a sudden jump from stage to stage. It rather supports the idea of a relatively fast transition from one stage to another because of the replicated use of elements, structures or strategies of a task. Also interestingly is the fact that a shared training setting does support the performance of the tested task sometimes even better. For example: We are testing the model of the Balance Scale Task of the second stage. And we train on the model of the Balance Task stage 1 and the model of the Sentence Task stage 3. We do see, that this combination gives a better result (62 % transfer), than training only on the lower stage of Balance Task stage 2, which would be Balance Task stage 1 (which is 46 %). We can conclude from this observation that it could be the case, that training on different tasks result in better performances, than only training on one specific task.

A limitation of this study is that the actual transitions between stages are not modeled, although ACT-R models of such progress are available for both the Balance task (van Rijn, van Someren & van der Maas, 2003) and the Sentence task (Zondervan & Taatgen, 2003). What this study shows, however, is a limitation of those models, and many other models that simulate developmental tasks, in that they describe progress on a task in isolation. What we shown here is the possibility that skills in general are inter-connected, and that progress on them should be studied in a broader context than just a single task.

In the end this is a theoretical account. The simulations were performed in an isolated environment and the tasks were modeled arbitrarily, of course with regard to earlier work that has been done on this field. Therefore the models were still plausible constructed based on already existing models of that kind of task. Furthermore the data we showed are theoretical predictions, but it is indeed interesting if this would be tested and be compared to empirical data. And if our prediction turn out to be true, or partially true, it would be interesting to look at possible implications for education, in a sense of teaching procedures or methods that are commonly used. At the very least this study suggests that diversity in training has benefits over just singular training.

Acknowledgments

This research was funded by ERC-StG grant 283597 awarded to Niels Taatgen.

References

Anderson, J.R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press..

Gopnik, A. (2003). The theory theory as an alternative to the innateness hypothesis. In L. Antony & N. Hornstein (Eds)., Chomsky and His Critics (pp. 238-254). New York: Basil Blackwell

Gittelson, L. & Taatgen, N.A. (2014). Transferrring primitive elements of skill within and between tasks. *Proceedings of the 36th annual meeting of the cognitive science society* (pp. 529-534). Austin, TX: Cognitive Science Society.

Piaget, J., & Inhelder, B. (1959). *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures* (Vol. 84). Routledge.

Karmiloff-Smith, A. (1979). A functional approach to child language. A study of determiners and reference. Cambridge, MA: Cambridge University Press.

Karmiloff-Smith, A. (1992). Beyond Modularity: A Developmental Perspective on Cognitive Science. *Behavioral and Brain Sciences, 17*, 693–745.

Katona, G. (1940). *Organizing and memorizing*. New York: Columbia University Press.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology, 8*(4), 481–520.

Siegler, R. S., Strauss, S., & Levin, I. (1981). Developmental Sequences within and between Concepts. *Monographs of the Society for Research in Child Development, 46*(2), 1–84.

Taatgen, N. A. (2013). The nature and transfer of cognitive skills. *Psychological Review, 120*(3), 439–471.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions: III. Functions involving attention, observation and discrimination. *Psychological Review, 8*(6), 553–564.

Van Rijn, H., Van Someren, M., & Van der Maas, H. (2003). Modeling developmental transitions on the balance scale task. *Cognitive Science, 27*(2), 227-257.

Zondervan, K., & Taatgen, N. A. (2003). The determiners model: a cognitive model of macro development and U-shaped learning in a micro domain. *Proceedings of the Fifth International Conference on Cognitive Modeling*, (pp. 225–230).