Controlling Your Mental Models

Treur, Jan

published in
Mental Models and their Dynamics, Adaptation and Control
2022

DOI (link to publisher)
10.1007/978-3-030-85821-6_4

document version
Publisher's PDF, also known as Version of record

document license
Article 25fa Dutch Copyright Act

Chapter 4
Controlling Your Mental Models: Using Metacognition to Control Use and Adaptation for Multiple Mental Models

Jan Treur

Abstract Learning processes can be described by adaptive mental (or neural) network models. If metacognition is used to regulate learning, the adaptation of the mental network becomes itself adaptive as well: second-order adaptation. In this chapter, a second-order adaptive mental network model is introduced for metacognitive regulation of learning processes. The focus is on the role of multiple internal mental models, in particular, the case of visualisation to support learning of numerical or symbolic skills. The second-order adaptive network model is illustrated by a case scenario for the role of visualisation to support learning multiplication at the primary school.

Keywords Metacognition · Control · Mental model · Multiple representation

4.1 Introduction

Metacognition (Darling-Hammond et al. 2008; Shannon 2008; Mahdavi 2014; Flavell 1979; Koriat 2007; Pintrich 2000) is a form of cognition about cognition. In (Koriat 2007) it is described as what people know about their own cognitive processes and how they put that knowledge to use in regulating their cognitive processing and behavior. A sometimes used closely related term is self-regulation and when the cognitive processes addressed by metacognition concern learning, the term self-regulated learning is used. For example, in (Pintrich 2000), self-regulated learning is described as an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behavior, guided and constrained by their goals and context.

J. Treur (✉)
Social AI Group, Department of Computer Science, Vrije Universiteit, Amsterdam, Netherlands
e-mail: j.treur@vu.nl

J. Treur et al. (eds.), *Mental Models and Their Dynamics, Adaptation, and Control*, Studies in Systems, Decision and Control 394,
https://doi.org/10.1007/978-3-030-85821-6_4

In learning, often different mental models play a role; e.g., (Gentner and Stevens 1983; Greca and Moreira 2000; Skemp 1971; Seel 2006). A specific case where the role of metacognition in learning processes is considered within educational science, is the use of multiple mental models such as in visualisation to support learning of more abstract (numerical or symbolic) skills; e.g., (Bruner 1966; Du Plooy 2016). An important metacognitive control decision in this context is whether or not and when to switch from one mental model to another one. In the educational science literature, much more can be found on this case, particularly for learning arithmetic or algebraic skills in primary or secondary schools supported by visualisation; see also (Bruner 1977; Bidwell 1972; Day and Hurell 2015; Freudenthal 1973; Freudenthal 1986; Koedinger and Terao 2002; Larbi and Mavis 2016; Lovitt et al. 1984; Renkema 2019; Roberts 1989).

From a network-oriented modeling perspective, learning is usually described by adaptive mental (or neural) network models, where some of the network characteristics such as connection weights or excitability thresholds change over time. If, in addition, metacognition is used to regulate or control the learning, this implies that the adaptation (by learning) of the mental network is itself adaptive as well, which is called second-order adaptation. Thus, a network model for such processes has to address such complex structures and behaviour. In the current chapter, using the modeling approach for higher-order adaptive networks from (Treur 2018,2020a, b), a second-order adaptive mental network model is introduced for metacognitive regulation of such learning processes. Here, the focus is on the role of multiple mental models in case of visualisation to support learning of more abstract (numerical or symbolic) skills. The adaptive network model is illustrated for a case study on the role of visualisation to support learning multiplication at the primary school as described, for example, in (Bruner 1966; Day and Hurrell 2015; Du Plooy 2016; Freudenthal 1973,1986; Rivera, 2011).

In this chapter, first in Sect. 4.2 more background knowledge is discussed on metacognition and the role of visualisation in learning processes. In Sect. 4.3 the network-oriented modeling approach used is briefly explained. Next, in Sect. 4.4 the introduced second-order adaptive network model is described in some detail. In Sect. 4.5, it is shown how this model was used to perform simulations for the illustrative example scenario. Finally, Sect. 4.6 is a discussion.

4.2 Metacognition and Multiple Mental Models

Literature on metacognition, sometimes also called self-regulation, can be found, for example in (Darling-Hammond et al. 2008; Shannon 2008; Mahdavi 2014; Flavell 1979; Koriat 2007; Pintrich 2000). The focus is here on the role of metacognition in learning. For example, in (Pintrich 2000, pp. 452–453) the following assumptions for self-regulated learning are described:

- It is a process whereby learners set goals for their learning and monitor and control their cognition, motivation, and behavior guided by these goals.
- Learners actively construct their own meanings, goals, and strategies.
- Learners can monitor, control, and regulate certain aspects of their own cognition, motivation, and behavior, and some elements of their environment.
- Some type of criterion or standard is used to assess whether the process should continue as is or if some type of change is necessary.
- Self-regulatory activities are mediators between personal and contextual characteristics and actual performance.

In line with these assumptions, in (Pintrich 2000, pp. 453–461, Table 1, p. 454), the following phases for self-regulation are described:

- Cognitive planning and activation
- Cognitive monitoring
- Cognitive control and regulation
- Cognitive reaction and reflection

In (Koriat 2007, p. 290), metacognition is described by what people know about cognition and in particular their own cognitive processes, and how they use that in regulating their cognitive processes and behavior. Assumptions mentioned there are:

- Self-controlled cognitive processes have measurable effects on behavior (Koriat 2007), pp. 292–293
- Feelings, such as the feeling of knowing are part of monitoring, and exert a causal role on the control of cognitive processing (Koriat 2007), p. 293, p. 314–315
- There is a causal relation from monitoring to control (Koriat 2007), p. 315

So, in both descriptions of Pintrich (2000) and Koriat (2007) on metacognition (as well as in most other literature on metacognition), monitoring and control of the own cognitive processes are central concepts (where Koriat also emphasizes the feeling or experiencing that comes together with monitoring). These processes work through a causal cycle where the own cognitive processes affect the metacognitive monitoring, this monitoring in turn affects the metacognitive control, and this control affects the own cognitive processes. This causal cycle will indeed be incorporated in the adaptive network model introduced in Sect. 4. Note that metacognitive monitoring is usually based on forming and maintaining a self-model describing a (subjective) estimation of some relevant aspects of the own cognitive processes.

In the area of learning using multiple mental models (Gentner and Stevens 1983; Greca and Moreira 2000; Skemp 1971; Seel 2006), metacognition plays an important role for the decisions about when to switch from one mental model to another one. In particular, this takes place when learning numerical or symbolic skills in arithmetic or mathematics is supported by visualisations; e.g., see (Bruner 1966,1977; Bidwell 1972; Day and Hurell 2015; Du Plooy 2016; Freudenthal

Table 4.1 The states in the adaptive network model

X_1	$\mathbf{N}_1$	Base state for number *a*
X_2	$\mathbf{N}_2$	Base state for number *b*
X_3	$\mathbf{N}_3$	Base state for number *c*
X_4	$\mathbf{S}_{23}$	Base state for number *b* + *c*
X_5	$\mathbf{P}_{12}$	Base state for number *a***b*
X_6	$\mathbf{P}_{13}$	Base state for number *a***c*
X_7	$\mathbf{PS}_{123}$	Base state for number *a**(*b* + *c*)
X_8	$\mathbf{SP}_{1213}$	Base state for number *a***b* + *a***c*
X_9	$\mathbf{RD}_{vert}$	Vertical dimension of the rectangles
X_{10}	$\mathbf{RD}_{hor1}$	Horizontal dimension of rectangle 1
X_{11}	$\mathbf{RD}_{hor2}$	Horizontal dimension of rectangle 2
X_{12}	$\mathbf{RD}_{hor3}$	Horizontal dimension of rectangle 3
X_{13}	$\mathbf{RA}_1$	Area of rectangle 1
X_{14}	$\mathbf{RA}_2$	Area of rectangle 2
X_{15}	$\mathbf{RA}_3$	Area of rectangle 3
X_{16}	$\mathbf{RA}_{12}$	Area of rectangles 1 and 2 together
X_{17}	$\mathbf{W}_{P112}$	Representation state for the weight of the connection from $\mathbf{N}_1$ to $\mathbf{P}_{12}$
X_{18}	$\mathbf{W}_{P212}$	Representation state for the weight of the connection from $\mathbf{N}_2$ to $\mathbf{P}_{12}$
X_{19}	$\mathbf{W}_{P113}$	Representation state for the weight of the connection from $\mathbf{N}_1$ to $\mathbf{P}_{13}$
X_{20}	$\mathbf{W}_{P313}$	Representation state for the weight of the connection from $\mathbf{N}_3$ to $\mathbf{P}_{13}$
X_{21}	$\mathbf{W}_{SP121213}$	Representation state for the weight of the connection from $\mathbf{P}_{12}$ to $\mathbf{SP}_{1213}$
X_{22}	$\mathbf{W}_{SP131213}$	Representation state for the weight of the connection from $\mathbf{P}_{13}$ to $\mathbf{SP}_{1213}$
X_{23}	$\mathbf{RW}_P$	Mental representation state concerning the weights of the connections to $\mathbf{P}_{12}$ and $\mathbf{P}_{13}$
X_{24}	$\mathbf{RW}_{SP}$	Mental representation state concerning the weights of the connections to $\mathbf{SP}_{1213}$
X_{25}	$\mathbf{WRD}_{vert}$	Representation state used for execution of control decision $\mathbf{CWRD}_{vert}$, representing the weight of the connection from $\mathbf{N}_1$ to $\mathbf{RD}_{vert}$
X_{26}	$\mathbf{WRD}_{hor1}$	Representation state used for execution of control decision $\mathbf{CWRD}_{hor1}$, representing the weight of the connection from $\mathbf{N}_2$ to $\mathbf{RD}_{hor1}$
X_{27}	$\mathbf{WRD}_{hor2}$	Representation state used for execution of control decision $\mathbf{CWRD}_{hor2}$, representing the weight of the connection from $\mathbf{N}_3$ to $\mathbf{RD}_{hor2}$
X_{28}	$\mathbf{RS}_{num}$	Representation of the self-model for the own numerical skills
X_{29}	$\mathbf{RS}_{geo}$	Representation of the self-model for the own geometric skills
X_{30}	$\mathbf{CWRD}_{vert}$	Control state for the switch to the geometric mental model: representation of the weight of the connection from $\mathbf{RW}_{PSP}$ to $\mathbf{WRD}_{vert}$
X_{31}	$\mathbf{CWRD}_{hor1}$	Control state for the switch to the geometric mental model: representation of the weight of the connection from $\mathbf{RW}_{PSP}$ to $\mathbf{WRD}_{hor1}$
X_{32}	$\mathbf{CWRD}_{hor2}$	Control state for the switch to the geometric mental model: representation of the weight of the connection from $\mathbf{RW}_{PSP}$ to $\mathbf{WRD}_{hor2}$

1973,1986; Koedinger and Terao 2002; Larbi and Mavis 2016; Lovitt et al. 1984; Renkema 2019; Roberts 1989). Here, when at some point during working with a numerical or symbolic mental model, a learner monitors that the cognitive processes get stuck, the control decision can be made by the learner to switch to working with a mental model based on visualisation, after which the outcomes can be fed back to the numerical or symbolic mental model. Within the literature in educational science as mentioned above, it is extensively described how such a detour via a visualisation can support the learning of numerical or symbolic skills. This type of use of metacognition for using multiple mental models is the main focus in the current chapter.

4.3 Higher-Order Adaptive Network Models

In this section, the network-oriented modeling approach used is briefly introduced. Following (Treur 2016,2020b), a temporal-causal network model is characterised by (here X and Y denote nodes of the network, also called states):

- *Connectivity characteristics*
  Connections from a state X to a state Y and their weights $\boldsymbol{\omega}_{X,Y}$
- *Aggregation characteristics*
  For any state Y, some combination function $\mathbf{c}_{\boldsymbol{Y}}$**(..)** defines the aggregation that is applied to the impacts $\boldsymbol{\omega}_{X,Y}X(t)$ on Y from its incoming connections from states X
- *Timing characteristics*
  Each state Y has a speed factor $\boldsymbol{\eta}_Y$ defining how fast it changes for given causal impact.

The following difference (or differential) equations that are used for simulation purposes and also for analysis of temporal-causal networks incorporate these network characteristics $\boldsymbol{\omega}_{X,Y}$, $\mathbf{c}_Y(..)$, $\boldsymbol{\eta}_Y$ in a standard numerical format:

$$Y(t+\Delta t) = Y(t) + \boldsymbol{\eta}_{\boldsymbol{Y}}[\mathbf{c}_{\boldsymbol{Y}}(\boldsymbol{\omega}_{X_1,Y}\boldsymbol{X_1}(\boldsymbol{t}), \ldots, \boldsymbol{\omega}_{X_k,Y}\boldsymbol{X_k}(\boldsymbol{t})) - \boldsymbol{Y}(\boldsymbol{t})]\boldsymbol{\Delta t} \tag{4.1}$$

for any state Y and where X_1 to X_k are the states from which Y gets its incoming connections. Within the software environment described in (Treur 2020b, Ch. 9), a large number of around 40 useful basic combination functions are included in a combination function library.

The above concepts enable to design network models and their dynamics in a declarative manner, based on mathematically defined functions and relations. Realistic network models are usually adaptive: often not only their states but also some of their network characteristics change over time. By using a *self-modeling network* (also called a *reified* network), a similar network-oriented conceptualisation can also be applied to *adaptive* networks to obtain a declarative description using mathematically defined functions and relations for them as well; see (Treur 2018, 2020a, b). This works through the addition of new states to the network (called *self-model states*) which represent (adaptive) network characteristics. In the graphical 3D-format as shown in Sect. 4, such additional states are depicted at a next level (called *self-model level* or *reification level*), where the original network is at the *base level*. As an example, the weight $\boldsymbol{\omega}_{X,Y}$ of a connection from state X to state Y can be represented (at a next self-model level) by a self-model state named $\mathbf{W}_{X,Y}$ (objective representation actually used) or $\mathbf{RW}_{X,Y}$ (subjective representation for a person-related self-model). Similarly, all other network characteristics from $\boldsymbol{\omega}_{X,Y}$, $\mathbf{c}_Y(..)$, $\boldsymbol{\eta}_Y$ can be made adaptive by including self-model states for them. For example, an adaptive speed factor $\boldsymbol{\eta}_Y$ can be represented by a self-model state named $\mathbf{H}_Y$ and an adaptive excitability threshold parameter $\boldsymbol{\tau}_Y$ can be represented by a self-model state named $\mathbf{T}_Y$.

As the outcome of such a process of network reification is also a temporal-causal network model itself, as has been proven in (Treur 2020b, Ch 10), this self-modeling network construction can easily be applied iteratively to obtain multiple orders of self-models at multiple self-model levels. In the current chapter, a multi-level self-modeling network will be applied to obtain a second-order adaptive mental network model addressing metacognitive control of learning in a multiple mental models context.

4.4 A Mental Network Model for Metacognitive Control of Learning from Multiple Internal Mental Models

In this section, the adaptive mental network model for metacognitive control on learning using multiple mental models is introduced. This adaptive mental network model has processes at three levels:

- The base level network for the (multiple) internal mental models used
- The first-order self-model level for the learning of the internal mental models by adaptations of them
- The second-order self-model level for control by adaptation of the first-order network for the learning

These three levels of processes have been modeled by a second-order adaptive self-modeling network (Treur 2018, 2020a,b) briefly described in Sect. 3; the connectivity of this network model is depicted in Fig. 4.2. The states used are explained in Table 4.1. For the example mental models at the base level, on the left hand side in Fig. 4.2 an internal numerical mental model for an arithmetic task is included and on the right hand side a visual, geometrical mental model for it. The example task is to show (in the numerical representation) for certain given natural numbers a, b and c that

$$a^*(b+c) = a^*b + a^*c \tag{4.2}$$

Note that this is often applied in calculations, for example, when calculating 9*48 by splitting it as $9 * 40 + 9 * 8 = 360 + 72 = 432$, or by calculating 27*7 + 27*3 as 27*(7 + 3) = 270.

The detour via visualisation considers two rectangles with vertical dimension a and horizontal dimensions b and c and their areas that are together equal to the area of a rectangle with vertical dimension a and horizontal dimension b + c, as shown in Fig. 4.1.

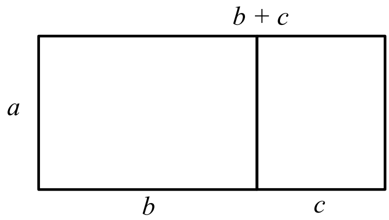

Fig. 4.1 Visualisation for the task expressed numerically by (2)

4.4.1 Network Characteristics: Connectivity and Timing

At the base level, for the numerical mental model, the base states $\mathbf{N}_1$, $\mathbf{N}_2$ and $\mathbf{N}_3$ represent the given numbers *a*, *b*, and *c*. Base states $\mathbf{P}_{12}$ and $\mathbf{P}_{13}$ represent the products *a***b* and *a***c*, respectively, whereas state $\mathbf{S}_{12}$ represents the sum *b* + *c*. Finally, base state $\mathbf{SP}_{1213}$ represents the sum of $\mathbf{P}_{12}$ and $\mathbf{P}_{13}$ which is *a***b* + *a***c*, while base state $\mathbf{PS}_{123}$ represents the product of $\mathbf{N}_1$ and $\mathbf{S}_{23}$ which is *a** (*b* + *c*). For the geometric mental model, base states $\mathbf{RD}_{\text{vert}}$, $\mathbf{RD}_{\text{hor1}}$, $\mathbf{RD}_{\text{hor2}}$, and $\mathbf{RD}_{\text{hor3}}$ represent the vertical and horizontal dimensions of the rectangles in Fig. 4.1, respectively. Moreover, $\mathbf{RA}_1$, $\mathbf{RA}_2$ and $\mathbf{RA}_3$ represent the areas of the three rectangles with horizontal dimension *b*, *c*, and *b* + *c*, respectively, and $\mathbf{RA}_{12}$ the area of the two smaller rectangles together.

At the first-order self-model level, the learning of the adaptive connections of the numerical mental model is modeled by the **W**-states and as input for the the self-model for the metacognitive monitoring the learnt relations as estimated by the learner are represented by the two (subjective) **RW**-states X_{23} and X_{24}. Moreover, the **WRD**-states X_{25} to X_{27} model the adaptive connections from the numerical mental model to the geometric mental model used to dynamically switch from one to the other; this is part of effectuating the metacognitive control.

At the second-order self-model level, the self-model for the status of the learning (for the own estimated learnt numerical and geometric skills) for the metacognitive monitoring is represented by the two **RS**-states X_{28} and X_{29} and the metacognitive control decisions (to switch to the geometric mental model) are modeled by the **CWRD**-states X_{30} to X_{32}, based on the impact from the self-model obtained by the metacognitive monitoring.

There are two types of connections: intra-level connections (in Fig. 4.2 depicted in black) and interlevel connections (depicted in blue for upward and in pink for downward). At the base level, within each of the two mental models, the connections define these mental models by their internal causal impacts. For example, the connections $\mathbf{N}_1 \rightarrow \mathbf{P}_{12}$ and $\mathbf{N}_2 \rightarrow \mathbf{P}_{12}$ define that within the numerical mental model the product of *a* and *b* represented by base state $\mathbf{P}_{12}$ depends on base states $\mathbf{N}_1$ and $\mathbf{N}_2$ representing these numbers.

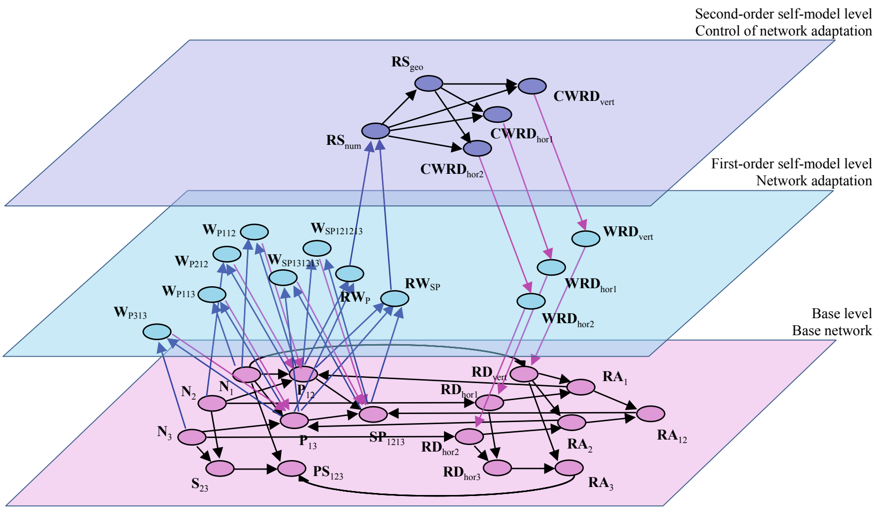

Fig. 4.2 Graphical representation of the connectivity of the second-order adaptive mental network model for metacognitive control of learning for multiple mental models

In addition, at the base level a number of connections define how the two mental models relate to each other. For example, the connection $\mathbf{N}_1 \rightarrow \mathbf{RD}_{\text{vert}}$ from the numerical mental model to the geometric mental model defines that the vertical dimension of the rectangles within the geometric mental model depends on the number *a* represented by numerical state $\mathbf{N}_1$. Moreover, a connection back from the geometric to the numerical mental model such as $\mathbf{RA}_{12} \rightarrow \mathbf{SP}_{1213}$ defines the influence of the outcomes of the geometric process on the numerical process as a form of reinforcement to amplify the learning of the numerical mental model.

The upward connections to the first-order self-model **W**-states provide impact to the **W**-states so that they can adapt over time, which is modeled according to a qualitative Hebbian learning (Hebb 1949) principle specified by (4.6) below. For example, connections $\mathbf{N}_3 \rightarrow \mathbf{W}_{\text{P313}}$ and $\mathbf{P}_{13} \rightarrow \mathbf{W}_{\text{P313}}$ provide impact to $\mathbf{W}_{\text{P313}}$ so that $\mathbf{W}_{\text{P313}}$ can adapt over time. On the other hand, the downward connection from a **W**-state makes that the value of it is actually used in the processing of the mental model. For example, the connection $\mathbf{W}_{\text{P313}} \rightarrow \mathbf{P}_{13}$ takes care for this for $\mathbf{W}_{\text{P313}}$ so that for the weight of the connection $\mathbf{N}_3 \rightarrow \mathbf{P}_{13}$ the value of $\mathbf{W}_{\text{P313}}$ is used. Furthermore, the upward connections to the first-order self-model **RW**-states make that a representation for the status of some connections of the numerical mental model is formed and maintained. This is a first step toward a self-model which is the basis of the metacognitive monitoring of the own cognitive processes.

At the second-order self-model level, based on impact from the **RW**-states at the first-order self-model level, the self-model is formed and maintained by the states $\mathbf{RS}_{num}$ and $\mathbf{RS}_{geo}$. Via their outgoing connections, the states $\mathbf{RS}_{num}$ and $\mathbf{RS}_{geo}$ of this self-model have their impact on the control decisions modeled by the **CWRD-**states. By their downward connections, the **CWRD-**states for control decisions determine the incoming connections to the corresponding **WRD**-states, so that the control decision is executed by realising that these **WRD**-states get values 1. In turn, once the **WRD**-state has a value 1, it makes that at the base level the corresponding connection from numerical mental model to geometric mental model is 1, which then leads to the geometric mental model states $\mathbf{RD}_{vert}$, $\mathbf{RD}_{hor1}$, and $\mathbf{RD}_{hor2}$ getting the appropriate values from states $\mathbf{N}_1$, $\mathbf{N}_2$, and $\mathbf{N}_3$ of the numerical mental model.

In Box 4.1 the complete role matrix specification of the connectivity and timing characteristics of the designed adaptive network model can be found. Here in each role matrix, each state has its row where it is listed which are the impacts on it from that role. Role matrix **mb** lists the other states (at the same or lower level) from which the state gets its incoming connections, whereas in role matrix **mcw** the connection weights are listed for these connections. Note that nonadaptive connection weights are indicated by a number (in a green shaded cell), but adaptive connection weights are indicated by a reference to the (self-model) state representing the adaptive value (in a peach-red shaded cell). For example, state X_5 (= $\mathbf{P}_{12}$) has incoming connections from X_1 (= $\mathbf{N}_1$), X_2 (= $\mathbf{N}_2$), and X_{13} (= $\mathbf{RA}_1$) with connection weights represented by X_{17} (= $\mathbf{W}_{P112}$) and X_{18} (= $\mathbf{W}_{P212}$) and 1, respectively. These two adaptive connection weights model the reinforced (by $\mathbf{RA}_1$) Hebbian learning. Also, the states $\mathbf{RD}_{vert}$, $\mathbf{RD}_{hor1}$, $\mathbf{RD}_{hor2}$ for the dimensions of the rectangles in the geometric mental model have adaptive connection weights. These adaptive connections are used to model the metacognitive control of the switch from numerical mental model to geometric mental model: if the control decision is made to switch, then these connection weights (represented by the **WRD**-states) quickly become 1 to transfer the numbers *a*, *b* and *c* to the geometric mental model. This rapid transition is specified in role matrix **ms** for the timing, where it is indicated that the speed factors of the **WRD**-states X_{25} to X_{27} are adaptive and immediately change from 0 to 1 as soon as the **CWRD-**states X_{30} to X_{32} for metacognitive control at the second-order self-model level change to 1.

Box 4.1. Role matrices for the connectivity and timing characteristics of the network model

mb	base connectivity	1	2	3
X_1	$\mathbf{N}_1$	X_1		
X_2	$\mathbf{N}_2$	X_2		
X_3	$\mathbf{N}_3$	X_3		
X_4	$\mathbf{S}_{23}$	X_2	X_3	
X_5	$\mathbf{P}_{12}$	X_1	X_2	X_{13}
X_6	$\mathbf{P}_{13}$	X_1	X_3	X_{14}
X_7	$\mathbf{PS}_{123}$	X_1	X_4	X_{15}
X_8	$\mathbf{SP}_{1213}$	X_5	X_6	X_{16}
X_9	$\mathbf{RD}_{\mathrm{vert}}$	X_1		
X_{10}	$\mathbf{RD}_{\mathrm{hor1}}$	X_2		
X_{11}	$\mathbf{RD}_{\mathrm{hor2}}$	X_3		
X_{12}	$\mathbf{RD}_{\mathrm{hor3}}$	X_{10}	X_{11}	
X_{13}	$\mathbf{RA}_1$	X_9	X_{10}	
X_{14}	$\mathbf{RA}_2$	X_9	X_{11}	
X_{15}	$\mathbf{RA}_3$	X_9	X_{12}	
X_{16}	$\mathbf{RA}_{12}$	X_{13}	X_{14}	
X_{17}	$\mathbf{W}_{\mathrm{P112}}$	X_1	X_5	X_{17}
X_{18}	$\mathbf{W}_{\mathrm{P212}}$	X_2	X_5	X_{18}
X_{19}	$\mathbf{W}_{\mathrm{P113}}$	X_1	X_6	X_{19}
X_{20}	$\mathbf{W}_{\mathrm{P313}}$	X_3	X_6	X_{20}
X_{21}	$\mathbf{W}_{\mathrm{SP121213}}$	X_5	X_8	X_{21}
X_{22}	$\mathbf{W}_{\mathrm{SP131213}}$	X_6	X_8	X_{22}
X_{23}	$\mathbf{RW}_{\mathrm{P}}$	X_5	X_6	
X_{24}	$\mathbf{RW}_{\mathrm{SP}}$	X_5	X_6	X_8
X_{25}	$\mathbf{WRD}_{\mathrm{vert}}$	X_{25}		
X_{26}	$\mathbf{WRD}_{\mathrm{hor1}}$	X_{26}		
X_{27}	$\mathbf{WRD}_{\mathrm{hor2}}$	X_{27}		
X_{28}	$\mathbf{RS}_{\mathrm{num}}$	X_{23}	X_{24}	
X_{29}	$\mathbf{RS}_{\mathrm{geo}}$	X_{28}		
X_{30}	$\mathbf{CWRD}_{\mathrm{vert}}$	X_{25}	X_{29}	
X_{31}	$\mathbf{CWRD}_{\mathrm{hor1}}$	X_{25}	X_{29}	
X_{32}	$\mathbf{CWRD}_{\mathrm{hor2}}$	X_{25}	X_{29}	

mcw	connection weights	1	2	3
X_1	$\mathbf{N}_1$	1		
X_2	$\mathbf{N}_2$	1		
X_3	$\mathbf{N}_3$	1		
X_4	$\mathbf{S}_{23}$	1	1	
X_5	$\mathbf{P}_{12}$	X_{17}	X_{18}	1
X_6	$\mathbf{P}_{13}$	X_{19}	X_{20}	1
X_7	$\mathbf{PS}_{123}$	1	1	1
X_8	$\mathbf{SP}_{1213}$	X_{21}	X_{22}	1
X_9	$\mathbf{RD}_{\mathrm{vert}}$	X_{25}		
X_{10}	$\mathbf{RD}_{\mathrm{hor1}}$	X_{26}		
X_{11}	$\mathbf{RD}_{\mathrm{hor2}}$	X_{27}		
X_{12}	$\mathbf{RD}_{\mathrm{hor3}}$	1	1	
X_{13}	$\mathbf{RA}_1$	1	1	
X_{14}	$\mathbf{RA}_2$	1	1	
X_{15}	$\mathbf{RA}_3$	1	1	
X_{16}	$\mathbf{RA}_{12}$	1	1	
X_{17}	$\mathbf{W}_{\mathrm{P112}}$	1	1	1
X_{18}	$\mathbf{W}_{\mathrm{P212}}$	1	1	1
X_{19}	$\mathbf{W}_{\mathrm{P113}}$	1	1	1
X_{20}	$\mathbf{W}_{\mathrm{P313}}$	1	1	1
X_{21}	$\mathbf{W}_{\mathrm{SP121213}}$	1	1	1
X_{22}	$\mathbf{W}_{\mathrm{SP131213}}$	1	1	1
X_{23}	$\mathbf{RW}_{\mathrm{P}}$	1	1	
X_{24}	$\mathbf{RW}_{\mathrm{SP}}$	1	1	1
X_{25}	$\mathbf{WRD}_{\mathrm{vert}}$	0		
X_{26}	$\mathbf{WRD}_{\mathrm{hor1}}$	0		
X_{27}	$\mathbf{WRD}_{\mathrm{hor2}}$	0		
X_{28}	$\mathbf{RS}_{\mathrm{num}}$	1	1	
X_{29}	$\mathbf{RS}_{\mathrm{geo}}$	1		
X_{30}	$\mathbf{CWRD}_{\mathrm{vert}}$	-0.1	1	
X_{31}	$\mathbf{CWRD}_{\mathrm{hor1}}$	-0.1	1	
X_{32}	$\mathbf{CWRD}_{\mathrm{hor2}}$	-0.1	1	

ms	speed factors	1
X_1	$\mathbf{N}_1$	0
X_2	$\mathbf{N}_2$	0
X_3	$\mathbf{N}_3$	0
X_4	$\mathbf{S}_1$	0.5
X_5	$\mathbf{P}_{12}$	0.5
X_6	$\mathbf{P}_{13}$	0.5
X_7	$\mathbf{PS}_{123}$	0.5
X_8	$\mathbf{SP}_{1213}$	0.5
X_9	$\mathbf{RD}_{\mathrm{vert}}$	0.5
X_{10}	$\mathbf{RD}_{\mathrm{hor1}}$	0.5
X_{11}	$\mathbf{RD}_{\mathrm{hor2}}$	0.5
X_{12}	$\mathbf{RD}_{\mathrm{hor3}}$	0.5
X_{13}	$\mathbf{RA}_1$	0.5
X_{14}	$\mathbf{RA}_2$	0.5
X_{15}	$\mathbf{RA}_3$	0.15
X_{16}	$\mathbf{RA}_{12}$	0.05
X_{17}	$\mathbf{W}_{\mathrm{P112}}$	0.02
X_{18}	$\mathbf{W}_{\mathrm{P212}}$	0.02
X_{19}	$\mathbf{W}_{\mathrm{P113}}$	0.02
X_{20}	$\mathbf{W}_{\mathrm{P313}}$	0.02
X_{21}	$\mathbf{W}_{\mathrm{SP121213}}$	0.02
X_{22}	$\mathbf{W}_{\mathrm{SP131213}}$	0.02
X_{23}	$\mathbf{RW}_{\mathrm{P}}$	0.1
X_{24}	$\mathbf{RW}_{\mathrm{SP}}$	0.1
X_{25}	$\mathbf{WRD}_{\mathrm{vert}}$	X_{30}
X_{26}	$\mathbf{WRD}_{\mathrm{hor1}}$	X_{31}
X_{27}	$\mathbf{WRD}_{\mathrm{hor2}}$	X_{32}
X_{28}	$\mathbf{RS}_{\mathrm{num}}$	0.1
X_{29}	$\mathbf{RS}_{\mathrm{geo}}$	0.5
X_{30}	$\mathbf{CWRD}_{\mathrm{vert}}$	0.5
X_{31}	$\mathbf{CWRD}_{\mathrm{hor1}}$	0.5
X_{32}	$\mathbf{CWRD}_{\mathrm{hor2}}$	0.5

4.4.2 Network Characteristics: Aggregation

The network characteristics for aggregation are defined by the selection of combination functions from the library and values for their parameters. First the six combination functions used for the model are specified by

$$\mathbf{mcf} = [1, 2, 39, 22, 23, 4]$$

$$= [\mathbf{eucl}, \mathbf{alogistic}, \mathbf{hebbqual}, \mathbf{complement-id}, \mathbf{product}, \mathbf{max-composition}]$$

Here the numbers are the numbers of the listed functions in the library. Next, it is specified which state uses which combination function. This can be seen in role matrix **mcfw** in Box 4.2.

Box 4.2. Role matrices for the aggregation characteristics: combination functions and their parameters

mcfw	**combination function weights**	1 **eucl**	2 **alogistic**	3 **hebb-qual**	4 **comp-id**	5 **product**	6 **max-comp**
X_1	$\mathbf{N}_1$	1					
X_2	$\mathbf{N}_2$	1					
X_3	$\mathbf{N}_3$	1					
X_4	$\mathbf{S}_{23}$	1					
X_5	$\mathbf{P}_{12}$						1
X_6	$\mathbf{P}_{13}$						1
X_7	$\mathbf{PS}_{123}$						1
X_8	$\mathbf{SP}_{1213}$						1
X_9	$\mathbf{RD}_{vert}$	1					
X_{10}	$\mathbf{RD}_{hor1}$	1					
X_{11}	$\mathbf{RD}_{hor2}$	1					
X_{12}	$\mathbf{RD}_{hor3}$	1					
X_{13}	$\mathbf{RA}_1$					1	
X_{14}	$\mathbf{RA}_2$					1	
X_{15}	$\mathbf{RA}_3$					1	
X_{16}	$\mathbf{RA}_{12}$	1					
X_{17}	$\mathbf{W}_{P112}$			1			
X_{18}	$\mathbf{W}_{P212}$			1			
X_{19}	$\mathbf{W}_{P113}$			1			
X_{20}	$\mathbf{W}_{P313}$			1			
X_{21}	$\mathbf{W}_{SP121213}$			1			
X_{22}	$\mathbf{W}_{SP131213}$			1			
X_{23}	$\mathbf{RW}_P$		1				
X_{24}	$\mathbf{RW}_{SP}$		1				
X_{25}	$\mathbf{WRD}_{vert}$				1		
X_{26}	$\mathbf{WRD}_{hor1}$				1		
X_{27}	$\mathbf{WRD}_{hor2}$				1		
X_{28}	$\mathbf{RS}_{num}$		1				
X_{29}	$\mathbf{RS}_{geo}$				1		
X_{30}	$\mathbf{CWRD}_{vert}$		1				
X_{31}	$\mathbf{CWRD}_{hor1}$		1				
X_{32}	$\mathbf{CWRD}_{hor2}$		1				

mcfp	**combination function parameters**	1 **eucl**		2 **alogistic**		3 **hebb-qual**		4 **comp-id**		5 **product**		6 **max-comp**	
		1 n	2 λ	1 σ	2 τ	1 μ	2	1	2	1	2	1 γ_1	2 γ_2
X_1	$\mathbf{N}_1$	1	1										
X_2	$\mathbf{N}_2$	1	1										
X_3	$\mathbf{N}_3$	1	1										
X_4	$\mathbf{S}_{23}$	1	1										
X_5	$\mathbf{P}_{12}$											23	1
X_6	$\mathbf{P}_{13}$											23	1
X_7	$\mathbf{PS}_{123}$											23	1
X_8	$\mathbf{SP}_{1213}$											1	1
X_9	$\mathbf{RD}_{vert}$	1	1										
X_{10}	$\mathbf{RD}_{hor1}$	1	1										
X_{11}	$\mathbf{RD}_{hor2}$	1	1										
X_{12}	$\mathbf{RD}_{hor3}$	1	1										
X_{13}	$\mathbf{RA}_1$												
X_{14}	$\mathbf{RA}_2$												
X_{15}	$\mathbf{RA}_3$												
X_{16}	$\mathbf{RA}_{12}$	1	1										
X_{17}	$\mathbf{W}_{P112}$					1							
X_{18}	$\mathbf{W}_{P212}$					1							
X_{19}	$\mathbf{W}_{P113}$					1							
X_{20}	$\mathbf{W}_{P313}$					1							
X_{21}	$\mathbf{W}_{SP121213}$					1							
X_{22}	$\mathbf{W}_{SP131213}$					1							
X_{23}	$\mathbf{RW}_P$			8	1.5								
X_{24}	$\mathbf{RW}_{SP}$			8	2								
X_{25}	$\mathbf{WRD}_{vert}$												
X_{26}	$\mathbf{WRD}_{hor1}$												
X_{27}	$\mathbf{WRD}_{hor2}$												
X_{28}	$\mathbf{RS}_{num}$			8	1.5								
X_{29}	$\mathbf{RS}_{geo}$												
X_{30}	$\mathbf{CWRD}_{vert}$			18	0.2								
X_{31}	$\mathbf{CWRD}_{hor1}$			18	0.2								
X_{32}	$\mathbf{CWRD}_{hor2}$			18	0.2								

The combination functions from the library used in the introduced network model are defined as follows:

- The *Euclidean combination function* $\mathbf{eucl}_{n,\lambda}(V_1, \ldots, V_k)$ is defined by

$$\mathbf{eucl}_{\boldsymbol{n},\lambda}(V_1, \ldots, V_k) = \sqrt[n]{V_1^{\boldsymbol{n}} + \ldots + V_k^{\boldsymbol{n}}} \quad (4.3)$$

where $\boldsymbol{n}$ is the order and $\boldsymbol{\lambda}$ a scaling factor and V_1, ..., V_k are the impacts from the states from which the considered state Y gets incoming connections. Note that if both parameters have value 1, then this is just the sum function and when there is only one incoming connection the identity function. This is always the case in the current model, as can be seen in role matrix **mcfp**.

- The *product combination function* **product(**V_1, V_2**)** is defined by

$$\mathbf{product}(V_1, V_2) = V_1 V_2 \quad (4.4)$$

- The *advanced logistic sum combination function* **alogistic**$_{\boldsymbol{\sigma},\boldsymbol{\tau}}$**(**$V_1$, ..., V_k**)** is defined by:

$$\mathbf{alogistic}_{\boldsymbol{\sigma},\boldsymbol{\tau}}(V_1, \ldots, V_k) = \left[\frac{1}{1 + \mathrm{e}^{-\boldsymbol{\sigma}(V_1 + \ldots + V_k - \boldsymbol{\tau})}} - \frac{1}{1 + \mathrm{e}^{\boldsymbol{\sigma},\boldsymbol{\tau}})}\right](1 + \mathrm{e}^{-\boldsymbol{\sigma},\boldsymbol{\tau}}) \quad (4.5)$$

  where $\boldsymbol{\sigma}$ is a steepness parameter and $\boldsymbol{\tau}$ a threshold parameter and V_1, ..., V_k are the impacts from the states from which the considered state Y gets incoming connections

- The *qualitative Hebbian learning combination function* **hebbqual**$_{\boldsymbol{\mu}}$**(**V_1, V_2, W**)** is defined by

$$\mathbf{hebbqual}(V_1, V_2, W) = V_1^* V_2^* (1 - W) + \mu W \quad (4.6)$$

  where $\boldsymbol{\mu}$ is a persistence parameter, W represents the weight of their connection, and V_i^* is 1 if $V_i > 0.1$ and else 0 (here V_1, V_2 are the activation levels of the connected states).

- The *complemental identity combination function* **complement-id(**V**)** is defined by

$$\mathbf{complement} - \mathbf{id}(V) = 1 - V \quad (4.7)$$

  where V is the incoming impact from a connected state

- The *max-composing combination function* **max-composition**$_{\boldsymbol{m},\boldsymbol{n}}$**(**$V_1$, V_2, V_3**)** is defined by

$$\mathbf{max} - \mathbf{composition}_{\boldsymbol{m},\boldsymbol{n}}(V_1, V_2, V_3) = \max(\mathrm{bcf}(\boldsymbol{m}, [1, 1], [V_1, V_2]), \mathrm{bcf}(\boldsymbol{n}, [1, 1], [V_3]) \quad (4.8)$$

  where bcf(i,p,v) is the i^{th} basic combination function from the library. This function composes two other combination functions from the library by using the max-function. It is actually defined as a special case using a more general function available in the combination function library that enables to create any function composition of any combination functions from the library: the function

$$\mathbf{composedbcfs}(h, p, nrs, ps, vs, ks)$$

which is defined as a function

$$\mathrm{bcf}(h, p,\ \mathrm{bcfvalues}(nrs, ps, vs, ks))$$

where m = length(nrs) is the number of functions composed with function number h, p is a list of parameter values of the composing function number h, nrs a list for the numbers of the composed functions, ps for their parameters, vs for their values and ks the numbers of their arguments, and (assuming two parameters per function)

$$\begin{aligned}\mathrm{bcfvalues}(nrs, ps, vs, ks) = [&\mathrm{bcf}(nrs(1), [ps(1), ps(2)], [vs(1), \ldots, v(ks(1))], \ldots,\\ &\mathrm{bcf}(nrs(m), [ps(2m-1), ps(2m)], \left[vs\left(1 + \Sigma_{i=1}^{m-1} ks(i)\right), \ldots,\right.\\ &\left.vs\left(\Sigma_{i=1}^{m} ks(i)\right)\right])]\end{aligned}$$

The combination function $\mathbf{eucl}_{n,\lambda}(\ldots)$ for n and λ both 1, is used to model addition, $\mathbf{product}(V_1, V_2)$ to model multiplication, $\mathbf{hebbqual}_{\mu}(V_1, V_2, W)$ to model learning of arithmetic operations, and $\mathbf{alogistic}_{\sigma,\tau}(V_1, \ldots, V_k)$ and **complement-id** (V) to model internal metacognitive monitoring and control states for the learning. The combination function $\mathbf{max\text{-}composition}_{m,n}(V_1, V_2, V_3)$ is used to reinforce the learning in the numerical mental model through the outcomes from the geometric mental model.

4.5 Example Simulation Scenarios

In this section, simulations of two example scenarios will be discussed to illustrate the introduced second-order adaptive network model. Both scenarios address the example task discussed in Sect. 3 (see also Fig. 4.1) for $a = 2$, $b = 3$, $c = 2$, which are used as constant values for base states $\mathbf{N}_1$, $\mathbf{N}_2$, and $\mathbf{N}_3$, respectively. The first scenario shows how someone who has good arithmetic skills addresses the task, without involving any switch to the geometric mental model; see Fig. 4.3. As can be seen, as one of the first, state $\mathbf{S}_{23}$ comes up which determines the sum of $\mathbf{N}_2$ representing b and $\mathbf{N}_3$ representing c, which correctly ends up in value 5 (the blue line). At about the same time state $\mathbf{P}_{12}$ (the red line) for the product of $\mathbf{N}_1$ and $\mathbf{N}_2$

representing a and b comes up, correctly ending up at 6. Similarly, $\mathbf{P}_{13}$ (the blue-green line) for the product of $\mathbf{N}_1$ (for a) and $\mathbf{N}_2$ (for c) correctly reaches 4.

Next, $\mathbf{PS}_{123}$ of $\mathbf{N}_1$ and $\mathbf{S}_{23}$ representing the product of a and $b + c$ is determined, which correctly ends up in 10 (the light dark green line). The determines the left hand side of the Eq. (4.2). At the same time, the right hand side of (2) is addressed. Therefore, $\mathbf{SP}_{1213}$ (again the dark green line) for the sum of $\mathbf{P}_{12}$ and $\mathbf{P}_{13}$ comes up and correctly reaches 10. This shows that the right hand side of (2) is indeed equal to the left hand side of (2), what solves the task. In the meantime it can be seen in Fig. 4.3 that the self-model about the numerical mental model is formed: the two lines for the two **RW**-states all end up at 1, and also based on them the third (orange) line for $\mathbf{RS}_{\text{num}}$, which as a form of metacognitive monitoring tells the learner that the arithmetic skills are OK. Therefore, in this case no control decision to switch to the geometric mental model is made, and also no further learning is needed.

The second scenario is the more interesting one (see Fig. 4.4). Here the learner has still good arithmetic skills (connection weights 1) to address the left hand side of (2), but not for the right hand side (connection weights are only 0.1). Therefore the light brown and purple lines in the upper graph in Fig. 4.4 are the same as in Fig. 4.3, but not the lines for $\mathbf{P}_{12}$, $\mathbf{P}_{13}$, and $\mathbf{SP}_{1213}$ needed for the right hand side of (2). Because that side gets stuck, and the self-model used for monitoring has low values showing a lack of arithmetic skills, the control decision is made to switch to the geometric mental model: all three **CWRD**-states come up soon and reach 1 shortly after time 5 (the purple line in the lower graph of Fig. 4.4). As a consequence, to execute this control decision, the **WRD**-states become 1 around time 5 (the red line in the lower graph of Fig. 4.4).

Because of that the **RD**-states representing the dimensions of the rectangles get their values 2, 3, and 5. Based on these, the **RA**-states for the areas of the rectangles are determined and get their values 4, 6 and 10. As these **RA**-states provide a reinforcing impact on the states $\mathbf{P}_{12}$, $\mathbf{P}_{13}$, and $\mathbf{SP}_{1213}$ in the numerical mental model,

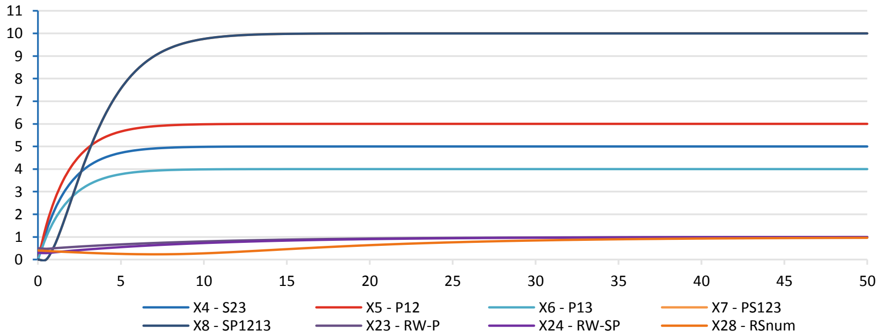

Fig. 4.3 Using the arithmetic mental model and formation of the self-model for metacognitive monitoring

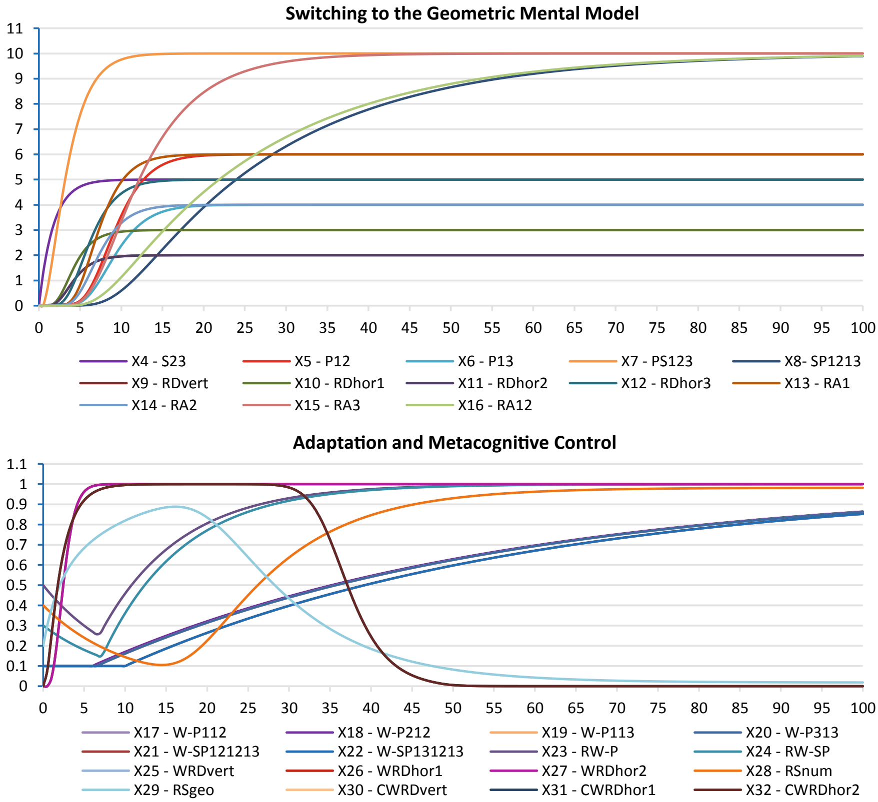

Fig. 4.4 Switching to the geometric mental model and adaptation and metacognitive control of it

it can be seen that with a small delay the latter states follow the **RA**-states to also reach values 4, 6 and 10 (the red, light blue, and dark green line in Fig. 4.4, upper graph). In the lower graph of Fig. 4.4 it can be seen what happens further concerning the adaptation levels. The lines starting at 0.1 are the **W**-states, and it is shown that after time 6 they start to increase to finally reach values close to 1. This is the reinforced Hebbian learning process for the numerical mental model: reinforced by the impact from the geometric mental model. Also the two **RW**-states and state $\mathbf{RS}_{\text{num}}$ for the self-model for the numerical mental model, starting at 0.3, 0.5 and 0.4, increase after time 6. Note that the $\mathbf{RS}_{\text{geo}}$ (light green line with peak near 0.9) also increases thereby supporting the decision to switch to the geometric mental model, but later on (after time 15) goes down just like the **CWRD**-states for the control themselves do (after time 25), as after learning the full arithmetic mental model, by the monitoring via the self-model the learner feels that there is no reason anymore to consider switching to the geometric mental model.

4.6 Discussion

Learning processes can be described by adaptive mental (or neural) network models. If metacognition is used to regulate learning (Pintrich 2000), the adaptation of the mental network becomes itself adaptive as well, so then it involves second-order adaptation. In this chapter, a second-order adaptive mental network model was introduced for metacognitive regulation of learning processes using multiple internal mental models. Part of the material was adopted from (Treur 2021).

The focus was on the role of multiple mental models (Gentner and Stevens 1983; Greca and Moreira 2000; Skemp 1971; Seel 2006), in particular, the case of visualisation to support learning of numerical or symbolic skills (Bruner 1966, 1977; Bidwell 1972; Day and Hurell 2015; Du Plooy 2016; Freudenthal 1973,1986; Koedinger and Terao 2002; Larbi and Mavis 2016; Lovitt et al. 1984; Renkema 2019; Roberts 1989). The second-order adaptive network model was illustrated for the role of visualisation to support learning multiplication at the primary school.

It was shown how a second-order self-modeling network model provides adequate means to model the different aspects that make the addressed topic complex: the network has a self-model about its own structure, it models mental models and their adaptation for learning, and it models dynamic metacognitive control of this adaptation. The model was applied to simulate some example scenarios that illustrate what the model does. In further work other scenarios can be addressed as well.

References

Bidwell, J.K.: A physical model for factoring quadratic polynomials. Math. Teach. **65**(3), 201–205 (1972)

Bruner, J.S.: Towards a theory of instruction. Harvard University, Cambridge, Mass (1966)

Bruner, J.S.: The Process of Education. Harvard University Press, London (1977)

Darling-Hammond, L., Austin, K., Cheung, M., Martin, D.: Thinking about thinking: Metacognition. (2008)

Day, L., Hurrell, D.: An explanation for the use of arrays to promote the understanding of mental strategies for multiplication. Aust. Prim. Math. Classr. **20**(1), 20–23 (2015)

Du Plooy, M.C.: Visualisation as a metacognitive strategy in learning multiplicative concepts: a design research intervention. Ph.D. Thesis, University of Pretoria, Department of Mathematics Education**.** https://repository.up.ac.za/handle/2263/51258 (2016)

Flavell, J.H.: Metacognition and cognitive monitoring: a new area of cognitive–developmental inquiry. Am. Psychol. **34**(10), 906–911 (1979)

Freudental, H.: Mathematics as an Educational Task. Reidel, Dordrecht (1973)

Freudental, H.: Didactical Phenomenology of Mathematical Structures. Reidel, Dordrecht (1986)

Gentner, D., Stevens, A.L.: Mental models. Erlbaum, Hillsdale NJ (1983)

Greca, I.M., Moreira, M.A.: Mental models, conceptual models, and modelling. Int. J. Sci. Educ. **22**(1), 1–11 (2000)

Hebb, D.O.: The organization of behavior: a neuropsychological theory. Wiley (1949)

Koedinger, K.R., Terao, A.: A Cognitive Task Analysis of Using Pictures To Support Pre-Algebraic Reasoning. In: Gray, W.D., Schunn, C.D. (eds.), Proceedings of the 24th Annual Conference of the Cognitive Science Society, CogSci'02, pp. 542–547. Lawrence Erlbaum Associates, Mahwah, NJ (2002)

Koriat, A.: Metacognition and consciousness. In: Zelavo, P.D., Moscovitch, M., Thompson, E. (eds.). Cambridge Handbook of Consciousness. New York, Cambridge University Press (2007)

Larbi, E., Mavis, O.: The Use of manipulatives in mathematics education. J. Educ. Pract. **7**(36), 53–61 (2016)

Lovitt, C., Marriot, C., Swan, K.: Lessons in algebra using algebra blocks. Laverton, Victoria: EDU-DOMES (1984)

Mahdavi, M.: An overview: metacognition in education. Int. J. Multidis. Current Res. **2**, 529–535 (2014)

Pintrich, P.R.: The role of goal orientation in self-regulated learning. In: Boekaerts, M., Pintrich, P., Zeidner, M.(eds.), Handbook of self-regulation research and applications, pp. 451–502. Academic Press, Orlando, FL (2000).

Renkema, W.: Effectiveness of Dynamic Visualisation in Video-based Animated Algebra Instruction. Thesis, Utrecht University, M.Sc (2019)

Rivera, F.D.: Visualization and Progressive Schematization: Framing the Issues. In: Toward a Visually-Oriented School Mathematics Curriculum. Mathematics Education Library, vol 49, pp. 21–58. Springer, Dordrecht. https://doi-org.vu-nl.idm.oclc.org/https://doi.org/10.1007/978-94-007-0014-7_2 (2011).

Rivera, F.D.: Toward a visually-oriented school mathematics curriculum: research, theory, practice, and issues. Math. Educ. Library **49**. https://doi.org/10.1007/978-94-007-0014-7_2 (2011)

Roberts, B.S.: The effects of algebra blocks on student achievement in algebra. B.Sc.Thesis, Edith Cowan University. https://ro.ecu.edu.au/theses_hons/173 (1989)

Seel, N.M.: Mental models in learning situations. In: Advances in Psychology, vol. 138, pp. 85–107. North-Holland, Amsterdam (2006)

Shannon, S.V.: Using metacognitive strategies and learning styles to create self-directed learners. Inst. Learning Styles J. **1**, 14–28 (2008)

Skemp, R.R.: The Psychology of Learning Mathematics. Penguin Books, Harmondsworth (1971)

Treur, J.: Network-Oriented Modeling: Addressing Complexity of Cognitive. Springer Publishers, Affective and Social Interactions (2016)

Treur, J.: Multilevel network reification: representing higher order adaptivity in a network. In: Aiello, L., Cherifi, C., Cherifi, H., Lambiotte, R., Lió, P., Rocha, L. (eds), Proc. of the 7th Int. Conf. on Complex Networks and their Applications, ComplexNetworks'18, vol. 1. Studies in Computational Intelligence, vol. 812, pp. 635–651, Springer Nature (2018)

Treur, J.: Modeling higher-order adaptivity of a network by multilevel network reification. Network Science **8**, S110–S144 (2020a)

Treur, J.: Network-oriented modeling for adaptive networks: Designing Higher-order Adaptive Biological, Mental and Social Network Models. Springer Nature Publishing, Cham, Switzerland (2020b)

Treur, J.: An adaptive network model covering metacognition to control adaptation for multiple mental models. Cogn. Syst. Res. **67**, 18–27 (2021)