

Novices' Reasoning Processes in a Psychological Modeling Task

Bachelor thesis

Psychology

Author: Otto Stellmacher (s0157465)
Supervisors: Prof. Dr. W. R. van Joolingen
Dr. A. Weinberger
Date: 03.07.2009



University of Twente

Abstract

This study investigated novices' reasoning processes during a psychological modeling task. Subjects of the experiment have been ten psychology novices. The task was to model of a psychological phenomenon, called 'sentence verification'. Data was gathered by the think-aloud technique. The results show consistent matches of all participants. Based on quantity and quality, models were grouped in strong and weak models. The Man-Whitney-U test showed that the strong group spent significant more time on orientation than the weak group. These findings are concordant with former research on modeling. However, interventions by the experimenter were found in the weak group more often. No significant differences were found for the other categories. In conclusion it can be stated orientation is an important activity during scientific reasoning, with implications for practice. A better comprehension and quality of the models can be achieved by supporting novices in their orientation process of modeling tasks.

Table of contents

Abstract.....	2
Table of contents	3
1. Introduction	4
1.1 <i>Science and scientific reasoning</i>	4
1.2 <i>What are models and modeling?</i>	4
1.3 <i>The role of models in science</i>	5
1.4 <i>Models in science education</i>	6
1.5 <i>The need for exploration</i>	6
2. Method	7
2.1 <i>Participants</i>	7
2.2 <i>Materials</i>	7
2.3 <i>Procedure</i>	7
2.4 <i>Data collection and analysis</i>	7
3. Results.....	9
3.1 <i>Reasoning processes</i>	9
3.2 <i>Modeling strategy</i>	10
3.3 <i>Models</i>	12
4. Discussion	13
5. References	14
Appendix A: Experiment material.....	16
Appendix B: Figures	19
Appendix C: Tables.....	27

1. Introduction

1.1 Science and scientific reasoning

Scientific reasoning plays an important role in psychology. Apart from being scientists themselves psychologists also want to study the products and the processes of science. Products in this case are the individual concepts and knowledge acquired. Processes, for example, are deducing, building, testing and evaluating hypotheses, theories or models. Research on these two different fields of processes and products often are divided in the exploration of two types of knowledge: domain-specific knowledge and domain-general strategies (Zimmerman, 2000).

The domain-specific knowledge approach deals with the content of scientific reasoning. What are the concepts, components and relations about phenomena in different domains? These elements constitute 'mental models'. Researchers try to analyze the difference between the models of people (Zimmerman, 2000). Domain-general strategies are concerned with the process of scientific reasoning. What are the underlying principles of model or theory building? Which reasoning strategies are used when solving a problem? (Koslowski, 1996)

For example, Zimmerman (2000) mentions designing experiments and evaluating results as basic skills for that kind of reasoning.

In general, scientific reasoning is the process of generating, testing, evaluating and interpreting theories and hypotheses. These processes must be learned since they are not a part of everyday reasoning strategies. Even if we reason about events, causes and relations, we do not build and evaluate theories consciously (Sodian & Bullock, 2008).

Thus, reasoning is a paramount cognitive ability, which is not innate in students (Ates & Cataloglu, 2007). Before prospective scientists can use this kind of reasoning as a thinking tool, they must pass through several steps. They think about descriptions of the world and build up small concepts and schemes. The more they find out about the relations of the concepts, the larger the schemes get. These schemes, when sufficiently developed, enable hypothetical and scientific reasoning. In consequence, students develop a better ability for doing science and the construction of science concepts (Lawson, Clark, Meldrum, Falconer, Sequist, & Kwon, 2000). Hence, Lawson et al. (2000) found a positive relationship between reasoning skills and achievement in education.

1.2 What are models and modeling?

Mental models are an important part of cognition and they can explain the development and organization of cognitive abilities. The origin goes back to Craik's work in 1943 who describes mental models as a storage tool for the content of thought (Craik, 1943). It represents what is thought rather than how it helps to understand something. This point of view is inducted by the second approach, which was introduced by Johnson-Laird (1983). He wanted to explain the process of building mental models as an instrument for thinking.

Bliss (1994) describes two questions: how are thoughts internalized and how do they return to an externalization. Both questions also refer to mental models.

Internalization is the process of acquiring and memorizing knowledge from the world. There are two different kinds of this process. First, knowledge can be instructed (explicit learning) or it can be internalized by an unconscious, implicit process. The structure of internalized knowledge and mental models are unconscious as well.

Externalization is the reverse process. It takes place mostly through the use of speech and actions. Thus, mental models are translated into an external representation which make it easier to reflect or think about the modeled phenomena (Bliss, 1994).

Models can be differentiated in physical models and conceptual models. Whereas physical models always require an external representation, however, conceptual models are models of the mind. Therefore, they do not need such a concrete representation (Webb, 1993).

Beside the distinctions in physical and conceptual models, one more is made by Clement (1989). He distanced expedient models and explanatory models. In the first category, the model and the

phenomenon have the same properties. However, it does not show the underlying properties and processes. This is what the exploratory model does. It shows how phenomena work and what the relations between the underlying processes are. These exploratory models can shape and improve science education (Clement, 1989). Thus, physical model have two practical roles. First, the model makes underlying relations and components visible so that this information can be used to investigate the phenomenon. Second, by investigating the model, it is possible to reason from the model about the phenomenon. The closer the model resembles the phenomenon, the better the world can be understood (Gregg & Simon, 1967).

The process of constructing a model is called modeling. It starts with the students' basic understanding of the world. These understandings direct how the student extrapolates components and relations of the phenomenon. In doing so, the student constructs, tests, interprets a model and, depending on the results, modifies it as well. That makes the student think about the underlying relations of a phenomenon and this should lead to an understanding of scientific practice (Penner, 2000).

In summary, models can appear in several forms: concepts in mind, physical objects to analyze phenomena and theories, concrete examples or tools for solving problems. Since models can be differentiated in various types, there are implications for the influences on science.

1.3 The role of models in science

Models are one of the basic instruments in modern science. Since the fundamental goal of science is the study and explanation of phenomena in the world, there are hardly any domains, in which models do not have influences. These influences emerge because of the different functions of models (Frigg & Hartmann, 2008).

Frigg and Hartmann (2008) present an outline of the types, the functioning and the impacts of models in science (for example scale models, idealized models, analogical models, phenomenological models, mathematical models, models of data, etc.). The answer to the question of ontology (what models actually are) is not easy as well, because of the different possibilities to define models. According to the authors, models are physical objects, fictional objects, set-theoretic structures, descriptions and equations. Even if there are some critics to all of these forms and types, there is no doubt that such manifoldness exists. Hence, there is a strong association between models and science theories (Frigg & Hartmann, 2008).

In psychology, models are of crucial importance. Showing the relation between mental concepts and behavior is one basic and historical important issue in this context. Without referring to models and representations, this question cannot be answered, because further considerations lead to the question of concept representation or the structure of cognition in general. The two main views in psychology are connectionism and the view of humans as symbol systems. Representation of knowledge is another essential field of psychology, since understanding these functions can lead to improved learning (Suppes, Pavel, & Falmagne, 1994). One of the representations is for instance Newell and Simon's production rules (1972). Furthermore, models and representations are indispensable in the field of measurement theory. Only with a good numerical representation of an empirical, observed structure, predictions or conclusions can be deduced. With the conclusions from measurement results, individuals, can be classified into certain dimensions (Suppes, Pavel, & Falmagne, 1994). The description of objects in terms of attributes certainly plays an important role in psychology.

There are several reasoning processes given by Sins, Savelsbergh and van Joolingen (2005) that are important in modeling: analyzing, inductive reasoning, quantifying, explaining and evaluating. While analyzing the phenomenon, the modeler tries to find all important components and relations. These are used as elements of the model. When the modeler examines if his model works in the way it should, inductive reasoning takes place. He needs to consider the relationship between the modeled phenomenon and the model itself. To make a model computable a process of quantification is needed. This more precise model can be tested in a mathematical manner. However, it is not a process that can always be found in modeling. In the explanation process, elements and their

relations are shown. Finally, in the evaluation process, the modeler tests his model against the phenomenon. This can (and should) lead to changes and improvements in the model (Sins, Savelsbergh, & van Joolingen, 2005).

Thus, in knowledge acquisition, models are an important feature, or as Frigg and Hartmann (2008) state: 'Models are vehicles for learning about the world.'

Since 'the most general goal of scientific investigation is to extend the knowledge of the world' (Zimmerman, 2000, p. 101) which can be accomplished by models, it is obvious that they are one of the most important components of scientific research.

1.4 Models in science education

According to Penner (2000), in educational science too many passive processes take place. Students should learn during experiments or demonstrations. However, these demonstrations only reproduce effects and findings (Penner, 2000). The common approach of science education is different from science in reality. For instance, the differences are that science is a dynamic process of creating and testing rather than a pure collection of facts. Making educational science more effective, students should be an active part of the processes involved in explanation and understanding (Holland, 1998). That can be described as inquiry learning.

White and Shimoda (1999) give one structure of ideal inquiry learning. This inquiry cycle includes hypothesizing, investigating, analyzing, modeling and evaluating. Besides this classification, many other structures exist. All have in common a same superior meaning of orientation, hypothesizing, experimenting, model evaluation and documentation. Since modeling can appear in the same structure it can be linked to inquiry learning. Whereas the inquiry cycle of White and Shimoda take modeling as a separate part, it can represent a whole cycle as well (Löhner, van Joolingen, Savelsbergh, & van Hout-Wolters, 2005).

Hence, models and modeling are one method of accomplishing such a dynamic process of investigation (Webb, 1993). Thus, students should build own models instead of trying to understand existing models (as it is usual in common education). According to Holland (1998), modeling processes have several advantages: for instance, students learn to select important components, which provide them an objective and reflective view about their models.

This different approach increases the knowledge but especially the understanding. This process could enrich learning sessions. The students' thinking could also change, because they have available a tool to work with scientific phenomena. That makes it easier to think about scientific topics in general as well (Stratford, Krajcik, & Soloway, 1998).

Sins, Savelsbergh and van Joolingen (2005) present an outline about the importance of modeling. Thus, modeling is a technique for building external representations and supports learning through the process of the modeling task. Students learn to think in a scientific way. For instance, this contains testing and reflecting their ideas (Sins, Savelsbergh, & van Joolingen, 2005).

Stratford et al. (1998) support this view. The authors define several 'Cognitive Strategies for Modeling' (p. 216), which include analyzing, relational reasoning, synthesizing, testing/debugging, and explaining. Following these steps, students should obtain a functioning model of the phenomenon. That includes the explanation of the underlying relationships and components.

1.5 The need for exploration

The importance of modeling and models in science leads to the need for exploration. As shown in the text above reasoning processes in modeling tasks were investigated; research pointed out the positive influence on understanding and insight of phenomena. Altogether, it clarified the important components of models and modeling and showed how they support domain specific knowledge. However, it is not enough to state the advantages of models and modeling. There is also the need to find out what happens during real behavior and if there are differences between modeling in one domain and another.

Numerous studies have been concerned with the research of reasoning strategies during modeling tasks in specific domains. Nevertheless, the domain of psychology has not been a target domain of the research so far.

That is surprising, because psychology is the central domain of cognitive research and uses models frequently to explain theories. Since it is vital for psychology to construct models, it would be interesting to investigate the process of modeling.

The basic question is how novices behave during the modeling process. The importance of models and modeling in scientific education is shown. Therefore, if learning sessions can benefit from models and modeling, it should be known how students use these processes. With more knowledge, researchers or teachers could provide more efficient help for students during modeling tasks. The research question therefore is:

‘What are the reasoning strategies of psychology novices during modeling a psychological domain?’

2. Method

2.1 Participants

Ten psychology novices participated in the study, all second and third year students at the University of Twente. All of them had already attended basic courses in cognitive psychology, thus they have been familiar with the research topic.

2.2 Materials

The modeling task was derived from a psychological experiment called ‘sentence verification’. The theory behind the experiment contains the assumption, that the farther two concepts (or properties) in a sentence lie apart in our semantic network, the longer it takes to recognize the truth-value of this sentence. The students had access to the experimental task and thus could experience the tasks from the viewpoint of an experimental subject. The task was provided by an interactional learning environment of the University of Twente, called ZAPs (Hulshof, Eysink, & de Jong, 2006).

The instructions included in addition to the theory and background information also the influential variables. The task for the participants was to explain the reaction time by making a model.

The instructions also included two training tasks. One puzzle to train the think aloud technique and one small modeling task to get to know the required form.

Paper and pencil were used to build the models, which seemed to be more practical compared to a computer program. Everything was recorded with a webcam and a microphone.

2.3 Procedure

The experiment started with a short introduction of modeling and information about the aim of the study. The participants started with the two small exercises, practicing the think aloud technique and making models. After that, they went over to proceed the ZAP experiment. It was a small simulation of the phenomenon that needed to be modeled later in the experiment. It should give the participants a feeling of the construct they were about to work with. Then they started with the modeling task without a time constraint. The experiment ended as soon as they stated that they could not find any model improvements.

2.4 Data collection and analysis

The think aloud protocols were split into text segments based on the participants natural breaks. They were coded in different categories based on the scheme created by Löhner et al. (2005). The first main category, scientific reasoning, has five sub-categories: orientation, hypothesizing, experimenting, model implementation and model evaluation. The second main category, other activities, has four sub-categories: actions, regulation, off tasks and experimenter. Depending on the length and content, the segments also were assigned to several categories. In the coding process, all

subcategories were taken into account. However, the analysis of the data refers to the main categories. The complete coding scheme and examples can be found in table 1.

Table 1
Coding Scheme based on Löhner et al. (2005)

Scientific Reasoning	
Orientation - Defining variables - Domain talk - Experience knowledge - Theoretical knowledge - Refer to instruction	<i>'Yes, according to me pattern matching amounts to the semantic gap, which is zero.'</i>
Hypothesizing - Predictions - Hypothesis generation	<i>'Ahm, the semantic gap has a positive influence, anyway. Thus, if the semantic gap is larger, then the reaction time should be longer, too.'</i>
Experimenting	<i>'Imagine you have a certain statement: canary is a bird. That is S-1. This is solved faster, than the P-sentence with the same semantic gap.'</i>
Model implementation	<i>'The arrow needs to be drawn from abstraction to the semantic gap and not directly to the reaction time.'</i>
Model evaluation - Interpretation model output - o Concluding - o Describing - model evaluation	<i>'Ahm, I cannot explain this with my current system.'</i>
Other activities	
Actions - Model syntax - Tool is not working - Tool use - Reading - Calculating	-
Regulation - Planning - Choose activity - Evaluation - Task - Frustration	<i>'I am going to read the text again to find things I forgot.'</i>
Off task	<i>'I think Otto is hungry.'</i>
Experimenter	<i>'What do you think?'</i>

Participant 5 needed to be excluded from the remaining analysis because of difficulties with the think aloud technique. A second rater coded the text segments of the nine useable participants, as well. The inter-rater reliability of the main categories was good (Cohen's kappa = 0.78). The mean time in percent of the single categories was computed and graphs of the strategies' process curves were plotted. The percentage of time was placed on the X-axis, the percentage of

shared processes on the Y-axis. In order to find differences between the participants, the models were classified in strong and weak models. Criteria were both the number of used variables and the importance of these variables for the model. The functioning of the model was taken into account as well. Thus, if the model showed important components of the phenomenon as well as the relation between them, it was a strong model. The less important components and relations were found, the weaker the model was. Consequently, strong models were a good representation of the phenomenon. Finally, five participants were assigned to the strong group and four to the weak group. These two groups were then compared using a Mann-Whitney U test.

3. Results

3.1 Reasoning processes

Table 2 shows the percentage of time for each participant and category. Almost all participants spent slightly more than half of the time on scientific reasoning and the other half on other activities. An exception was participant 6. He spent 40 % of the time on scientific reasoning and 60 % on other activities. Within scientific reasoning, the largest category was orientation with a mean of 16.6 % and a standard deviation of 6.4. The largest category within other activities was actions (32.0 %, SD = 5.7) and is thus the largest category in all. There are partly large differences between participants. For instance, participant 10 spent 29.2 % of the time on orientation, whereas participant 6 spent only 9.5 %.

Table 2
Percentage of time participants spent on reasoning processes

Participant	1	2	3	4	6	7	8	9	10	mean
Scientific reasoning (%)	47.1	56.2	51.5	52.9	40.0	55.4	50.5	44.7	51.2	49.9
Orientation (%)	17.6	23.8	12.1	17.1	9.5	15.8	10.3	13.6	29.2	16.6
Hypothesizing (%)	6.5	6.9	8.7	11.2	5.7	6.3	15.5	5.3	8.6	8.3
Experimenting (%)	3.9	0.0	6.8	4.7	2.9	4.3	2.1	4.2	0.5	3.3
Model implementation (%)	13.7	19.2	11.2	12.9	12.4	16.0	14.4	12.5	8.6	13.4
Model evaluation (%)	5.2	6.2	12.6	7.1	9.5	13.0	8.2	9.1	4.3	8.4
Other activities (%)	52.9	43.8	48.5	47.1	60.0	44.6	49.5	55.3	48.8	50.1
Actions (%)	32.7	39.2	24.8	35.9	36.2	22.3	35.1	28.0	34.0	32.0
Regulation (%)	9.2	3.1	12.6	8.2	13.3	16.0	6.2	13.6	12.0	10.5
Off tasks (%)	3.3	1.5	2.4	0.0	4.8	4.1	2.1	4.9	1.9	2.8
Experimenter (%)	7.8	0.0	8.7	2.9	5.7	2.2	6.2	8.7	1.0	4.8

As a consequence of these differences, the models were arranged according to quality. Table 3 shows the comparison of the upper 5 models and the 4 at the bottom at the arrangement. The Mann-Whitney-U test ($n = 9$) indicates two significant differences. Participants with strong models spent more time in the orientation category (20.7 %, SD = 5.7) than participants with weak models (11.4 %, SD = 1.8; $U = 0.00$, $p = 0.014$). Interventions by the experimenter were less found in the first group (2.8 %, SD = 3.0) than in the second group (7.3 %, SD = 1.6; $U = 2.00$, $p = 0.049$).

Table 3

Percentage of time the strong group and weak group spent on the processes during the task

	strong group	weak group			
	M (SD)	M (SD)	U	p	
Orientation (%)	20.7 (5.7)	11.4 (1.8)	0.00	0.014	*
Hypothesizing (%)	7.9 (2.1)	8.8 (4.7)	9.00	0.806	
Experimenting (%)	2.6 (2.2)	4.0 (2.1)	8.00	0.624	
Model implementation (%)	14.1 (3.9)	12.6 (1.3)	6.00	0.327	
Model evaluation (%)	7.2 (3.4)	9.8 (1.9)	4.00	0.142	
Actions (%)	32.8 (6.4)	31.0 (5.5)	9.00	0.806	
Regulation (%)	9.7 (4.8)	11.4 (3.5)	7.00	0.462	
Off tasks (%)	2.2 (1.6)	3.5 (1.5)	4.00	0.142	
Experimenter (%)	2.8 (3.0)	7.3 (1.6)	2.00	0.049	*

U and p scores were obtained with a non-parametric Mann-Whitney U-test

**p = 0.05*

3.2 Modeling strategy

In order to find regularities in the strategy, a timeline was computed for the different processes. Figure 1a – 1i (Appendix B) show the participants' reasoning strategies. All figures start with high orientation, which slightly decreases over time, but with always recurring peaks (above all at the end). Exceptions are the figures of participants 4, 7 and 6. Here are the highest peaks in the middle of the timeline or at the end. Whereas orientation decreases, implementation increases. However, this effect is not very clear. Approximately 5 % - 15 % after the first increase of implementation evaluation starts increasing as well. It is noticeable that the slope is almost the same as by the preceding implementation. Exceptions are the figures of participant 2, 7 and 9. A clear and large increase in evaluation can be seen in the middle. The smallest modification has experimenting. In the figures of participant 1 and 2, experimenting is even completely missing. No regularity can be found for hypothesizing.

The summary of all strategies are shown in figure 2 (Appendix B). Figure 3a and 3b show the different strategies of the strong and weak model group. It is obvious that the strong group has a different strategy than the weak group. The strong group has over the complete time line higher and more orientation values, whereas the weak group has a large drop in the first part. The implementation curve makes relatively large changes as well as the evaluation curve. However, the strong group holds in these categories a much more regular level. Larger amplitudes are found more often in the weak group than in the relatively consistently strong group.

Figure 3a

Progressing average of relative frequencies of reasoning processes in the strong group

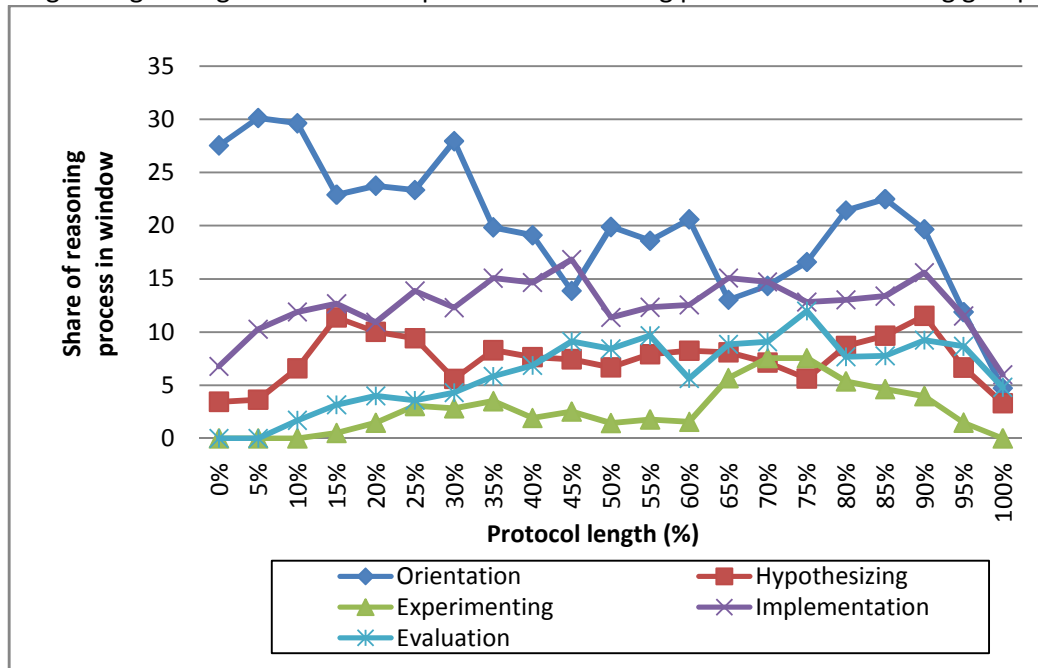
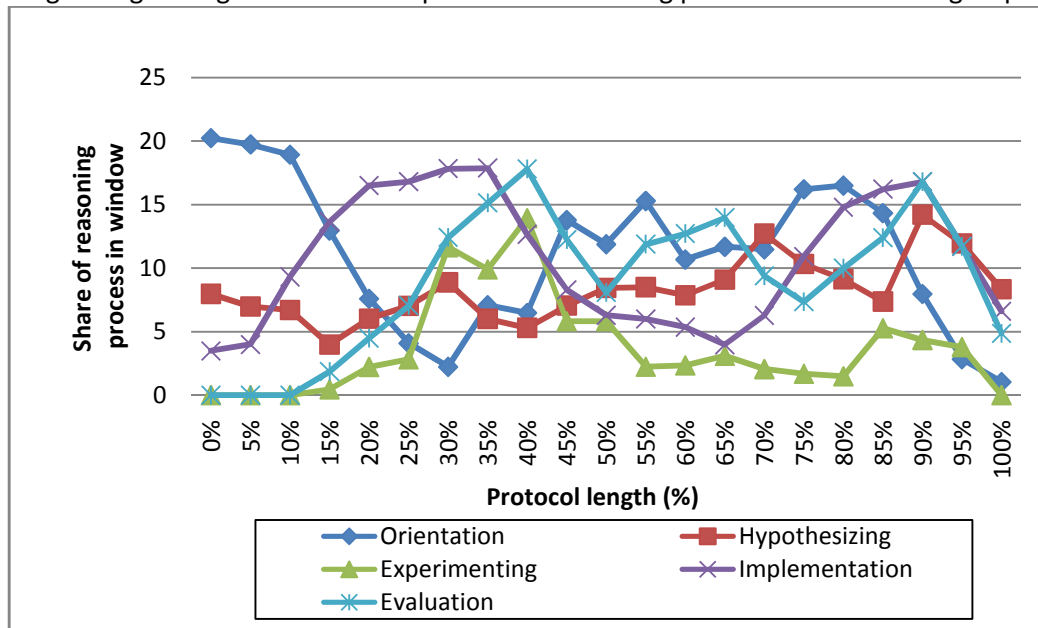


Figure 3b

Progressing average of relative frequencies of reasoning processes in the weak group



3.3 Models

Figures 4a and 4b show the two models that were rated best and worst, as typical examples of these model. All other models are included in Appendix B. The model of participant 2 is more complex in comparison to the model of participant 3. The participant used more different variables and precisely specified the relations between them. The importance of the used variables differs as well. Whereas participant 2 found all variables that are important for a good functioning of the model, participant 3 used only four important variables ('P', 'S', 'B' and 'R') but added two non-essential variables ('A' and 'D').

Figure 4a

Model of participant 2: example of a strong model

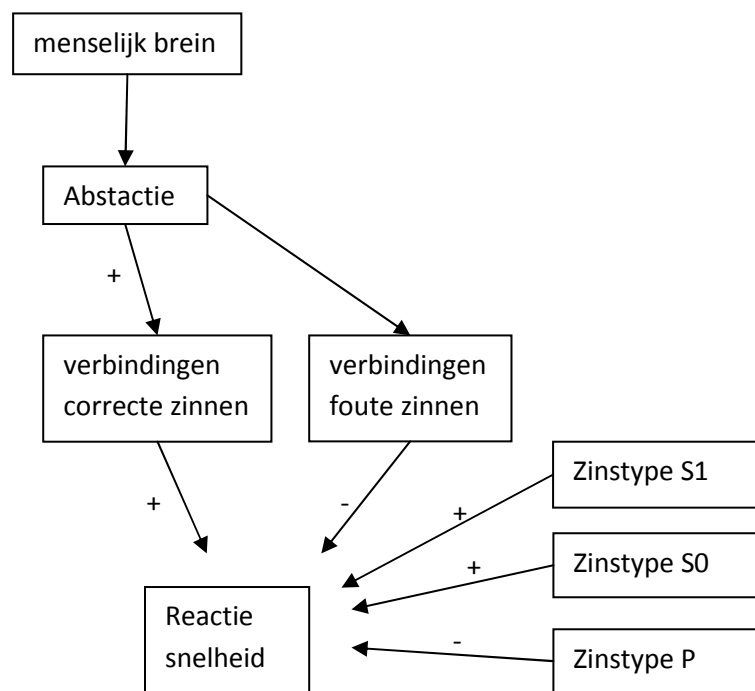
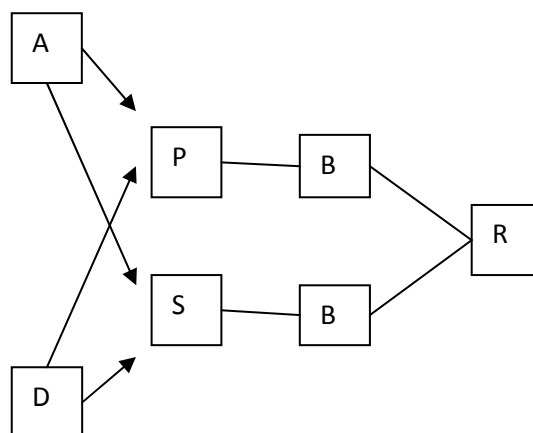


Figure 4b

Model of participant 3: example of a weak model



4. Discussion

The research question of this study was:

‘What are the reasoning strategies of psychology novices during modeling a psychological domain?’

The results show that there is a strategy during the process of modeling. Since orientation represents the largest category in the scientific reasoning domain, further analysis seems justified. It shows the highest peaks over the whole time line. As orientation decreases, implementation increases. The parallel run of the implementation and evaluation curve is also conspicuous. Almost no high peaks are visible at the experimenting curve during the whole time line.

The arrangement in strong and weak models emphasizes the importance of orientation. Whereas the high orientation remains in the strong group over time, it disappears in the weak group. Since there were found only significant differences between the groups for orientation, the other changes need to be carefully analyzed. It might be that these differences play a role in understanding and achieving a modeling task, but based on these findings there cannot be made any deductions. However, since the significance test did not take the time line into account (only the amount of usage) an attempt can be made to explain the distribution. Hence, without a good orientation, participants interrupt more often and try different things until they find something sufficient. That might be a possible explanation for the fluctuation of the other processes.

The importance of orientation is supported by a work of Van den Broek (2007). He studied modeling behavior of experts and novices in the domain of management. He found that experts spent more time on orientation than novices did. A further conformity of the results is the mean orientation curve of his expert modelers and the orientation curve of the strong group of this study. Both curves had equal runs and their peaks nearly at the same periods of time. However, the strong group of this study shared much more processes at all periods.

This study contained only novices but some of them acted like experts. This might be through the type of the task. Van den Broek (2007) used a difficult, ill-defined task, whereas in this study a relatively easy and less complex task had to be solved. Thus, it was easier to become an expert. The structure and content of the task might also be responsible for the large amount of time the participants spent on other activities. All information was given in text form and as a consequence participants read a lot while solving in the task. As most of the participants read silently, it had implications for recording the data. Possible additional activities may not have been recorded. In particular, model evaluation could have had influence that does not emerge in the analysis. Participants could test and evaluate their models only by reading the information and silently comparing it to their models. The same applies to orientation, experimenting and hypothesizing. Participants might have got ideas during the reading process and could have silently tested them against the text.

Experimenting was the smallest category in scientific reasoning. Here again, the relatively simple task structure could have played an important role. They did not see the need for experimentation, as they easily comprehended their models.

This study has some limitations and suggestions for future studies in this research field. First, processes may have taken place silently. If the participants read the text aloud, some of possible hidden processes could have been detected.

Another limitation of this study concerns the small complexity. Almost no participants were engaged in the task very long, which has implications for the data collection. If they were engaged longer, the additional gained data could have generated different and more significant outcomes.

Furthermore, it can be criticized that the nine models were arranged in 5 strong models and 4 weak models. The model in the middle of this arrangement also could be assigned to the weak group. However, if the model excluded from the analysis, no essential changes in the results are found. The significant difference regarding interventions of the experimenter disappears ($U = 2.00$, $p = 0.114$), but the important significant difference of orientation remains ($U = 0.00$, $p = 0.021$). There are also

no changes in the strategy of the strong group. All process curves maintain the same run. Thus, there are no serious implications for the conclusion. The table comparing the strong group and the weak group without the fifth model is shown in table 4 (Appendix C).

In summary, there are obvious patterns in the modeling process. The orientation pattern seems to be most important, since it is positive associated with the quality of the models. The results give a notion what novices do during modeling a psychological domain. Nevertheless, they are only tendencies and therefore they need to be treated carefully. Suggestion for supporting novices in modeling tasks can only be made for orientation activity. Based on the results, it is assumed, supporting the orientation process could enhance the quality of the models. If novices get support in defining variables or talking about the domain, it will possibly lead to better models. In respect of science education, this could lead to better comprehension of the issue, thus resulting in better learning performance. Hence, optimizing the support for students referring to all categories, it can have considerably positive implications for science education in general. However, further research is required to confirm the results of this study and to clarify the roles of the other categories.

5. References

- Ates, S., & Cataloglu, E. (2007). The effects of students' reasoning abilities on conceptual understandings and problem-solving skills in introductory mechanics. *European Journal of Physics* , 1161-1171.
- Bliss, J. (1994). From mental models to modeling. In J. Bliss, H. Mellar, R. Boohan, J. Ogborn, & C. Timpsett, *Learning with artificial worlds: computer based modeling in the curriculum* (pp. 27-32). London: The Falmer press.
- Clement, J. (1989). Learning via model construction and criticism. In J. A. Glover, R. R. Ronning, & C. R. Reynolds, *Handbook of creativity* (pp. 341-381). New York: Plenum.
- Craik, K. J. (1943). *The Nature of Explanation*. Cambridge: University Press.
- Frigg, R., & Hartmann, S. (2008, September 21). *Models in Science*. Retrieved March 14, 2009, from The Stanford Encyclopedia of Philosophy: <http://plato.stanford.edu/archives/fall2008/entries/models-science>
- Gregg, L. W., & Simon, H. A. (1967). Process models and stochastic theories of simple concepts. *Journal of Mathematical Psychology* , 4, 246-276.
- Holland, J. H. (1998). *Emergence: from chaos to order*. Oxford: Oxford University Press.
- Hulshof, C., Eysink, T., & de Jong, T. (2006). The ZAP project: Designing interactive computer tools for learning psychology. *Innovations in Education and Teaching International* , 43 (4), 337-351.
- Johnson-Laird, P. (1983). *Mental Models*. Harvard: University Press.
- Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning*. Cambridge: MIT Press.
- Lawson, A. E., Clark, B., Meldrum, E. C., Falconer, K. A., Sequist, J. M., & Kwon, Y.-J. (2000). Development of scientific reasoning in college biology: do two levels of general hypothesis-testing skills exist? *Journal of Research in Science Teaching* , 37 (1), 81-101.

- Löhner, S., van Joolingen, W. R., Savelsbergh, E. R., & van Hout-Wolters, B. (2005). Students' reasoning during modeling in an inquiry learning environment. *Computers in Human Behavior* , 21, 441-461.
- Newell, A., & Simon, H. (1972). *Human Problem Solving*. Englewood Cliffs: Prentice Hall.
- Penner, E. P. (2000). Cognition, computers, and synthetic science: building knowledge and meaning through modeling. *Review of research in education* , 25, 1-35.
- Sins, P. H., Savelsbergh, E. R., & van Joolingen, W. R. (2005). The Difficult Process of Scientific Modelling: An analysis of novices' reasoning. *International Journal of Science Education* , 27 (14), 1695-1721.
- Sodian, B., & Bullock, M. (2008). Scientific reasoning—Where are we now? *Cognitive Development* , 23 (4), 431-434.
- Stratford, S. J., Krajcik, J., & Soloway, E. (1998). Secondary Students' Dynamic Modeling Processes: Analyzing, Reasoning About, Synthesizing, and Testing Models of Stream Ecosystems. *Journal of Science Education and Technology* , 7 (3), 215-234.
- Suppes, P., Pavel, M., & Falmagne, J. -C. (1994). Representations and models in psychology. *Annual Review of Psychology* , 517-44.
- Van den Broek, A. Tijs (2007). *How experts reason during modeling in an ill-structured domain*. Manuscript submitted for publication, Twente University, Enschede.
- Webb, M. (1993). Computer-based modelling in school science. *School Science Review* , 74, 33-46.
- White, B. Y., & Shimoda, T. A. (1999). Enabling students to construct theories of collaborative inquiry and reflective learning: computer support for metacognitive development. *International Journal of Artificial Intelligence in Education* , 10, 151-182.
- Zimmerman, C. (2000). The Development of Scientific Reasoning Skills. *Developmental Review* , 99-149.

Appendix A: Experiment material

Beste proefpersoon,

Modeleren is een aanpak om fenomenen (of theorieën, ideeën, gedachten) te externaliseren. Door de onderliggende variabelen uit te vinden, in samenhang met elkaar te brengen en te schetsen ontstaat een externe representatie, die de structuur en de werking van een fenomeen verheldert. Op die manier kan men hypothesen afleiden, modellen testen of gewoon begrijpen hoe iets in elkaar zit. Het is een methode om leren en begrip te ondersteunen en te verdiepen. Dit onderzoek wil de structuren van het modelleren proces uitvinden.

In het volgende zal je aan hand van een klein experiment en een theorie daarvan zelf een model maken, dat verbanden tussen variabelen verklaart. Het is de bedoeling dat je hierbij de hele tijd hardop gaat denken.

Voordat je met het modelleren echt gaat beginnen, krijg je twee kleine opgaven om het modelleren proces en het hardop denken te oefenen.

Opdracht 1: Puzzel

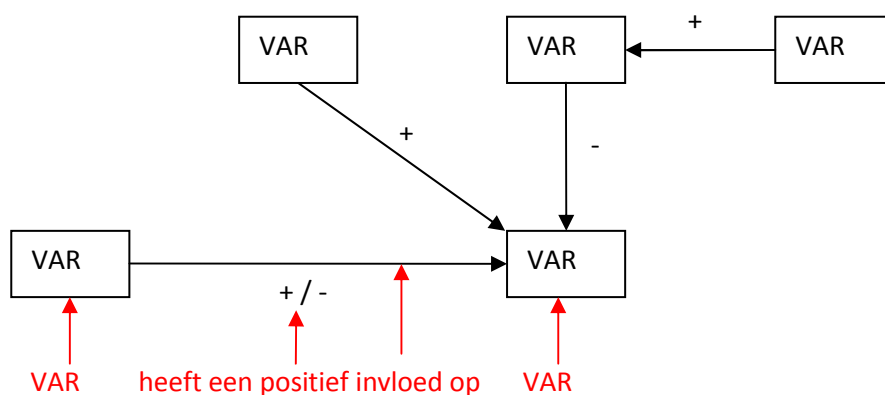
Maak uit de 4 puzzeldelen een vierkant. Probeer de hele tijd hardop te denken.

Opdracht 2: Een lekkende emmer water

Onder een lopende kraan staat een wateremmer, dus er stroomt water in de emmer. Onderaan, op de zijkant, heeft de emmer een gat waar ook weer water uitstroomt. Hoeveel water er door het gat uit de emmer stroomt, hangt af van twee dingen. Ten eerste: hoe groter het gat, hoe meer water er doorheen kan en ten tweede: hoe meer water er in de emmer zit, hoe groter de druk van de waterkolom, waardoor het water harder uit het gat zal stromen.

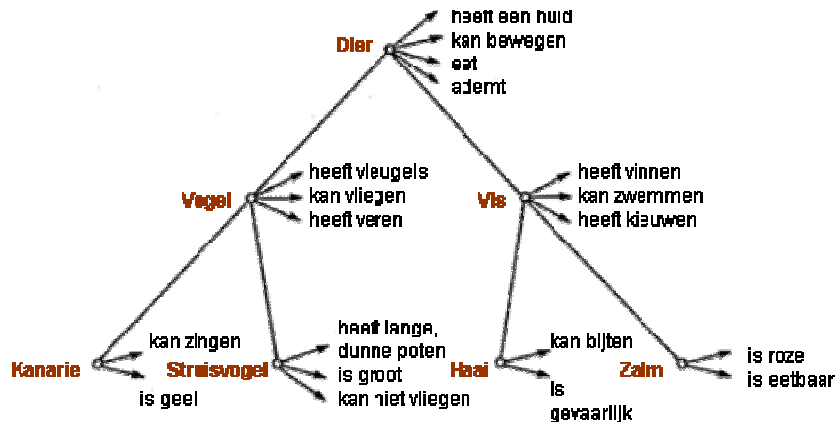
Modelleer de lekkende emmer water en probeer weer de hele tijd hardop te denken.

Voorbeeld van elementen in een model:



Theorie:

De basisaanname van het semantisch geheugen volgens Collins en Quillian is dat elk woord apart is opgeslaan met een verwijzing naar andere woorden. Daardoor ontstaat een netwerk met concepten en eigenschappen als onderdeel van deze concepten.



Invloedrijke variabelen

Abstractie

Het netwerk is hiërarchisch gestructureerd. Dit betekent dat alle concepten en eigenschappen die onder een ander, hoger gelegen concept staan, ermee verbonden zijn en tot eenzelfde klasse behoren. Als je het plaatje hierboven bekijkt dan zie je bijvoorbeeld dat de concepten "kanarie" en "struisvogel" behoren tot de klasse van de vogels en tot de klasse van dieren. Hoe hoger iets in de hiërarchie zit, hoe abstracter is het.

In het netwerk geldt het principe van overerving. Dit betekent dat eigenschappen maar één keer worden opgeslagen en wel zo hoog mogelijk in de hiërarchie. In het plaatje hierboven zie je bijvoorbeeld dat "heeft veren" niet opgeslagen wordt bij "kanarie" maar bij "vogel".

Deze hiërarchie (ofwel abstractie) is het kader voor twee bijkomende variabelen: semantische afstand en het aantal verbindingen tussen de concepten. Hoe hoger het niveau van de abstractie, hoe groter de afstand ofwel hoe groter het aantal verbindingen.

Semantische afstand / aantal verbindingen

Als bewijs voor het bestaan van dit netwerk worden reactietijden van proefpersonen gemeten op zinnen die semantische (feitelijke) relaties uitdrukken zoals "Een kanarie heeft een huid". Van groot belang hierbij is het begrip semantische afstand. De semantische afstand tussen twee woorden wordt bepaald door de verbindingen in het netwerk en wordt uitgedrukt in de kortste weg tussen deze twee woorden. De semantische afstand tussen bijvoorbeeld "kanarie" en "vogel" is 1, terwijl die tussen "kanarie" en "dier" 2 is. De semantische afstand in de zin "Een kanarie heeft een huid" is dus 2.

Uit onderzoek blijkt dat mensen bij correcte zinnen langer doen over hun beslissing naarmate de semantische afstand toeneemt. Beslissen of een kanarie een huid heeft, kost meer tijd dan beslissen of een kanarie geel is.

Bij incorrecte zinnen geldt net het omgekeerde: hier zullen reactietijden langer zijn naarmate de semantische afstand tussen twee woorden kleiner is. Zo is de reactie op de zin "Een kanarie is een struisvogel" trager dan de reactie op de bewering "Een kanarie is een vis". Woorden met kleine semantische afstanden zijn namelijk omwille van hun sterker verwante eigenschappen, moeilijker van elkaar te onderscheiden bij incorrecte beweringen. Er moet dus een groter aantal verbindingen worden afgezocht om te beslissen dat de stelling fout is.

Zintypes

Verder maakten Collins en Quillian onderscheid tussen twee zintypes. Zinnen die op een concept ("Superset"; S) en zinnen die op eigenschappen ("properties", P) verwezen. "Een kanarie is een vogel" is dus een S1-Sentence, want het heeft een semantische afstand van 1 en heeft betrekking op het concept vogel.

Deze onderscheiding is belangrijk, omdat er een verschil in de reactietijd ontstaat. S-zinnen werden steeds sneller beoordeeld, dan P-zinnen. De grafische weergave van de uitkomsten schetst een goed overzicht.

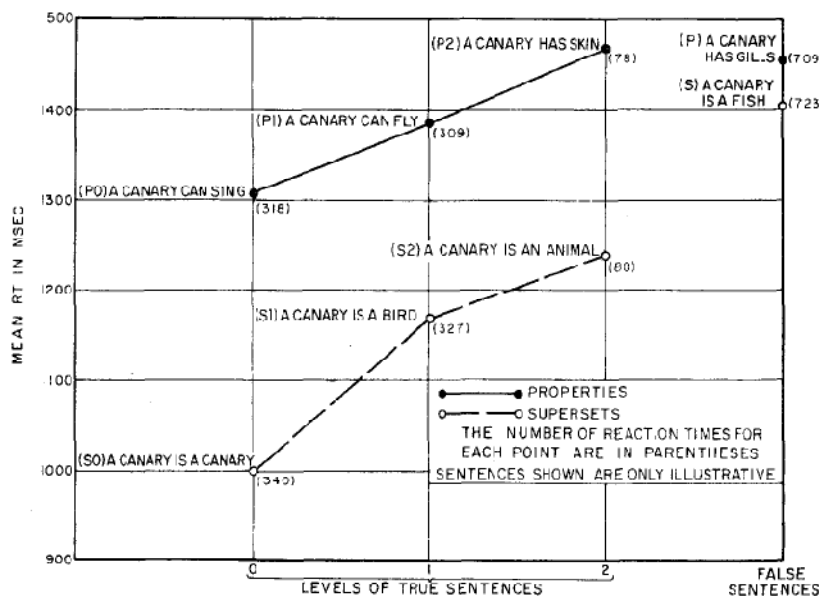


FIG. 2. Average reaction times for different types of sentences in three experiments.

Voor dit verschil konden zij echter geen duidelijke verklaring vinden. Een mogelijke idee was, dat de S-zinnen vaker in het dagelijkse leven voorkomen, waardoor de reactietijd werd verlaagd. Helemaal bevestigd werd het dus niet!

Pattern Matching

Een andere opzichtigheid is, dat de S0-zin sneller werd herkend, dan Collins en Quillian hadden verwacht. Zij gingen daarvan uit dat P en S-zinnen helemaal parallel verlopen. Door interviews met de proefpersonen concludeerden zij dat dit verschil op "Pattern Matching" terug te voeren is, omdat zij bij deze soort zinnen niet naar de betekenis moesten kijken.

"Pattern Matching" betekend dat bijv. de zin 'A canary is a canary' meteen als "goed" wordt herkend, want men moet niet naar de betekenis kijken, maar ziet meteen dat er twee keer 'canary' staat. Dit soort beoordeling gaat sneller, dan wanneer men naar de betekenis van de zinnen moet kijken.

Individuele eigenschappen

Het menselijke brein is niet bij iedereen gelijk opgebouwd of gestructureerd. Dat heeft ook invloed op andere, bovenstaande variabelen.

Opgave:

Maak een model, dat de reactietijd voor de stellingen kan verklaren. Maak het zo compleet als mogelijk en denk daarbij ook aan eventueel andere variabelen, die niet in de tekst staan of niet expliciet gekentekend zijn. De vorm van het model zou dezelfde zijn als bij de lekkende emmer water. Vergeet ook het hardop denken niet.

Succes!

Appendix B: Figures

Figure 1a

Progressing average of relative frequencies of reasoning processes of participant 1

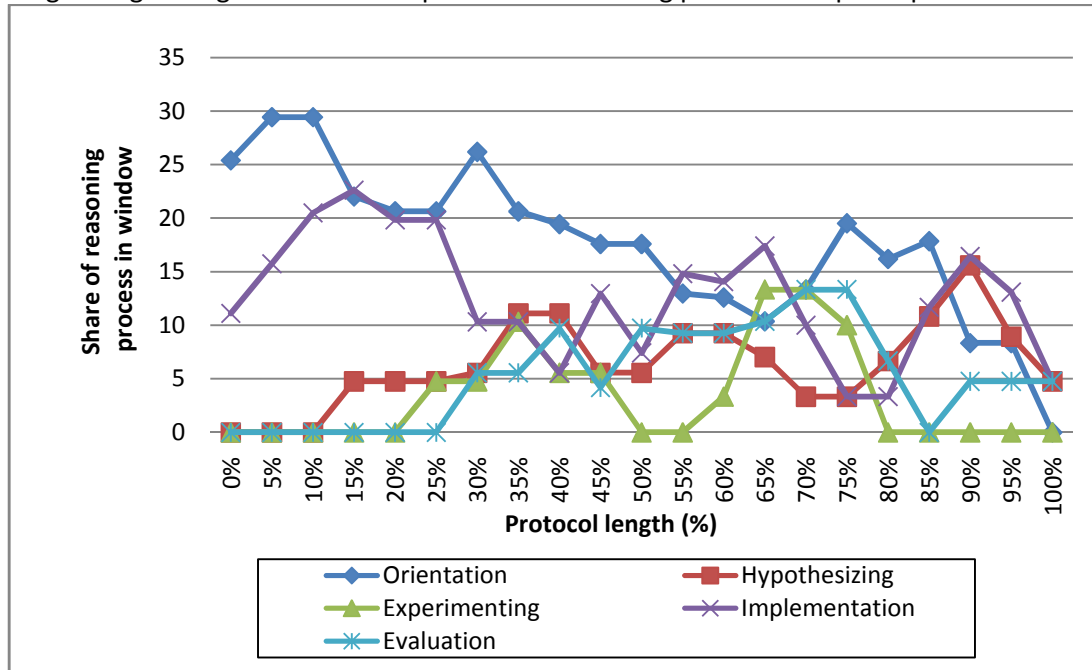


Figure 1b

Progressing average of relative frequencies of reasoning processes of participant 2

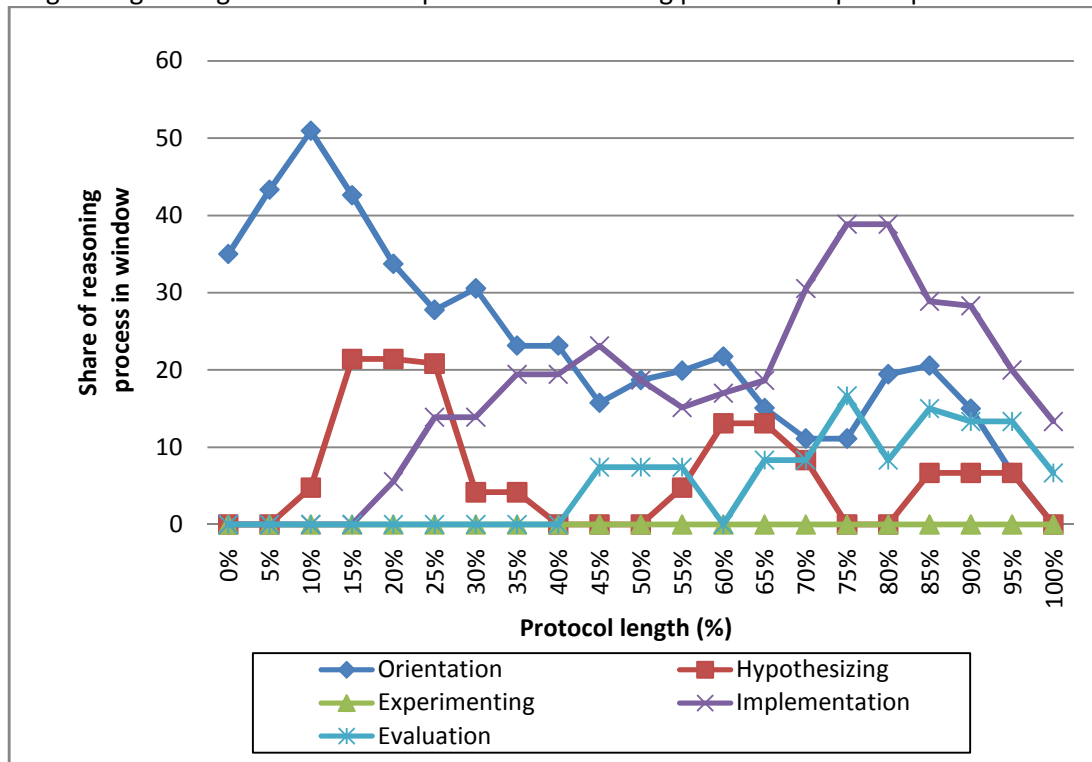


Figure 1c

Progressing average of relative frequencies of reasoning processes of participant 3

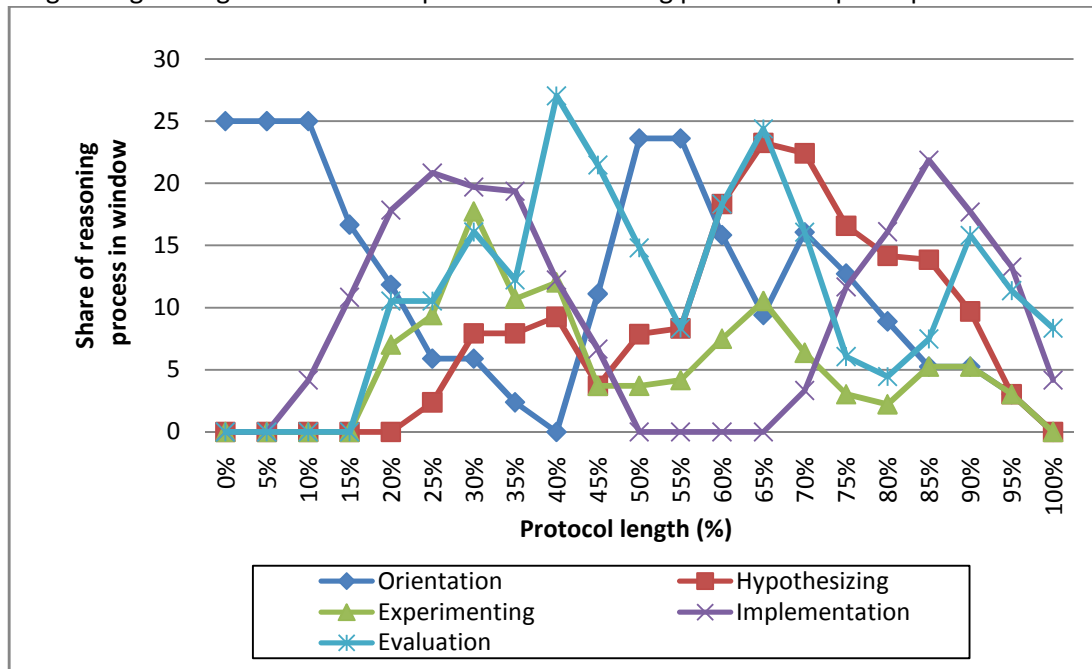


Figure 1d

Progressing average of relative frequencies of reasoning processes of participant 4

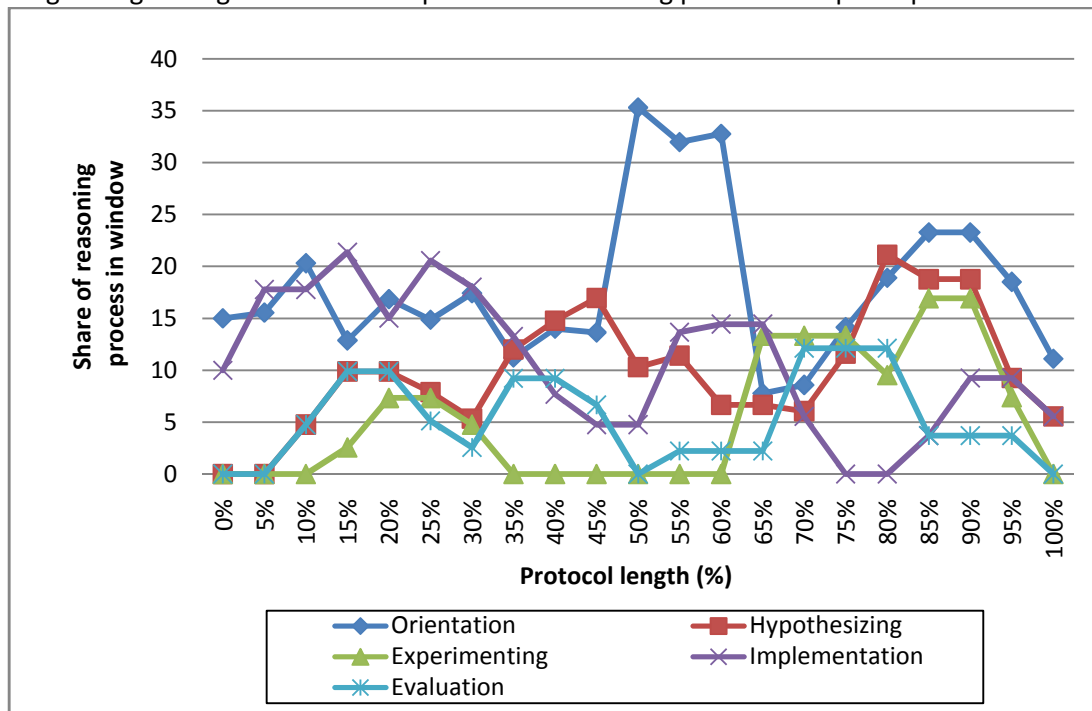


Figure 1e

Progressing average of relative frequencies of reasoning processes of participant 6

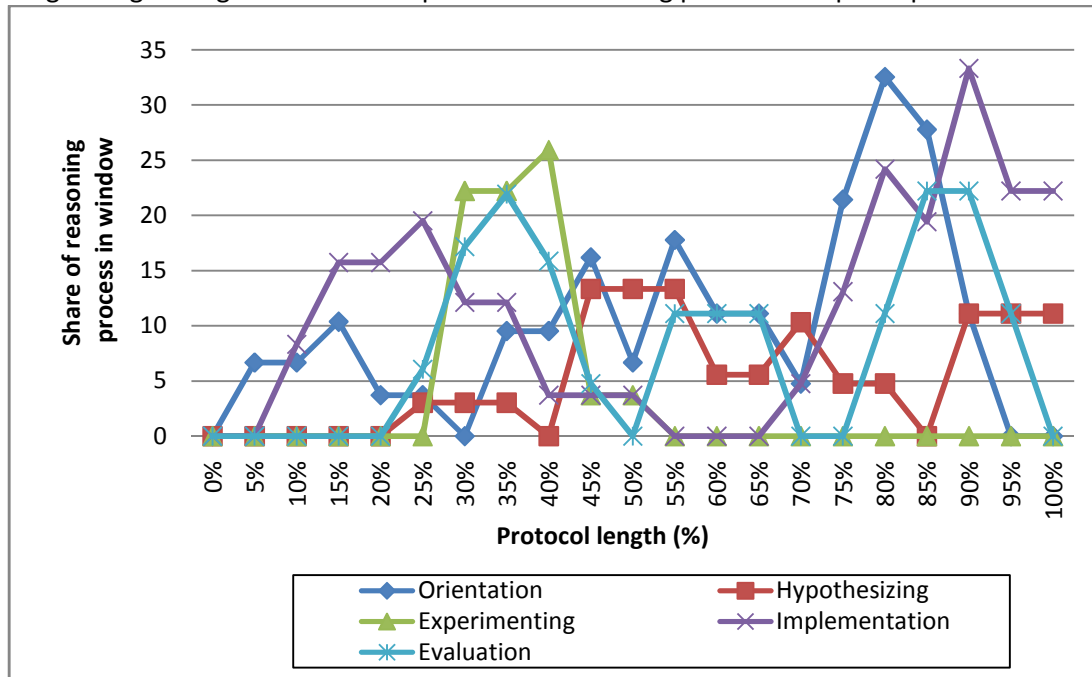


Figure 1f

Progressing average of relative frequencies of reasoning processes of participant 7

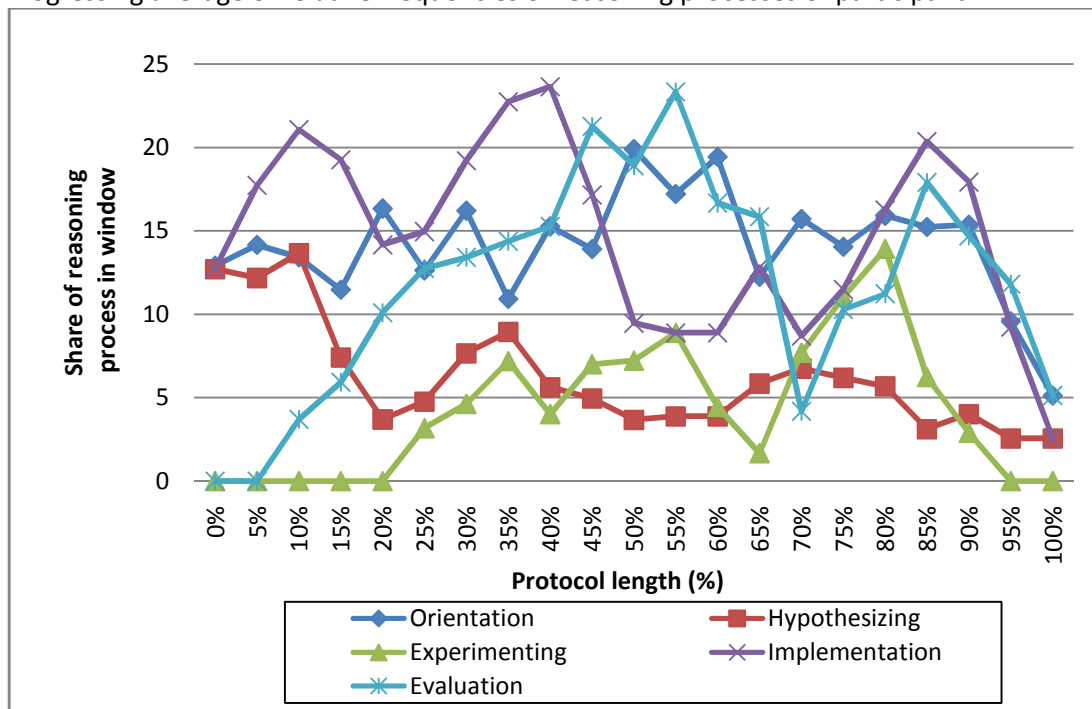


Figure 1g

Progressing average of relative frequencies of reasoning processes of participant 8

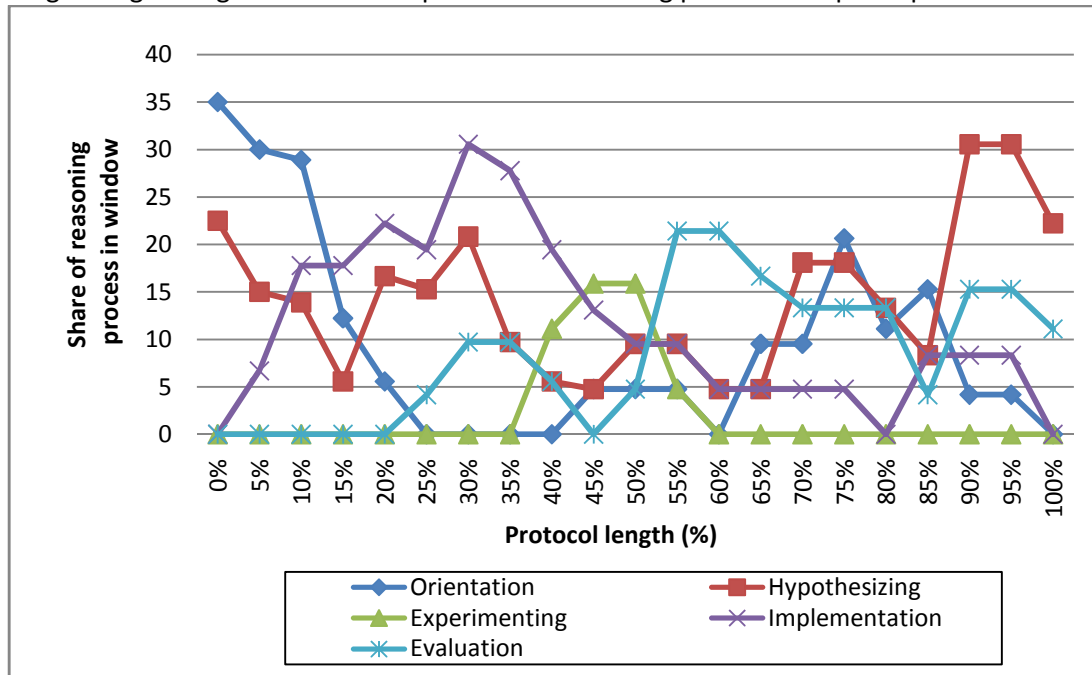


Figure 1h

Progressing average of relative frequencies of reasoning processes of participant 9

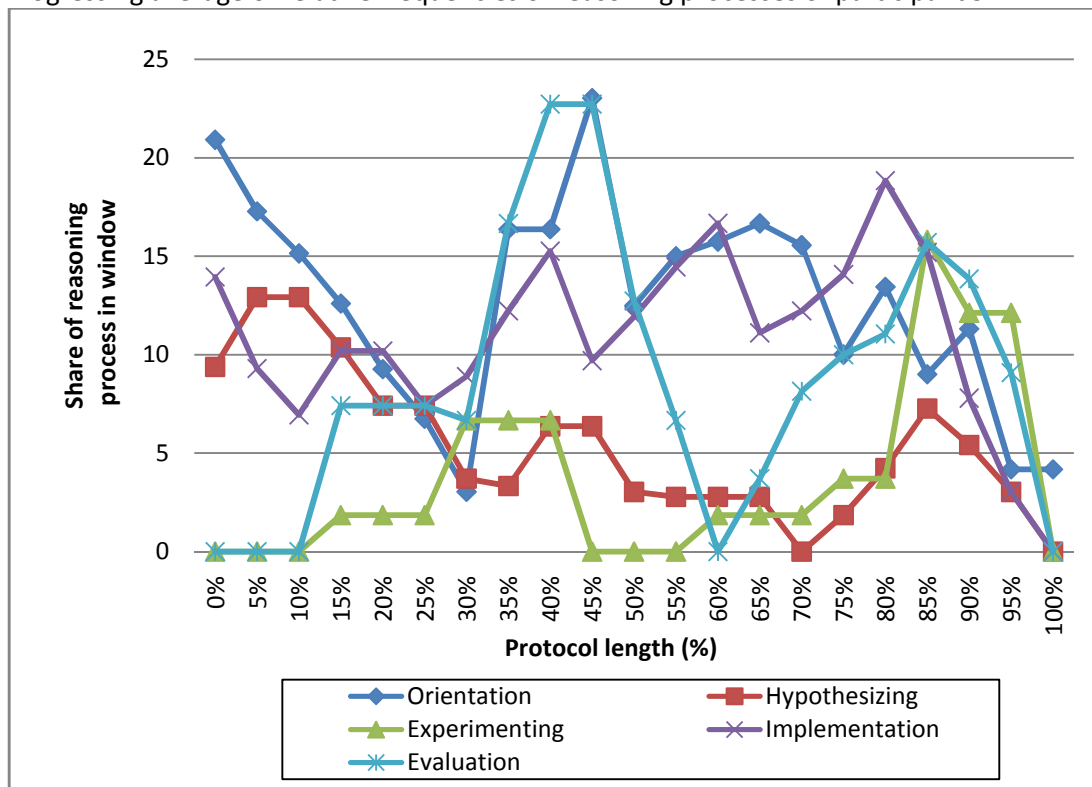


Figure 1i

Progressing average of relative frequencies of reasoning processes of participant 10

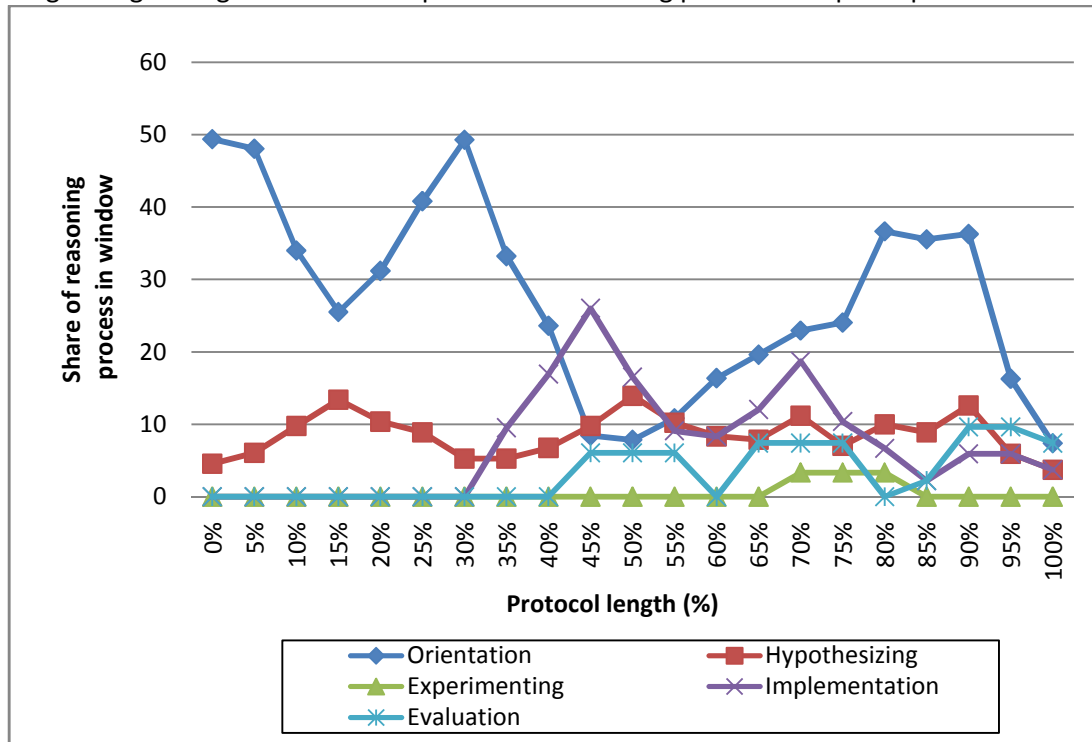


Figure 2

Progressing average of relative frequencies of reasoning processes of all participants

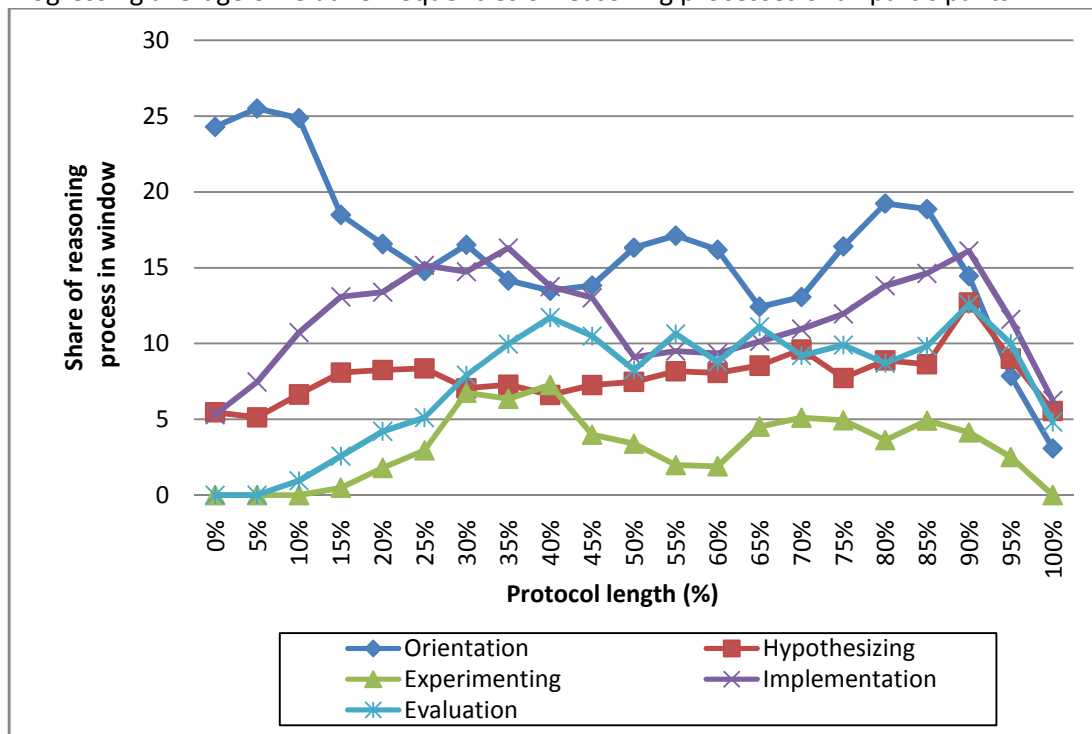


Figure 4c
Model of participant 1

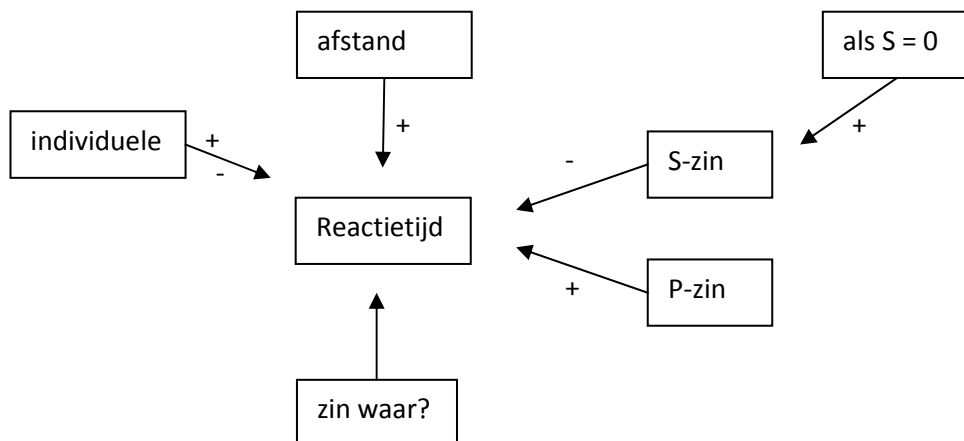


Figure 4d
Model of participant 4

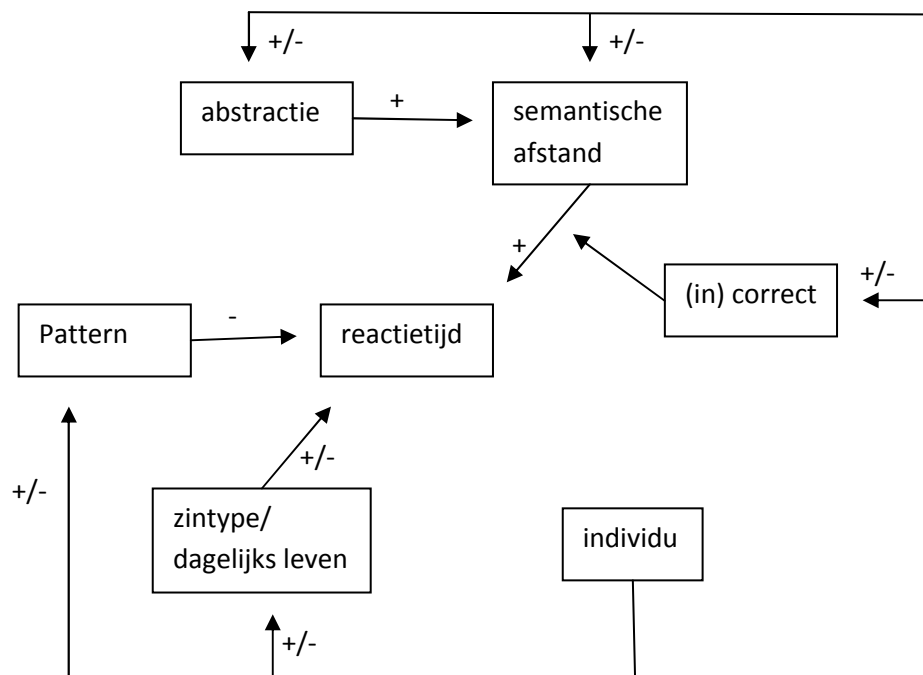


Figure 4e
Model of participant 6

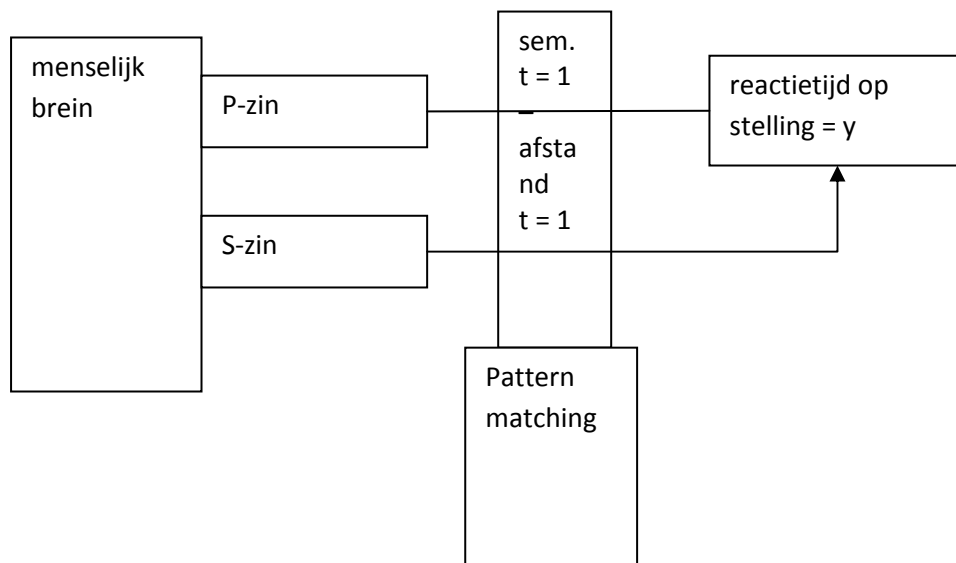


Figure 4f
Model of participant 7

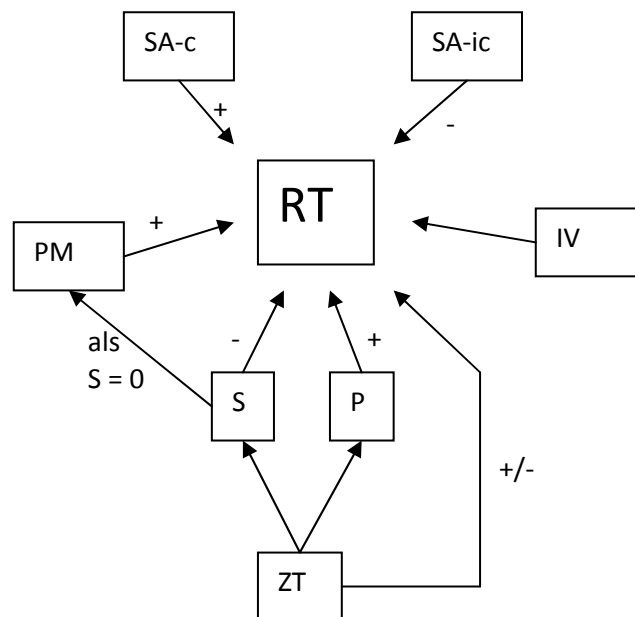


Figure 4g
Model of participant 8

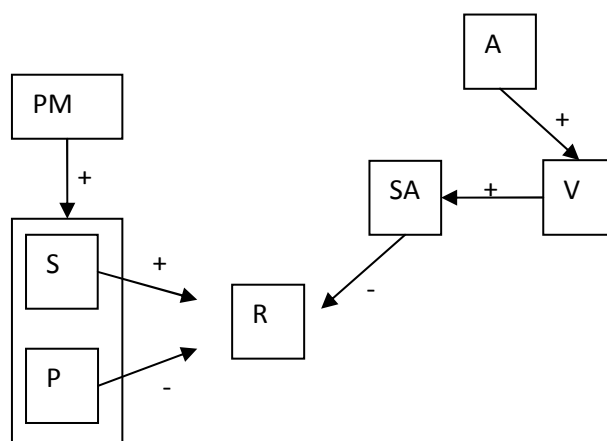


Figure 4h
Model of participant 9

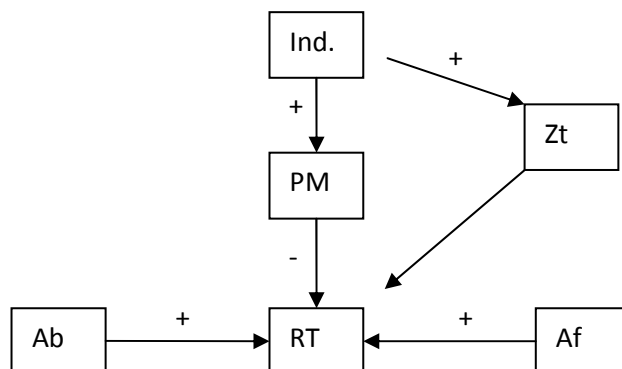
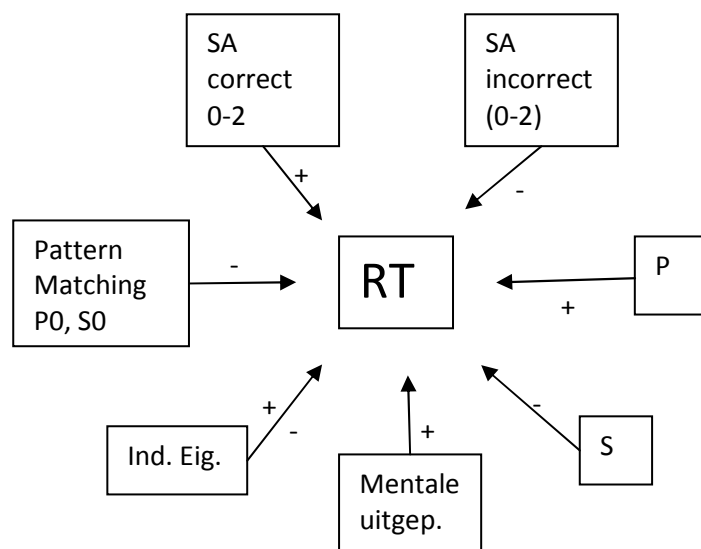


Figure 4i
Model of participant 10



Appendix C: Tables

Table 4

Percentage of time the strong group and weak group spent on the processes during the task (without the fifth model)

	strong group	weak group			
	M (SD)	M (SD)	U	p	
Orientation (%)	21.7 (6.1)	11.3 (1.8)	0.00	0.021	*
Hypothesizing (%)	7.0 (1.0)	8.8 (4.7)	8.00	1.000	
Experimenting (%)	2.1 (2.2)	4.0 (2.1)	5.00	0.386	
Model implementation (%)	14.3 (4.4)	12.6 (1.3)	5.00	0.386	
Model evaluation (%)	7.1 (3.9)	9.8 (1.9)	4.00	0.248	
Actions (%)	32.0 (7.1)	31.0 (5.5)	8.00	1.000	
Regulation (%)	10.1 (5.4)	11.4 (3.5)	6.00	0.564	
Off tasks (%)	2.7 (1.2)	3.5 (1.5)	4.00	0.248	
Experimenter (%)	2.7 (3.5)	7.3 (1.6)	2.00	0.114	

U and p scores were obtained with a non-parametric Mann-Whitney U-test

**p = 0.05*