Students' reasoning during scientific computer-based modeling : the impact of epistemology, motivation and communication mode

Sins, P.H.M.

STUDENTS' REASONING DURING COMPUTER-BASED SCIENTIFIC MODELING

The impact of epistemology, motivation and communication mode

UNIVERSITY OF AMSTERDAM
Graduate School of Teaching and Learning

The research reported in this dissertation was supported by the European Community under the Information Society Technology (IST) School of Tomorrow programme (contract IST-2000-25035).

ico

This research was carried out in the context of the Interuniversity Centre for Educational Research.

Cover design: Monique van Hootegem
Printed by PrintPartners Ipskamp B.V., Amsterdam.

ISBN-10: 90-78087-05-6
ISBN-13: 978-90-78087-05-2

STUDENTS' REASONING DURING SCIENTIFIC COMPUTER-BASED MODELING

The impact of epistemology, motivation and communication mode

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
Prof. mr. P.F. van der Heijden
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit

op donderdag 18 mei 2006, te 12.00 uur

door

Patrick Henricus Maria Sins

Geboren te Maastricht

Promotiecommissie

Promotor:	Prof. dr. B.H.A.M. van Hout-Wolters
Co-promotores:	Dr. W.R. van Joolingen Dr. E.R. Savelsbergh
Overige leden:	Prof. dr. A.J.M. de Jong Prof. dr. A.L. Ellermeijer Prof. dr. H.G. Schmidt Dr. J.H. van Driel Prof. dr. B.J. Wielinga
Faculteit:	Maatschappij-en Gedragswetenschappen
	Het promotie-onderzoek is voorbereid aan het Instituut voor de Lerarenopleiding van de Universiteit van Amsterdam

TABLE OF CONTENTS

VOORWOORD

Het was met andere woorden dan de volgende dat mijn familie- en vriendenkring in 2001 mij op extremistische wijze waarschuwde: ‘Kind, waarom ga je in Gods- en hemelsnaam naar het immens gevaarlijke Amsterdam? Onze voorvaderen hebben toch niet voor niets muren om ons Mestreech gebouwd!’.

Maar door een onverklaarbare wending in de loop der dingen (ik werd aangenomen als aio bij de Universiteit van Amsterdam) besloot ik het er toch maar op te wagen, om de stap over de grote rivieren heen te zetten, op weg naar hippieland. Beladen met: a) een zwaar doch immens schattig accent, b) een kortgeschoren blond geverfde haardos (wat heeft mij in die tijd toch bezield?) en c) een netjes gestreken klederdracht nam ik mijn intrek in Amsterdam-Noord; in een stulp waar een populatie kakkerlakken hoogtij vierde (vandaar de uitdrukking: ‘Als de Zuid-Limburger van huis is, dansen de kakkerlakken een kleine tango op de tafel’).

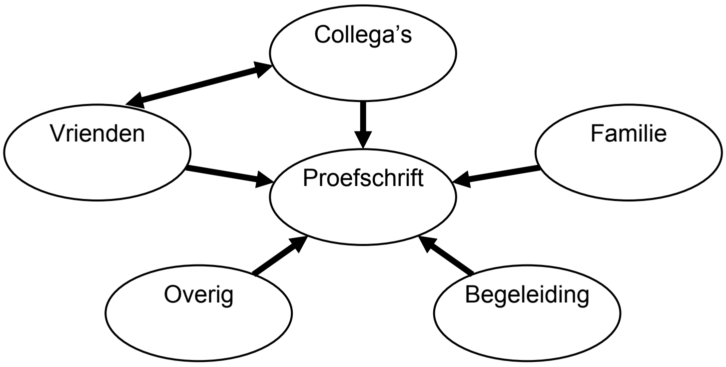
Op het Instituut voor de Lerarenopleiding waar ik mijn promotieonderzoek zou gaan uitvoeren, voerde ik in den beginne een titanenstrijd met mijn articulatievermogens, aangezien woorden als ‘Bijlmer’, ‘Wouter’, ‘herhalen’, ‘gewone kroketten’, respectievelijk als ‘buimach’, ‘wooteh’, ‘halen’, ‘wat zeg je me daar!??’ werden opgevat. Zelfs na het uitspreken van een alledaags woord als ‘stoel’ lagen mijn collega’s al in een deuk onder de spreekwoordelijke tafel. Ook de Mokumse straten troffen mij als een culturele mokerhamer. Zo heeft bijvoorbeeld iedereen een mening die hij of zij niet onder stoelen of banken steekt, is hier totaal geen carnaval, en nog schandelijker: in Amsterdam wordt de welbekende Vlaai waanzinnig aangeduid als: ’Hé, lekker taart!’.

Nee, ik wist niet hoe snel ik weer in het weekend naar die Heimat kon afreizen. Maar toch is er, in die bijna 5 jaar dat ik hier thans heb vertoefd, een ontwikkeling gaande geweest die ervoor heeft gezorgd dat: a) mijn haar in een oneindige groeifase is beland, b) mijn klederdracht niet meer netjes gestreken is, c) ik ook ineens een paar meningen heb (zie de stellingen), d) ik Maastricht zie vanuit het perspectief van een inheemse toerist en e) ik nog steeds een accent van hier tot Tokio heb.

Het heeft zelfs zover mogen komen dat ik mijn promotieonderzoek, waarvan u de verslaglegging nu in de knuisten heeft, succesvol heb kunnen afronden. Mijn onderzoek heb ik gelukkig niet in een vacuüm moeten uitvoeren, anders zou het figuurlijk maar ook zeker letterlijk niet mogelijk zijn geweest om op een fatsoenlijke wijze te kunnen functioneren. Verscheidene mensen hebben die vacuüm voor mij (on-) vrijwillig opgevuld. Zoals het een zichzelf relativerende onderzoeker betaamt heb ik een eenvoudig conceptueel model ontwikkeld dat een overzichtelijk en ook een leuk beeld geeft van al die mensen (zie Figuur 1). Ik zal in een semi-onwillekeurige volgorde elk element uit mijn model bespreken.

Leest u mee?

Goh, fijn zeg!



Figuur 1. Eenvoudig conceptueel model

1. FAMILIE

Aangezien het plantje uit het zaadje komt, begin ik met mijn familie. Mijn familieleden, zijnde: De opa's Josje en Paul, de oma's Gretha en Corry, Tante Wil, Greetje, Annemieke, Siemen, Klaas, Marleen, Benni, Yvonne, Harry, Denise, Anique, Jordy en Mattie, bleven altijd vol interesse luisteren als ik weer een onbegrijpelijke uitleg gaf over mijn onderzoek. Ook wil ik van de gelegenheid gebruik maken om Gerrit, Henny en Robert (en natuurlijk Floris en Lotje) te bedanken voor de gezellige visites die ik hen heb mogen brengen. Als laatste ben ik ontzaglijk veel dank verschuldigd aan mijn ouders Hugo en Harma en mijn broeder Mark voor alles wat ze voor mij hebben gedaan de afgelopen jaren en daarvoor. Jullie zijn echt fantastisch!

2. OVERIG

Geheel tegen de verwachting en het gebruikelijke in, bespreek ik het element 'overig' als tweede. Ten behoeve van mijn onderzoek heb ik bijna alle docenten natuurkunde 5 VWO in Amsterdam lastig gevallen met de vraag of ze hun leerlingen voor even wilden afstaan. Hierbij zou ik dan ook Bart Rijkenberg, Frans Koopman, Michiel Boonzaaier, Bert Bödicker, Erick Wormhout, Frans Hochstenbach, Jan Bos, Gerard ter Horst, Frans Eerkens, Piet Molenaar, Hans van Riet, Hans van Dijk, Peter Uylings en natuurlijk al hun leerlingen hartelijk willen bedanken voor hun deelname aan mijn onderzoek.

Ook hebben enkele andere personen hun waardevolle bijdrage geleverd aan dit proefschrift: Het Amstel instituut financierde voor de helft het laatste jaar van mijn aanstelling; Amber Walraven werkte als stagiaire mee aan onder andere de ontwikkeling van het analyseschema; José van Gelderen hielp mee tijdens enkele (instabiele) Co-Lab sessies en Iris Dicke typte haar vingers blauw om mij van de nodige transcripten te voorzien.

Tenslotte wil ik de personen die werkzaam waren binnen het Co-Lab project en die (in)direct hebben bijgedragen aan het welslagen van mijn onderzoek bedanken, maar dat doe ik voor de vorm even in het Engels.

I am grateful for the support I received from the following persons who contributed to the Co-Lab project: Anjo Anjewierden, Ton Ellermeijer, Martin Beugel, Thorsten Bell, Davide Biolghini, Ulrich Bosler, Alberto Ceccarelli, Miguel Celdran, Marijn van Eupen, Leendert van Gastel, Ernesto Martin, Toni Martínez Carreras, Eduardo Martinez Gracia, Laura Miani, Tina Miggiano, Manuel Mora, Maarten Pieters, Sascha Schanze, Jakob Sikken, Antonio Gomez Skarmeta, Roberto de Souza, Peter Uylings, Pascal Wilhelm, Sarah Manlove, Ard Lazonder, Ton de Jong and Thilo Wünschler.

3. BEGELEIDING

Zoals gezegd wist ik in het begin van mijn prille aio-bestaan van toeten noch blazen. Ik zou onderzoek gaan verrichten naar de wijze waarop leerlingen binnen het vak natuurkunde modellen maken en testen op de computer. Elwin en Wouter wisten mij enigszins wijs te maken binnen de complexe wereld van het modelleren en die van de natuurkunde. Maar gelukkig kon ik nog mijn veelal cognitief psychologisch ei kwijt bij Bernadette. Ik zal kort een ruwe typering geven van mijn drie begeleiders, zodat de lezers thuis weten hoezeer ik van mijn begeleidingsteam (oftewel 'de Bermuda driehoek') heb geleerd.

Een van de meest relativerende en problematiserende onderzoeker die ik ken is Elwin, ik noem hem dan ook de denker. Ik waardeer het enorm dat Elwin betamelijk veel denktijd in mijn onderzoek heeft willen steken en dat hij mij heeft geleerd dat als je iets zegt of doet, dat je dan altijd moet kunnen beargumenteren *waarom* je dat zegt of doet. Daarnaast heeft Elwin mij ook geholpen met het uitvoeren van de experimenten en van de vele data-analyses. Wouter (de pragmaat) toonde zich kundig in het doorvoeren van heldere revisies, om de veelal wollige papers die ik aanleverde nog strakker te maken. Daarnaast ben ik Wouter erkentelijk voor de intensieve ondersteuning tijdens het laatste gedeelte van mijn promotietraject. Tenslotte kan Bernadette (de psycholoog) zeker niet onvermeld blijven. Zij was degene met het spreekwoordelijk luisterend oor maar ook degene die ten gunste van de makke aio (lees Patrick) inhoudelijke knopen kon doorhakken, bedankt daarvoor. Zoals bij elke puzzel is 'het geheel meer dan de som der delen' en dat geldt dus ook voor mijn begeleidingsteam.

4. COLLEGA'S

Het element 'collega's' bestaat uit een zéér grote populatie en vormt als zodanig het sociale vangnet van de promovendus. Ten eerste heb ik op allerlei nationale en internationale congressen als de ORD, JURE, EARLI, AERA en natuurlijk de VPO dagen verscheidene interessante medepromovendi (ik noem hier geen namen anders wordt dit bescheiden nog langer en dat kan toch echt niet) mogen ontmoeten waarmee ik vele fijne koffie-en drankorgeluren heb kunnen doorbrengen. Ten tweede ben ik al de leden van de HogCog zéér erkentelijk voor de immer verfrissende en kritische discussies die tijdens de bijeenkomsten op de donderdagmorgen werden gevoerd.

Tenslotte wil ik stilstaan bij de werkpaarden van het ILO, waarvan de meesten inmiddels ook onder het element ‘vrienden’ vallen, vandaar de pijl tussen ‘collega’s’ en ‘vrienden’ in Figuur 1. Want is het niet zo, dat je veelal meer tijd spendeert met directe collega’s dan met de meeste vrienden? En het zou toch van de zotte zijn, zouden deze collega’s de uitdrukking: ‘Een goede buurman is beter dan een verre vriend’ niet waar kunnen maken? Nou, ik kan zeggen dat ik met mijn directe ILO collega’s enorm in mijn sas was en nog steeds ben. Ik neem een kleine greep uit de ‘Goh, heb je dat met je collega’s gedaan?’ doos: a) duizenden kilos aan lunch hebben we samen verorberd, b) miljoenen luchtmoleculen hebben we met elkaar uitgewisseld tijdens de meestal genoeglijke gesprekken, c) duizenden liters alcoholische sappen hebben we inmiddels genuttigd, d) kilometers dansvloer hebben we onveilig gemaakt, enzovoorts enzovoorts. Als dankzegging zal ik jullie allemaal bij naam (in een semi-onwillekeurige volgorde) noemen: Hein en Nadira (altijd leuk om bij jullie binnen te wippen), Stan, Simone, Monique, Marleen (je was een hele leuke kamergenote), Annoesjka (ik hoop dat ik van je kan blijven winnen met squash), Anne (je tolerantie jegens mijn goede eigenschappen is ongekend), Lap (thank you for being such a nice roommate), Irma, Mariëlle (jij helpt altijd, dankjewel), Martine, Joost, Annemieke, Louk, Carla, Jannet, Wilfried, Rijkje en Lenie (jullie plezierige aanwezigheid bij de lunch is 100%), Lizan, Hanneke, Gert, Jannis, Wiel, Geert, Saskia, Anne-Martine, Jeroen en Marcelle, Marianne, Marcel, Lucianne, Nico, Inez, Glenn, Bert, Bas, Petra, Pauline, Rob en Anita. Als laatste wil ik mijn huidige en sympathieke kamergenoten Hélène en Jaap ontzettend bedanken voor het feit dat ze het toch inhumaan lang met me hebben kunnen volhouden.

5. VRIENDEN

‘Geen straat is te lang met een vriend naast je’ zeg ik altijd als ik onder de douche lig en zo is het! Ik heb in Amsterdam en daarbuiten een hoop vrienden kunnen maken en ik hoop dat ik nog enkele jaren met ze mag slijten. Hoewel er hoogstwaarschijnlijk een direct negatief verband bestaat tussen de hoeveelheid tijd doorgebracht met vrienden en de tijd die besteed is aan het schrijven van dit proefschrift, hebben de volgende vrienden zeker indirect bijgedragen aan de totstandkoming van dit bescheiden: Helena, Iris, Vanessa, Elodie, Markéta, Rik, Merel-Naomi, Merel, Imre, Vincent, Patrick, Erik, Hester, Veronique, Luc, Stephan, Sjoertje, Sofie (2X), Maaike, Rosemarie, Marie-José, Kelly, Muriel, de mensen van theatersportgroep Placebo en de mensen van de verschillende toneelgroepen.

Annemarie, ik wil jou als laatste bedanken voor de tijd die je met me hebt durven doorbrengen en ook voor je rechtschapen twijfels betreffende onderwijskundig onderzoek in het algemeen en het mijne in het bijzonder.
Iedereen bedankt!!

Patrick Sins

Amsterdam, 2006

CHAPTER 1

INTRODUCTION

Many phenomena of scientific interest, such as climate change, population dynamics, or mechanical oscillations, are of a complex and dynamic nature. Complex in the sense that (non-)linear interactions between multiple variables explain the behavior of the phenomenon and dynamic because these variables change over time. Traditionally, in secondary science education these phenomena are addressed only in strongly simplified form because of the advanced mathematics needed to compute the behavior of these systems. Consequently, students attain little insight in the behavior of such phenomena. Computer models overcome these problems, since the computer takes over the tedious task of solving differential equations, allowing learners to experiment with and to visualize the behavior of the phenomenon being modeled (Niedderer, Schecker, & Bethge, 1991; Schecker, 1993; Whitfield, 1988).

Several claims have been put forward with regard to the educational value of computer-based modeling (Milrad, Spector, & Davidsen, 2002; Penner, 2001; Raghavan, Satoris, & Glaser, 1998; Schecker, 1993; Stratford, 1997). First, some authors emphasize the opportunity for students to think scientifically about the behavior of these phenomena (e.g., Bliss, 1994; Hestenes, 1997; Justi & Gilbert, 2002; 2003; Schwarz & White, 2005; Stratford, Krajcik, & Soloway, 1998; Wild, 1996). They argue that creating computer models involves various activities (e.g., identifying variables, considering and implementing connections among variables and verifying models by comparing their behavior to experimental data) that provide students with opportunities to reason about phenomena in a way that resembles the practice of scientists. Engaging students in these modeling activities enables them to understand and experience the issues associated with the construction of models in science. For instance, students may come to appreciate the tentative nature of scientific models, in that models are subject to continual revision and evaluation (e.g., Gilbert & Rutherford, 1998). In addition, students may acknowledge that models are built for a specific scientific purpose, such as to predict how a phenomenon will behave under certain circumstances. Acquiring such a comprehensive understanding of scientific models may be important in advancing students' thinking during modeling (Gobert & Discenna, 1997; Schwarz, 2002; Schwarz & White, 2005).

Second, computer models are tools that allow students to externalize their understanding of a particular phenomenon. Some authors claim that this affordance of computer models supports and advances students' reasoning, because it enables students to elaborate and reflect upon their mental model of the phenomenon (e.g., Coon, 1988; Fretz, Wu, Zhang, Davis, Krajcik, & Soloway, 2001; Gilbert, Boulter, & Rutherford, 1998; Jackson et al., 1996; Jonassen, Strobel, & Gottdenker, 2005; Ogborn & Wong, 1984; Penner, 2001). In addition, the fact that computer models can be simulated provides students with the opportunity to reason about cause-effect relations among variables and about how these relations interact over time (Hogan & Thomas, 2001). Furthermore, computer models enable students to test their own mental models by comparing the model output (e.g., graphs or tables) with empirical data.

Finally, computer models can serve to make ideas accessible to criticism from peers (Devi, Tiberghien, Baker & Brna, 1996; Rouwette, Vennix, & Thijssen, 2000; Suthers, 1999). The external nature of models allows them to serve as artifacts to reason with, not only on the individual level but also on the group level. According to this argument, models promote collaboration between a group of students, which may lead to a more elaborate understanding of the phenomenon compared to individual modeling. Also, constructing models collaboratively highlights the socially mediated side of science.

Despite all high expectations and optimistic claims, for secondary students, the process of modeling poses a highly demanding learning task, and the expected benefits are unlikely to occur unless all circumstances, such as the assignment(s), the modeling formalism, and the computer software, will be tailored towards the students needs. In order to identify these needs and to eventually provide suggestions for support, a thorough exploration of students' reasoning during modeling is necessary. Moreover, the degree to which computer-based modeling leads to deep reasoning and better achievement may be dependent on: a) characteristics of the students, such as their motivation (i.e., achievement goal orientation and self-efficacy) and their epistemological understanding of models and on b) characteristics of the modeling environment, such as the mode students use to communicate (i.e., chat versus face-to-face communication).

In this dissertation we focus on the collaborative modeling of students who, working in dyads, construct and revise models of dynamic scientific phenomena with the help of a computer. We will investigate the nature of students' reasoning during modeling and we will examine whether and how attributes and beliefs of the students and of the collaborative learning environment influences students' reasoning.

1. THE COMPLEX NATURE OF COMPUTER MODELING

Studies have reported that students encounter a number of typical difficulties in performing a computer-based modeling task. First, students have problems in comprehending the complex and dynamic behavior of phenomena that are modeled (e.g., Hogan & Thomas, 2001; Kurtz dos Santos & Ogborn, 1994; Riley, 1990; Sweeney & Sterman, 2000). They have difficulties in addressing variables that change over time and in thinking about complex interactions between variables. For instance, students have trouble understanding feedback mechanisms (e.g., the amount of interest on a bank account leads to an increase of the main sum, which in turn leads to a higher interest next year) which involves complex reasoning, and resort to linear causal thinking (i.e., trace one cause to one effect) instead. Second, students have difficulties in applying the representational formalism employed by some computer modeling tools, such as STELLA (Steed, 1994) and Powersim (Byrkness & Myrtveit, 1997). For instance, many students struggle in translating their mental model into a computer model using the symbols (e.g., graphical icons) employed by the modeling tool to select and to define factors, especially in ways that make them adaptable to quantification (e.g., Bliss, 1994; Kurtz dos Santos & Ogborn, 1994; Tinker, 1993). Finally, establishing the connection between model output and experimental data is an important step in evaluating and in revising a model, but this step is not easily taken by students (Doerr, 1996; Steed, 1994; Whitfield, 1988). Consequently, students may employ an engineering approach to modeling in which they attempt to fit the output of their model with the data (i.e., model fitting behavior), without reasoning about how the model yielded that particular output (Ogborn, 1999).

In order to frame the difficulties students encounter and to provide support, it is important to investigate how they reason *during* modeling. Therefore, in Chapter 2, we analyzed the reasoning processes of students engaged in a modeling task. This leads to an analysis scheme that measures the *occurrences* as well as the *quality* of the reasoning processes students employ during modeling, based on their verbal interactions (cf. Hogan & Thomas, 2001). Among other things this analysis scheme is capable of distinguishing deep versus surface reasoning in the modeling process by the criterion that deep reasoning processes involve integration with existing knowledge.

This online measure of students' reasoning is also employed in the studies reported in Chapters 3 through 5 of this dissertation. Through these investigations we hope to learn more about the nature of students' reasoning during modeling and about which variables contribute to successful modeling.

2. EPISTEMOLOGICAL UNDERSTANDING AND STUDENTS' REASONING

Models are powerful tools which enable scientists to generate predictions, as well as guide explanation, interpretation, understanding and discovery in science (Crawford & Cullin, 2004; Giere, 1990; Hestenes, 1997). Although, there is no single consensus scientific epistemology that scientists agree upon, an expert conception of the role and/or purpose of models in science can be summarized as follows:

> '1. Models exist as aids to understanding phenomena and this understanding can be checked or verified by comparing the results obtained by manipulating the model to observations obtained in the real world
> 2. A primary guideline for making a model is to consider its purpose
> 3. A scientist can have more than one model for the same thing because different models can be used to address different specific interests or questions about the referent
> 4. A scientific model can change and be replaced by one that is better for answering questions' (Crawford & Cullin, 2004; 1382-1383).

Most students hold conceptions that correspond with a 'naïve realist' epistemology of models and modeling (e.g., Barowy & Roberts, 1999; Gobert & Discenna, 1997; Talsma, 2000; Treagust, Chittleborough, & Mamiala, 2002; Van Driel & Verloop, 1999). This means that students mostly regard scientific models as concrete replicas of the real object, albeit on a different scale. More advanced students, however, appear to understand that models are designed for specific purposes, and that models may be revised when, for example, new empirical data has been analyzed (e.g., Schwarz & White, 2005; Spitulnik, Krajcik, & Soloway, 1999). Almost no student reports that scientists use models to test their ideas or that the function of models is explaining and/ or predicting the behavior of scientific phenomena.

Understanding the nature and purpose of models as well as how models are constructed may help students use and build models. Moreover, students' epistemological understanding about models and modeling may have an effect on how they approach the material and ultimately on what they learn (Hammer, 1994; Schwarz & White, 2005; Van Driel & Verloop, 1999). Therefore, it seems plausible that students' epistemological understanding influences the reasoning processes they employ during modeling (e.g., Hofer, 2001; Hofer & Pintrich, 1997; Kardash & Howell, 2000; Windschnitl & André, 1998). Although there is ample evidence to suggest that students have a poor epistemological understanding of models and modeling, there has been little evidence directly linking students' epistemology to the processes they employ during modeling. If such a relation exists, students should be scaffolded to develop a more sophisticated epistemological understanding of models and modeling. In the study reported in Chapter 3, we focus on examining the relation between students' epistemological understanding of models and the reasoning processes they employ during modeling.

3. COMMUNICATION MODE AND STUDENTS' REASONING

In the studies reported in this dissertation, students constructed and revised models in dyads. As noted above, models can play a role in the communication between students, as they are an external artifact that can be both means and subject of communication. Communication takes place through these common artifacts as well as other means of communication provided by the collaborative learning environment, such as text-based chat. Although the availability of chat in such environments makes collaboration between dispersed groups possible, when compared to direct face to face communication, the set of modalities by which learners can communicate is reduced. This raises the question of what the impact is of chat on students' reasoning during modeling compared to face-to-face interaction.

Chat imposes constraints on communication, in that written text relies on fewer channels than face-to-face interaction for transmission of a message (Baltes, Dickson, Sherman, Bauer, & LaGanke, 2002; Doerry, 1996; McGrath & Hollingshead, 1993; Olson & Olson, 1997). In addition, this medium requires greater effort to compose and send messages compared to face-to-face interaction. It is also impossible for students to respond while messages are being written. These communication constraints of chat may hinder students during modeling. Alternatively, dyads communicating through a chat tool may be pressured to increase the efficiency of their communications by compressing their communication (Condon & Cech, 1996a; 1996b; Jonassen & Kwon, 2001). In addition, the affordance that sent chat messages stay available on students' screens may also advance students' reasoning.

Unfortunately, findings from studies investigating the influence of chat versus face-to-face communication on students' performance are rather contradictory. The purpose of the study reported in Chapter 4 is to investigate the effects of communication mode (i.e., chat versus face-to-face communication) on : a) students' activities during modeling, b) students reasoning during modeling, and c) the quality of the models students construct.

4. MOTIVATION AND STUDENTS' REASONING

According to current models of self-regulated learning (e.g., Covington, 2000; Pintrich, 2003; Schunk, 2005; Zimmerman & Schunk, 2001), motivational factors are likely to influence students' reasoning during learning tasks. Constructs which have received the most attention within the field of motivation research are: a) type of students' achievement goal orientation, and b) students' self-efficacy. Students' achievement goal orientation involves their beliefs regarding learning success and their motives to engage in learning activities, such as modeling. Two distinct types of achievement goal orientations are traditionally distinguished (e.g., Jagacinski, 1992; Nicholls, 1989): *mastery goal orientation* and *performance goal orientation*. These goals differ primarily in terms of whether learning is perceived and valued as an end in itself or as a means to a goal external to the task. Mastery goal orientation involves the belief that effort leads to improvement in performance and that competence is malleable. Students who are mastery oriented, focus on the development of new skills and knowledge, try to elaborate on the learning task and try to reach their own learning goals. Students with a performance goal orientation, in contrast, believe that competence can be demonstrated by performing better than peers and tend to focus on attaining normative learning goals (e.g., Butler, 1991; Covington, 2000; Dweck & Leggett, 1988; Elliot & Dweck, 1988; Meece & Holt, 1993; Nicholls, 1989).

The type of achievement goal orientation may influence the nature of students' reasoning during modeling. For instance, students who are mastery goal oriented attempt to gain a rich insight in the given learning material and may therefore employ more deep reasoning processes to increase their comprehension. Performance goal oriented students, in contrast, try to obtain good results without having to invest too much cognitive effort. As a result, these students may mainly employ surface reasoning processes during modeling.

Another motivational variable that is considered to have an effect on students' reasoning is their self-efficacy. Self-efficacy has been defined as the learner's belief regarding their performance capabilities in a particular domain (Bandura, 1982; 1986). If students are high on self-efficacy, it can be expected that they aim to elaborate more on the learning material compared to students who are lower on self-efficacy (Pintrich & Schrauben, 1992). Therefore, self-efficacy may have a positive influence on students employing deep reasoning processes during a modeling task.

Goal of the study reported in Chapter 5 is to investigate the relations among achievement goal orientation, self-efficacy, reasoning and model quality of students who are working in dyads on a computer-based modeling task. In contrast to most studies in the field of motivation, we did not assess students' reasoning based on self-report measures, but rather we base our assessments on the process observations as described in the previous chapters. This is first because, the validity of self-report on these issues has become under doubt (Hogan & Thomas, 2001; Pintrich, 2000a), and furthermore because in the present context, where it comes to interpreting the detailed reasoning processes of students in a collaborative setting, more fine-grained measures are needed and the interaction between the collaborating partners must be taken into account.

5. OVERVIEW OF THE DISSERTATION

Main purpose of this dissertation is to investigate whether and how students' reasoning during computer-based modeling is affected by: a) their epistemological understanding of models, b) the mode students use to communicate and c) their motivation. Participants in all studies were students from eleventh-grade pre-university education, with a major in science. They had no prior experience with computer models and their age ranged between 16-18 years. In Chapters 2 and 3 students worked on a model of the distance covered by an ice-skater using the systems-oriented modeling tool Powersim (Byrkness & Myrtveit, 1997). In Chapters 4 and 5 students' worked on a model of the temperature of an irradiated black sphere using a similar modeling tool implemented in the Co-Lab environment (Van Joolingen, et al., 2005).

In the chapters in this dissertation, we will address the following four main research questions:

Which reasoning processes, employed by students during a computer-based modeling task, need support?

Aim of the study reported in Chapter 2, is to assess the specific reasoning processes students employ during computer-based modeling. In order to identify the features that are needed to describe students' reasoning processes, an analysis scheme is developed. This scheme is used in scoring the protocols of interacting dyads to identify the reasoning processes students' typically employ during modeling and to determine what processes are difficult for them to perform.

What is the relation between students' epistemological understanding of models and modeling and their reasoning during a computer-based modeling task?

The relation between students' epistemological understanding and the reasoning processes they employ during modeling is examined in Chapter 3. On the basis of the analysis scheme that is described in Chapter 2, we distinguish between deep versus surface reasoning in the modeling process by the criterion that deep reasoning processes involve integration with existing knowledge. Correlations are used to examine the relation between students' epistemological understanding and the level of their reasoning during modeling.

What is the effect of chat communication versus face-to-face communication on students' modeling activities, on their reasoning during modeling, and on the quality of their model?

Goal of the study reported in Chapter 4 is to investigate the differential effect of face-to-face communication versus chat communication on: a) students' modeling activities, b) students' reasoning during modeling, and c) the quality of the students' model. The two conditions are compared on these variables.

What is the relation between students' motivation and their reasoning during a computer-based modeling task and is the relation between students' motivation and model quality mediated by their reasoning?

Chapter 5 is concerned with investigating whether students' achievement goal orientation and self-efficacy (i.e., students' motivation) influences their reasoning during modeling and whether students' reasoning mediates the relation between motivation and model quality. A series of multiple hierarchical regression analyses are performed to analyze the relations between these variables.

Chapter 6 presents and discusses the main results obtained from the studies reported in Chapters 2 through 5[1]. In addition, this chapter provides some limitations of the methodologies we had chosen for our studies. Finally, implications for educational practice are suggested.

[1] *The studies reported in these chapters have either been accepted for publication or have been submitted to international journals. Therefore, some overlap in texts between chapters may occur.*

CHAPTER 2

THE DIFFICULT PROCESS OF SCIENTIFIC MODELING: AN ANALYSIS OF STUDENTS' REASONING DURING COMPUTER-BASED MODELING*

Although computer modeling is widely advocated as a way to offer students a deeper understanding of complex phenomena, the process of modeling is rather complex itself and needs scaffolding. In order to offer adequate support, we need a thorough understanding of the reasoning processes students employ and of difficulties they encounter during a modeling task. Therefore, in this study twenty-six students, working in dyads, were observed while working on a modeling task in the domain of physics. A coding scheme was developed in order to capture the types of reasoning processes used by students. Results indicate that most students had a strong focus on adjusting model parameters to fit the empirical data with little reference to prior knowledge. The successful students differed from the less successful students in using more prior knowledge and in showing more inductive reasoning. These observations lead to suggestions for the design of appropriate scaffolds.

1. INTRODUCTION

1.1 Educational value of modeling

The educational value of computer-based dynamic modeling has been advocated by many authors. Some emphasize the importance of the model as an artifact that allows explicit visual representation of complex relations (e.g., Mandinach, 1988; Schecker, 1993; Bliss, 1994; Steed, 1994; Hogan & Thomas, 2001). Others put more stress on the activity of constructing a model as a meaningful learning experience (e.g., Stratford, 1997; Spector, 2000; Penner, 2001; Milrad, Spector, & Davidsen, 2002). This activity offers the opportunity for students to think scientifically about the behavior of complex phenomena (e.g., Bliss, 1994; Jackson et al., 1996; Wild, 1996; Hestenes, 1997); to reflect upon their own understanding (e.g., Schecker, 1993; Gilbert et al., 1998; Jonassen,

---

* Sins, P.H.M., Savelsbergh, E.R., & Joolingen, W.R. van (2005). The difficult process of scientific modeling: An analysis of novices' reasoning during computer-based modeling. *International Journal of Science Education, 14*(8), 1695-1721.

et al., 2005; Raghaven et al., 1998;); and to test their mental models (Coon, 1988; Doyle & Ford, 1998; Penner, 2001). Crucial in this process is that modeling tools help students to externalize their ideas, so that they are open to criticism and discussion (Devi et al., 1996; Suthers, 1999; Rouwette et al., 2000). However, constructing a dynamic model is a complex task, and it may not be surprising that students encounter problems in performing this task. In the literature several types of difficulties are reported both with regard to the task perception, the content addressed and the tools used.

At the level of *task perception*, it has been found that students tend to view a modeling task as an engineering problem rather than a scientific one. Instead of focusing on the meaning of the model they focus on its output, and without proper reasoning the behavior turns to model fitting, i.e., tuning model parameters until the model output resembles the observed empirical data. Apart from a probable lack of success of this behavior, no deep reasoning on the model elements or model structure will take place (Stratford et al., 1998; Ogborn, 1999; Hogan & Thomas, 2001). Such model fitting behavior leads to a disconnection between the model and content knowledge: the model becomes an artifact that has to 'work', not something that provides explanatory power in understanding a phenomenon (e.g., Bliss, 1994; Hogan & Thomas, 2001). Moreover, some students even do not detect mismatches between the model output and the expected behavior of the phenomenon being modeled (Whitfield, 1988; Steed, 1994; Doerr, 1996). Finally, using empirical data gathered to generate hypotheses is a difficult process for students (De Jong & Van Joolingen, 1998).

At the *content* level, students have difficulties conceptualizing the complex phenomena which are typically addressed in computer-based modeling. Typical difficulties are the time dependence of variables and multiple processes that cancel out. Students tend to consider the influences of individual variables separately (e.g., Doerr, 1995; Stratford, et al., 1998; Hogan & Thomas, 2001; Kainz & Ossimitz, 2002). Also feedback mechanisms are profoundly difficult. If there is a feedback loop present in the system (e.g., the amount of interest on a bank account leads to an increase of the main sum, which in turn lead to a higher interest next year) students often fail to reason about interactive variables and display linear causal thinking instead, which means that students trace one cause to one effect (e.g., Riley, 1990; Kurtz dos Santos & Ogborn, 1994; Zohar, 1995; Löhner, Van Joolingen, & Savelsbergh, 2003).

At the level of *the tool*, students find it difficult to express their ideas in a modeling formalism. A typical modeling formalism, also used in the study presented in this chapter, is system dynamics (Forrester, 1961). System dynamics employs Stocks (represent the central quantity of the model which build up over time), flows (determine how quickly the stock changes), auxiliaries (factors that influence the flows), constants (factors with fixed values), and connectors (linkage arrows to show direction of proposed relations in a system). Expressing ideas in this (or other) formalisms requires learning and involves gaining some experience. It has been found, for example, that students find it particularly complex to represent flows as concrete elements in their model (Kurtz dos Santos & Ogborn, 1994; Hogan & Thomas, 2001). Also, students have difficulties in deciding the type of the variable they would like to implement into their model and, in addition, students frequently struggle in specifying the mathematical relations between variables in the model (Tinker, 1993; Cox & Webb, 1994; Sweeney & Sterman, 2000;

Ossimitz, 2002). At last, students find it difficult to translate their own knowledge of the phenomenon into a computer model using the formalism of an icon-based modeling tool (Kurtz dos Santos & Ogborn, 1994; Ainsworth, 1999).

These studies point to several structural difficulties students encounter during the modeling process. However, to frame these problems and to provide appropriate support, we need a thorough exploration of the reasoning processes students employ during modeling. As such, process-oriented studies can shed more light onto how levels of reasoning processes interact and contribute to successful or less successful modeling.

1.2 Reasoning processes during modeling activities

Several studies have aimed at describing students' reasoning processes. Most studies have taken inductive viewpoints, either in the form of individual case studies (Schecker, 1998; Stratford et al., 1998; Hogan & Thomas, 2001), or quantitative analyses (e.g., Zhang, Wu, Fretz, Krajcik, Marx, Davis & Soloway, 2002; Fretz et al., 2003; Löhner, Van Joolingen, Savelsbergh & Van Hout-Wolters, 2005). Other researchers have taken a normative stance, based on expert views, leading to descriptions of the ideal modeling process (Hestenes, 1987; Schecker, 1998; White & Frederiksen, 1998; De Jong, Van Joolingen, Lazonder, Ootes, Savelsbergh & Wilhelm, 2002). An overview of the outcomes from the aforementioned studies was presented by Löhner et al. (2005). This framework provides the basis for our explorative analysis. We reconsidered some labels and definitions to encompass representative global *reasoning processes* (see Figure 1).

This led to the following categories: Analyze, Inductive reasoning, Quantify, Evaluate, and Collect data. Collecting data involves the gathering of information through experimental inquiry. Because our focus is on the modeling process and because data collection and interpretation through inquiry involves many difficulties of its own (e.g., Kuhn, 1989; Chinn & Brewer, 1993; De Jong & Van Joolingen, 1998), collecting data will not be further addressed in this chapter. The category *Explain* was added to this framework to involve processes in which students articulate explanations to others. The way in which students justify model actions demonstrate how elaborate they reason about the material (e.g., Schecker, 1998; Stratford et al., 1998).
The reasoning processes specifically associated with modeling are:

- *Analyze*: When students are analyzing, they decompose the phenomenon they are studying into parts and identify important model elements (i.e., quantities, or relations between quantities) to be implemented in their model. Also, students interpret model output or empirical data that is presented in tables or graphs. Most of the modeling activities associated with this reasoning process are performed during the orientation phase of a modeling task. During Hestenes' (1987) *Description stage*, for example, students decide on the type of model that which will be constructed. In addition, they identify the variables and relations that have to be implemented in their model.

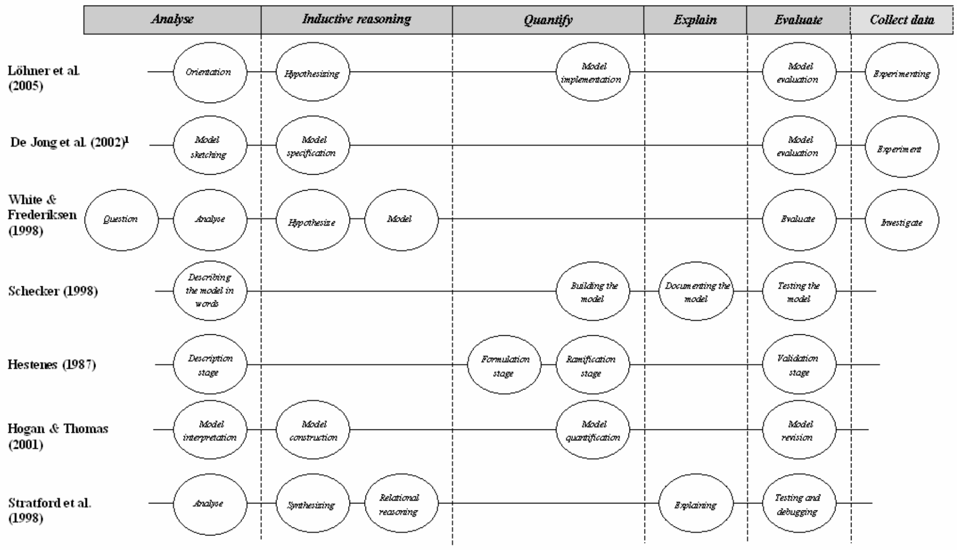


Figure 1. Overview of reasoning processes found in studies on computer-based modeling (adapted from Löhner et al., 2005).

This involves students analyzing graphs and tables identifying how variables increase or decrease over time. Both definitions were taken in our synthesis as indications of the process of analyzing.

- *Inductive reasoning*: Inductive reasoning occurs when students conjecture hypotheses on how model elements interact and on how the model should behave. This process implies a great deal of elaboration on the relations between the model structure and the behavior of the phenomenon being modeled, which makes it a complex process for students to perform (e.g., Schecker, 1993; Tinker, 1993). Stratford et al. (1998) use the term *Synthesizing*, which involves students making statements about the content, behavior, or structure of the model as a whole (e.g., considering how the model should behave, discussing the representation of their model, and discussing new relations between quantities in the model), which can be categorized as a process of inductive reasoning.
- *Quantify*: When students construct a preliminary model, they can make their ideas about model elements and relations more precise by expressing them into an executable mathematical format. This implies that quantities in the model are specified with a starting-value and relations are worked out in equations. This process of quantifying a model comprises the processes employed during Hestenes' (1987) *Formulation stage* and *Ramification stage* to a great extent. *The formulation stage*, according to Hestenes (1987), involves students using knowledge of physical laws to determine definite equations for the phenomenon that is modeled. During the *Ramification stage* the special mathematical properties and implications of the model are worked out. This stage implies that equations are solved out by experimenting with parameters and formulas.

- *Explain*: Involves the process in which students clarify to each other why model elements are related, that is, they document the reason(s) why one factor causes changes in another. The phase Schecker (1998) designates as *Documenting the model* clearly falls under this process.
- *Evaluate*: Finally, students have to connect between the output from their model and results obtained from experiments in order to evaluate and ultimately test their model. In evaluating their model, students determine if their model is consistent with their own beliefs, with data obtained from experiments and/ or with descriptions of behavior about the phenomenon being modeled. Model evaluation leads to model revision activities which involves modifying parts of the model so that it better describes or explains a given situation. The process of *Model revision* in Hogan & Thomas' (2001) study involves similar activities. They define *Model revision* as a process in which students assess the degree of fit between model output and expected or empirically-confirmed patterns.

The purpose of the present study is to understand the specific reasoning processes that play a role during students' activities. In order to gain insight in these reasoning processes we need to investigate the occurrences as well as the quality of these processes. More in particular, as the paragraph on student difficulties made clear, students may focus on particular aspects of their model, while ignoring other aspects (such as individual variables versus interacting variables or global model structure) as well as use inappropriate arguments, or even no arguments at all, to justify their reasoning. Therefore, in addition to identifying types of reasoning, the focus of the conversation and the types of argumentations used need to be understood. For these features – focus and argumentation – we developed a scheme for analysis in a more inductive fashion on basis of the obtained protocol data. The main research question in this study therefore is:

> *Which reasoning processes, employed by students during a computer-based modeling task, need support?*

This question encompasses the following sub questions:

1) What features are relevant to describe students' reasoning processes during computer-based modeling?
2) What distinguishes successful from less successful students?
3) Which reasoning processes are difficult for students to perform?

2. METHOD

2.1 Participants

This study involved thirty-eight students from eleventh-grade pre-university education, with a major in science. Students had no prior experience with system dynamics models. Students' age ranged between 16-18 years. During the task, participants worked in pairs, which were composed by having the students choose their own partners from within a group of familiar students.

2.2 Materials

Participants were presented with a task asking them to explore and revise a model that described the distance covered by an ice-skater[2]. Since participants had no prior experience with modeling, a completely open modeling task would be too complex for them to be successful within the time constraints of the experiment. Therefore, participants were given an incomplete model as a starting point. Such a model revision task enables students to concentrate on trying to comprehend and improve a model without having to start from scratch. The modeling task was implemented in Powersim. Powersim is a modeling tool based on system dynamics (see Figure 2). Powersim uses the five model building blocks characteristic for system dynamics modeling: Stocks, rates, auxiliaries, constants, and connectors. To insert a modeling element, students can drag and drop the icons on the screen they think are relevant for the phenomenon being modeled, creating a qualitative diagram of the phenomenon. While creating this diagram, students can quantify these elements by entering values and formulas. Once the model is quantified it can be executed. When students run their model, Powersim automatically generates the differential equations required to perform calculations. The results of simulations runs over time can be displayed as graphs or tables.

The modeling task was presented in a cover story in which a scientist attempted to construct a model of this phenomenon. Participants were provided with measurements that were obtained by the scientist, which they could use to test the model. The empirical data was presented in two graphs, one for the distance covered by the skater (see upper right hand side of Figure 2) and one graph for the velocity of the skater (lower right-hand side of Figure 2). Participants' task was to revise their model in such a way in that it would provide a good match with the data. Successful completion of the task would require the identification of two friction forces and of a feedback loop which runs from velocity to air resistance. This feedback implies that a skater at a higher velocity experiences more air friction, which consequently leads to a more rapid decrease in velocity.

The present study made use of data which were collected in order to compare the effects of different initially provided models and the effects of two sets of data which differed in quality. Therefore, dyads received slightly different versions of the modeling task. For the present purpose data could be pooled, since no significant differences were found between these settings on the dependent measures employed in this study.

---

[2] *Adapted from 'Computerondersteund modeleren natuurkunde: Een sportieve beweging' ['Computer-based modeling physics: A sportive movement'] (courtesy of Koos Kortland, Kees Hooyman, and Development Group Dynamic Modeling, University of Utrecht).*

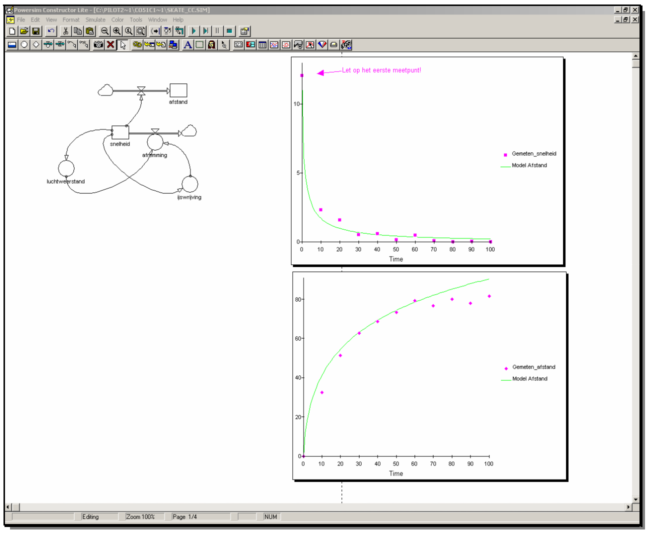


Figure 2. Screenshot of a model in Powersim.

2.3 Procedure

In order to get acquainted with system dynamics modeling in Powersim, each student individually worked through an instruction manual. This manual was adapted from: *Computer-based modeling: Manual Powersim*[3], which is developed and disseminated to Dutch schools by the Centre for Science and Mathematics Education at the University of Utrecht. In this instruction manual, students are presented with an example model of a water tank. The simplest system imaginable to illustrate fundamental aspects of modeling and the behavior of dynamic systems is that of a water tank containing a faucet and a drain. The volume of water in the water tank is represented by the stock variable (i.e., reservoir variable), and the flows represent respectively the inflow of water into the water tank via the faucet and the outflow of water from the water tank via a drain. The water tank model, its elements (i.e., variables and relations between variables) and how it can be built in Powersim were explained to students in the instruction. Also, students could execute and revise parts of the water tank model in the Powersim environment. The instruction took about one hour. Subsequently, participants were grouped into dyads. Participants were informed that they were going to explore and subsequently revise models working in couples. Next, dyads read the modeling task and were presented with the initial model version and with the data. Dyads were asked to collaboratively revise the model for approximately one and a half hour.

---

3 Translated from Dutch: 'Computerondersteund modeleren: Basishandleiding Powersim' which is available at: http://www.cdbeta.uu.nl/model/literatuur/basishandleiding.pdf

2.4 Data collection

The primary source of data consists of the verbal protocols of the collaborating dyads. Modeling actions and verbal communication between students were obtained using the program Lotus ScreenCam™. This program recorded all onscreen actions and audio. Verbal protocols were obtained by transcribing these recordings. After an initial qualitative, exploratory analysis, the transcripts were subsequently segmented into episodes. Episodes were scored using an analysis scheme that was developed on the basis of the qualitative analyses of the protocol data using our framework of reasoning processes as reference. Scoring the protocols was performed employing the program MEPA (Erkens, 1998). Unfortunately, because of recording software failure the verbal protocols of only thirteen dyads could be analyzed.

The quality of revised models was also assessed. The scoring was based on both the degree of model fit and the conceptual structure of the model. The model fit score, ranging from one to five points, was based on face value. The score for structure consisted of a score for correct/ incorrect quantities and a score for correct/ incorrect relations between quantities. For each correct quantity or relation in the model, one point was awarded. For each incorrect quantity or relation two points were subtracted. The model structure score was subsequently rescaled to a five-point scale. Addition of the two components led to a total model quality score on a ten-point scale.

3. RESULTS

3.1 Exploratory case studies

The goal of the data analysis was to characterize the reasoning processes students employ during computer-based modeling activities. To that end, we started with an exploratory qualitative analysis of the protocols of selected dyads. Selection was based on their model scores: we chose a high, a middle and a low scoring dyad. As an initial scheme of analysis we used our synthesis of reasoning processes identified in the literature as a starting point. In examining these protocols, focus of students' discussions and type of argumentation were also considered. Based on these ideas and a detailed study of the protocol data we subsequently modified the analysis scheme in several rounds of analysis.

3.1.1 Case 1: High performing dyad

Dave and Roel created an excellent model (model score: ten), which was also reflected in the quality of the processes they engaged in. Dave and Roel were very systematic in their approach, very elaborate in their reasoning, and critical in their discussions about their model. While implementing model revisions they carefully considered the semantics of the variables, relations and the behavior of their model. Also, they attempted to keep their model as simple as possible during the whole task.

The critical stance of Dave and Roel towards their model is exemplified by the following episode (Excerpt 1). Directly before this episode they were trying to fit the output of their model to the given data by adjusting parameters. Since their attempts remained without success, they engaged in an elaborate model evaluation.

1 D: Because ice friction is a force, deceleration is acceleration, how much does the skater decelerate per time unit
2 R: Negative acceleration
3 D: Negative acceleration
4 R: So the ice friction
5 R: But look, deceleration in the model is equal to ice friction, while the ice friction also has to be negative, so that is not right
6 D: So that is a mistake in our model
7 R: Let's read the assignment again
8 D: Let's start with the most important, we have the ice friction
9 D: Because we have to go from force to uhh velocity
10 R: Force, is it not possible to do that with the Work law?
11 D: Yes, but it would come in handy if we had the formula's, but I do not know them by heart
12 R: It says here that ice friction is equal to deceleration, that is not right
13 D: That is absolutely not possible, there has to be a step in between

Excerpt 1. (Protocol: ac_16, start: 0'16'20, duration: '52s, Process-code: evaluate; Focus-code: relation; Argumentation-code: physics knowledge).[4]

In this excerpt, Dave and Roel are *evaluating* a *relation* between variables. They conclude that the relation is incorrect and try to come up with a correct relation between these variables *using their knowledge of physics* as argumentation. They argue that the two variables in their model (i.e., 'ice friction' and 'deceleration') cannot be related in the way it was implemented (lines 1-5). Therefore, they try to find a physical law or formula which includes both variables in order to be able to compute one from the other (lines 8-13). Although they seem to know the relevant formulae, it appears that they had difficulty in translating these formulas into their computer model.

Dave and Roel acknowledged the importance of focusing on how the different relations in their model affected the model output, in order to figure out how their model works. For instance, in the following episode (Excerpt 2), Dave and Roel *reason inductively* about relations in their *model*, concluding that they have to proceed investigating individual relations in order to comprehend the behavior of their model. Therefore, Roel and Dave delete the variable 'air friction' from their model in order to separately investigate how ice friction affects the velocity of the ice-skater.

[4] *Below each protocol segment the following information is provided: protocol number, start time, duration of the episode and the final scoring of the episode employing our analysis scheme.*

R: So we have to ask ourselves when the skater is at 0 m/s and how much he must skate
D: Yes, only he is not constant
D: If velocity decreases
R: The deceleration is constant
D: No, not any more
R: Why not?
D: Because air friction is now dependent on your velocity
R: Oh, let's delete air friction then
D: Let us first begin with uh
R: It is not realistic
D: Let's first investigate the influence of ice friction on the skater's velocity

Excerpt 2. (Protocol: ac_16, start: 0'37'40, duration: '25s, Process-code: Inductive reasoning; Focus-code: model structure; Argumentation-code: experimental data).

An episode (Excerpt 3) in which Dave and Roel *reasoned inductively* about a key *relation* in their model, shows that they also base their arguments on *everyday experience*. In this segment, Dave and Roel reason about the relation between 'air friction' and 'velocity': air friction increases with higher velocity as we know from everyday experience. This leads Dave and Roel to recognize the need for the implementation of a feedback loop (i.e., between velocity and air friction) in their model.

Dave and Roel engaged in an elaborate approach to the modeling task, which was reflected in the use of high quality reasoning processes. They often reasoned about relations between quantities in their model and critically evaluated how their model worked. Mostly, their focus was on relations between variables or on the structure of their model. In addition, they frequently referred to experiential knowledge as an argumentation for model revisions. Finally, they often referred to physics formulas they had learned in class.

D: Yes do you see, you have to do something with air friction in order to relate it to velocity
D: Air friction depends on your velocity, if you go faster then your air friction is higher. If the wind goes faster, if you have a higher wind velocity, then your air friction is higher. Then there is no wind, you do not notice air friction if you are riding your bike, when wind force is 12 then you will not be able to go forward
D: If you are riding your bike and you ride with 1 km/h than you would not feel a thing
R: The harder you ride, the better
D: But if you go with 20 km/h, for example from a bridge, then you will notice the air friction
D: So the air friction comes in handy, now you have to combine it with velocity

Excerpt 3. (Protocol: ac_16, start: 0'42'12, duration: '42s, Process-code: Inductive reasoning; Focus-code: relation; Argumentation-code: experiential knowledge).

3.1.2 Case 2: Medium performing dyad

Laurette and Hinde created a reasonable model (model score: seven). They mainly analyzed and identified individual elements in their model without elaborating much on how their model could be improved. Instead of employing their own knowledge during modeling, they often referred to the degree of model fit as argumentation for model revisions. They also showed a concern for their model being a realistic representation of the phenomenon they were modeling.

In a significant part of their protocol Laurette and Hinde were engaged in unsuccessful model fit behavior:

L: The green line {the model graph}has to run as the other
H: Yes! It has to go through the little dots {the data}
L: Oh
H: But uh
H: Yes, you know if you notice the graph, like the one we just had, than we had given ice-friction the value 1
L: Yes
H: And then the graph went through the first dots, this graph, look

Excerpt 4. (Protocol: an_23, start: 0'10'17, duration: '24s, Process-code: quantify; Focus-code: quantity; Argumentation-code: correspondence model graph and data).

In this episode (Excerpt 4) Laurette and Hinde *quantify* the *quantity* 'ice friction' in their model to improve the *fit between the model graph and the data points*. Note that Laurette and Hinde implement this revision without further thinking about why their model yields this particular output.

Next to this superficial model fitting, they primarily *analyzed* and discussed (relevant) modeling entities (i.e., *quantities*):

L: This one {the model graph} is more correct
H: Hmmhmm
L: Shall we include another variable, did we forget something?
H: Yes, you also have air resistance, is that right?
L: Eeeuh yes but
H: But that has nothing to do with the ice-skater in this problem

Excerpt 5. (Protocol: an_23, start: 0'17'58, duration: '12s, Process-code: analyze; Focus-code: quantity; Argumentation-code: experiential knowledge).

In this episode (Excerpt 5), Laurette and Hinde identify a relevant variable that may be included in their model (i.e., 'air resistance') but refute the idea on the basis of *experiential knowledge*.

In a few episodes (e.g., Excerpt 6) in which Laurette and Hinde were *reasoning inductively* about their *model*, but they did not refer to relevant school physics knowledge:

H: And this arrow indicates that it, that the velocity
L: That it is a cycle, so if velocity is becoming less
H: The ice friction
L: Influences the distance
L: Yes, the distance becomes less but the velocity
with which the distance decreases
L: That changes
H: Eeeeuhm
L: But
H: A little
L: Deceleration is equal to ice friction
H: Do we have something that also influences the deceleration?

Excerpt 6. (Protocol: an_23, start: 0'18'23, duration: '31s, Process-code: Inductive reasoning; Focus-code: model structure; Argumentation-code: none).

In general, Laurette and Hinde were mainly engaged in analyzing modeling elements and did not justify their model revisions. The focus of Laurette and Hinde was similar to that of Dave and Roel, but the quality of their reasoning processes was less, because they hardly employed any argumentation. Nonetheless, Laurette and Hinde did show a preference for applying experiential knowledge in their thinking about their model.

3.1.3 Case 3: Low performing dyad

Marije and Lola constructed a poor model (model score: three). They spent much of their time attempting to adjust the model parameters to match the graph to the experimental data. They hardly came up with structural model revisions themselves. They relatively often requested guidance from the experimenter. These students did not seem to understand the purpose behind their model in specific and behind computer modeling in general. The main part of the protocol obtained from this dyad consists of episodes (e.g., Excerpt 7) in which they were *quantifying* individual *quantities*:

M: Here, wait try this one
M: No, that one
L: No, that is not correct
M: No
L: There has to be something, ice friction
L: I set ice friction on 8
M: Let's see what it {graph} does

Excerpt 7. (Protocol: an_20, start: 0'08'02, duration: '17s, Process-code: quantify; Focus-code: quantity; Argumentation-code: none).

Marije and Lola frequently quantified quantities in their model without consideration of why they chose particular values. It was quite difficult for Marije and Lola to see what specific kind of change was needed in order for them to get the model output fit the empirical data:

M: No
L: He {the graph} is becoming longer, much longer
M: Much longer indeed
L: This is not correct
M: Huh?
L: This is not okay
M: Eeeuhm
L: It must decrease much faster from here

Excerpt 8. (Protocol: an_20, start: 0'19'29, duration: '13s, Process-code: evaluate; Focus-code: model fit/ model output; Argumentation-code: none).

This protocol fragment (Excerpt 8) shows Marije and Lola *evaluating* the degree of *model fit*, but they do not know how to revise their model such that the fit improves.

In cases in which Marije and Lola *identified quantities* to include in their model, they did not provide any argument for this revision. For instance, in the following episode (Excerpt 9), Marije and Lola suggest to include the variable 'air-friction' as a constant to their model, without clarifying why they want to add this quantity.

M: Shall we make another square for the air-friction?
L: huh?
M: A square{i.e., constant}for air-friction?
L: Yes

Excerpt 9. (Protocol: an_20, start: 0'20'28, duration: '06s, Process-code: analyze; Focus-code: quantity; Argumentation-code: none).

In general, Marije and Lola's approach to modeling can be represented as model fit behavior. Their main focus was on individual quantities and they did not take into account relations between variables. Finally, they did not refer to their own experiential knowledge during modeling and did not provide evidence for their claims. Marije and Lola only shallowly processed the model task and they did not make the link between the behavior and structure of their model.

These case descriptions indicate the range of reasoning processes that were to be found in the protocols of the other dyads participating in our study (see Table 1). In the example excerpts given above, it becomes clear how these episodes can be attributed to the reasoning categories listed in the introduction. However, for a great deal of episodes no clear reasoning process could be discerned, as can be seen in Table 1. Consequently, new categories had to be added to our initial scheme in order to capture these activities.

The cases also show the need for additional scoring of students focus (i.e., the model elements students consider) during modeling and scoring of prior knowledge used, to grasp the differences in quality of the reasoning processes. For example, the less successful group (i.e., Marije and Lola) tended to focus on individual variables, which suggests that they did not see how model behavior depends on the influence of interacting variables. This is in contrast to Dave and Roel who did reason about relations. From the above it becomes clear that, in order to describe the reasoning processes of students and to identify difficulties, we need a coding scheme that takes into account the *focus* of reasoning as well as the underlying *arguments* that are put forward by the modeler.

Table 1.Percentage of time spent on global reasoning processes and model score (max. 10) for each dyad

Dyad	Global reasoning processes						Model score
	Analyze	*Inductive reasoning*	*Quantify*	*Explain*	*Evaluate*	*Other*	
Mark & Hugo	2.21	12.95	31.46	0.58	11.46	41.34	8
Jordy & Harry	7.04	9.96	22.68	1.93	8.97	49.42	7
Dave & Roel	7.75	16.22	35.44	1.98	3.39	35.22	10
Harma & Annemarie	6.99	7.41	32.82	0	8.51	44.27	5
Jaap & Robert	7.70	5.08	35.99	0	13.03	38.20	6
Rik & Corry	5.13	6.62	37.52	0	5.24	45.49	7
Gretha & Anique	3.31	8.19	39.07	0	4.88	44.55	5
Laurette & Hinde	9.27	11.36	35.57	0	4.40	39.40	7
Jos & Nadira	3.50	9.06	37.34	0	2.90	47.20	6
Marije & Lola	6.38	3.86	49.34	0	11.05	29.37	3
Paul & Benni	0.56	7.26	43.31	0	3.79	45.08	4
Denise & Astrid	4.23	4.62	38.53	1.10	6.46	45.06	4
Stephan & Sjoertje	2.38	0.74	31.03	0	1.94	63.91	8

3.2 Coding the data

Processes like analyzing or explaining involve several turns by both partners in a dyad. Therefore, the unit of analysis was determined to be at the 'episode' level, an episode being a period of coherent continuous talk on a single issue, rather than single utterances. As a practical operationalization, episodes were segmented on the basis of the following non-content criteria (cf. Chi, 1997):

- Following each run of the students' model.
- Following an interval of more than 15 seconds during which nothing is said.
- The interval during which the experimenter intervenes in the modeling process is a segment by itself.

In the majority of cases these criteria would lead to acceptable segments, however, some segments were of a much too long duration. Therefore, an additional criterion was applied:

- The maximum length for a segment is one minute. If a segment lasts longer, the segment will be more closely analyzed in order to see whether segmentation is possible on the basis of changes in reasoning process or changes in focus (often signaled by words as: 'Okay…', or 'Now….').

The segments obtained using this procedure were classified according to the reasoning process that was employed by students. We employed the definitions of the five clusters of reasoning processes (i.e., Analyze, Reason, Quantify, Explain, and Evaluate) identified in our synthesis of the literature. Four categories were added to the coding scheme in order to characterize episodes in which no clear reasoning process was shown, namely:

- *Guiding by experimenter:* Guidance by experimenter involves: a) providing an explanation of the assignment, b) explaining the tools and formalism used in Powersim and c) encouraging the collaboration.
- *Read & paraphrase:* Students read, paraphrase or discuss the assignment or modeling actions. When students are talking about the assignment or when they mention the actions they are performing, that episode is coded with read & paraphrase.
- *Off task:* Students talk about subjects that are unrelated to the assignment at hand.
- *Other:* Other processes that are not included in the analysis scheme. For example, inaudible murmur is coded as other.

For coding the focus of reasoning we introduce the following categories:

- *Quantity:* Students discuss quantities (i.e., constant or stock).
- *Relation:* Students discuss relations or interaction(s) between quantities.
- *Model output:* Students discuss the output generated by running the model, including the degree of fit between the output generated by their model and the experimental dataset.
- *Data points:* Students explicitly discuss the experimental data in the graph.
- *Model structure:* Students discuss the global structure of the model, for example: how the quantities are linked. In order to score this category, the episode has to involve more than one relation in the model.
- *Modeling actions/ the tool:* Students mention modeling actions, such as adding or deleting quantities, running the model, or adding or deleting relations. This category is also scored when students are trying to figure out how Powersim (i.e., tools, buttons, and formalism) works.
- *The assignment:* Students talk about the modeling assignment (i.e., the ice-skater problem).

Finally, we categorized the types of argumentation employed in reasoning:

- *None:* Students make no reference to specific knowledge in their reasoning.
- *Physics knowledge:* Students use terminology, concepts, or formulas that are used in physics. For example, if units for variables in the model are mentioned.
- *Mathematics knowledge:* Students use terminology, concepts, or formulas that are used in mathematics. For example, if students talk about the mathematical function of the model graph.
- *Experiential knowledge:* Students refer to experiential knowledge. For example, if students mention their own experiences.
- *Correspondence between model graph and data*: Students refer to (mis)match between the model output and experimental data. For example, students quantify a quantity in their model and refer to the degree of match between the data and the model graph (i.e., model fit behavior).

- *Experimental data:* Students refer to the data points without explicit reference to the model fit.

In Appendix A the complete coding scheme and the criteria for assigning the codes are presented. Interrater reliabilities for each of the three sub codes were determined by comparing the ratings of two independent judges. (n = 202; Process-code, Cohen's kappa: .74; Focus-code, Cohen's kappa: .76; and Argumentation-code, Cohen's kappa: .53). The interrater reliability for the first two categories can be regarded as satisfactory (Heuvelmans & Sanders, 1993). The Cohen's kappa for the argumentation code is low, and consequently, findings with respect to this aspect should be taken with care.

3.3 Quantitative results

The coding scheme was now applied to the entire set of protocols. Because protocols differ in length, and because protocol episodes are of different length as well, reporting frequencies give a skewed image. Therefore, frequencies were converted to proportion of total time for each dyad, and further analyses are based on these proportions. Data from these analyses are presented in Table 2 and Table 3.

From Table 2 it becomes clear that the amount (re)reading and paraphrasing (of the modeling actions and/ or modeling tool and of the assignment) and experimenter guidance is high for all dyads. Also, the time spent on quantifying is rather high. Many quantifying episodes have their focus on a single quantity, rather than relations between quantities. Quite often in these episodes reference is made to the degree of correspondence between model graph and data, whereas only a few refer to prior knowledge (see Table 3). The conjunction of these features clearly indicates model fit behavior. By contrast, most dyads did not spend much time on inductive reasoning; explaining seems to be almost a missing category; and use of prior knowledge is relatively rare.

In order to examine whether students follow a systematic approach in that there is a preferred sequence of activities, a transition analysis was performed. The analysis was conducted on transitions between episodes, which means that a significant transition occurs when the number of observed transitions between two reasoning processes is significantly higher than may be expected on the basis of the distribution of coding (Test of uni-directional transitions using z-scores, taking into account the conditional probabilities for every transition). The significant successions between the global reasoning processes found for the dyads in our study are shown in Figure 3. Line thickness indicates the magnitude of the difference between observed frequency and expected frequency.

Table 2. Percentage of time spent by type of reasoning process and type of focus averaged for all dyads

	Focus								
Reasoning process	*Quan-tity*	*Rela-tion*	*Modelfit/ model output*	*Data points*	*Model structure*	*Modeling actions/ the tool*	*The assign-ment*	*n.a.*	*Total*
Guidance	-	-	-	-	-	-	-	13.97	13.97
Evaluate	0.15	0.39	5.72	-	0.37	-	-	-	6.63
Explain	0.15	0.09	-	0.04	-	0.15	-	-	0.43
Quantify	32.80	3.41	-	-	-	-	-	-	36.21
Inductive reasoning	1.34	3.78	0.77	0.25	1.82	-	-	-	7.96
Analyze	2.16	0.63	0.68	0.52	1.03	-	-	-	5.02
Read & paraphrase	-	-	-	-	-	10.70	9.08	-	19.78
Off task	-	-	-	-	-	-	-	3.95	3.95
Other	-	-	-	-	-	-	-	5.95	5.95
Total	36.60	8.30	7.17	0.81	3.22	10.85	9.08	23.87	100

Table 3.Percentage of time spent by type of reasoning process and type of argumentation averaged for all dyads

	Argumentation							
Reasoning process	*None*	*Physics knowledge*	*Mathematics knowledge*	*Experiential knowledge*	*Correspondence between model graph & data*	*Experimental data*	*n.a.*	*Total*
Guidance	-	-	-	-	-	-	13.97	13.97
Evaluate	4.42	0.28	0.60	0.70	-	0.63	-	6.63
Explain	0.22	0.08	-	0.13	-	-	-	0.43
Quantify	18.59	3.40	0.93	1.61	11.55	0.13	-	36.21
Inductive reasoning	2.50	1.32	0.47	2.52	0.76	0.39	-	7.96
Analyze	3.80	0.48	0.07	0.27	0.21	0.19	-	5.02
Read & paraphrase	19.71	-	-	-	-	0.07	-	19.78
Off task	-	-	-	-	-	-	3.95	3.95
Other	-	-	-	-	-	-	5.95	5.95
Total	49.24	5.56	2.07	5.23	12.52	1.41	23.87	100

What is apparent from Figure 3 is that students frequently switch from analyzing, inductive reasoning, evaluating, or reading to guidance from the experimenter and vice versa. This means that after episodes in which students employed these processes they inquired for support from the experimenter. In addition, Figure 3 reveals that quantifying, reading, and inductive reasoning are persistent activities, in that they tend to extend across multiple episodes.

Finally, to investigate whether particular types of reasoning processes are associated with the quality of the model students created, we computed Spearman rank correlations between types of reasoning processes and the quality of the students' model. There was a significant negative correlation between the model score and the amount of time spent on quantifying quantities without argumentation (r = -.51, p = .04). In addition, a significant positive correlation was found between model score and inductive reasoning about relations between quantities with reference to experiential knowledge (r = .46, p = .049) and an almost signification positive correlation was found for inductive reasoning with reference to physics knowledge (r = .45, p = .08).

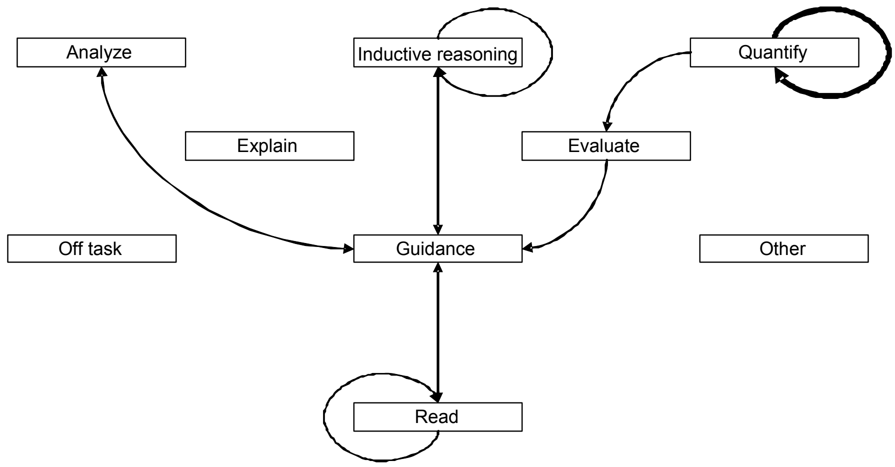


Figure 3. Transition diagram, displaying significant transitions between reasoning processes.

4. CONCLUSION AND DISCUSSION

The first part of our research concerned the features of reasoning processes that are needed to describe students' computer-based modeling processes. In order to identify these features, we started from a framework based on available research, and then refined this framework on the basis of a qualitative analysis of protocol data. This resulted in a framework containing the following dimensions: a) type of reasoning process, b) topic focus, and c) type of argumentation. Individual case studies made clear that all three dimensions are needed: apart from the *occurrence* of reasoning processes, it is important to also assess their *quality.* The first two dimensions could be reliably scored, whereas the type of argumentation remained difficult to decide on.

The second research question was how successful and less successful students differ in their reasoning. This question was first answered qualitatively, on the basis of three case studies, and next these answers were corroborated using a quantitative analysis. In the qualitative analysis we found that the more successful students, in contrast to the less successful ones, tended to justify their reasoning in terms of both experiential and physics prior knowledge. The less successful students were more narrowly focused on the model and the model output. Moreover, the more successful students regarded the model more as a whole, taking into account the model structure, whereas the less successful students mostly considered only one quantity at a time. In sum, the weaker students spent a relatively large part of their time manipulating parameters in order to let the model fit the given data, showing model fitting behavior. The quantitative analysis of the protocols confirmed this picture. Correlation analysis revealed that students who spent much time on quantifying quantities without argumentation (indicative of model fitting behavior), arrived at lower quality models. Students who spent their time on inductive reasoning with reference to prior knowledge arrived at better models.

The third research question was which reasoning processes are difficult for students to perform. It was found that, in general, students encountered a great deal of difficulties during the modeling task. This was indicated by the finding that the percentages of time spent on episodes in which students prompted for support or (re-) read the assignment were high. In addition transition analysis showed that episodes in which students were engaged in analyzing, inductive reasoning, evaluating, or reading were often followed by episodes in which they asked for support from the experimenter.

The particular problems students had during the modeling task are discussed according to the three levels at which difficulties are known to occur. At *the level of task perception*, we found frequent evidence of model fitting behavior. In addition, our case study revealed that the less successful students were found to be more engaged in model fitting compared to the more successful dyads. This model fitting behavior was, furthermore, found to be negatively associated with the quality of the students' model. Similar patterns were found in a case study by Ogborn (1999). This indicates that most students were not able to go beyond employing the model as an artifact instead of using the model as a means to comprehend the behavior of complex phenomena. At *the content level*, the quantitative analysis indicated that on average, students did not often connect to prior knowledge, which was taken as an indication that students had difficulties with relating their own prior knowledge to the phenomenon being modeled (cf. De Jong & Van Joolingen, 1998). This is in line with findings in earlier case studies (Stratford et al., 1998; Hogan and Thomas, 2001; Zhang et al., 2002). In addition, it was found that less successful students focused more often on individual variables in their model, implying that they had difficulties with considering interactions between variables. Finally, at *the level of the tool*, the case study and the quantitative analysis revealed that, even though students had received an instruction in dynamic modeling with Powersim, they still had difficulties in grasping the formalism used by Powersim. This was reflected in the finding that students' reasoning focused on the modeling tool during a great deal episodes (cf. Tinker, 1993; Cox & Webb, 1994).

From these findings it can be concluded that modeling of dynamic phenomena is a complex undertaking for students, and that probably more experience is needed in order to obtain a learning benefit. Consequently, appropriate support should be provided, either in the modeling tool or in the classroom context to scaffold students' reasoning. Note that our sample consisted of students with no prior experience with computer-based models. Thus, results of the present study and implications for scaffolding are applicable only to modelers who start to learn to use system dynamics models.

In the present study, it was found that when students employ their own knowledge during modeling activities, they constructed models of higher quality. Scaffolds should, thus, encourage students to activate their prior knowledge not only during modeling but also before engaging in any modeling activities. When students are initially prompted to think about variables and relations that could play a role in explaining the behavior of a dynamic phenomenon, activating whatever knowledge resources they have available, this may serve as an anchor for the further modeling process (e.g., Clement, Brown, & Zeitsman, 1989; Hammer, 2000). Ideally, this knowledge activation takes place within collaborative settings in which students can discuss their models with other groups (e.g., Rouwette et al., 2000).

The identification and articulation of specific modeling (sub-)goals also forms an essential aspect during the process of prior knowledge activation. When students have clear (sub)goals to attain, the modeling process will be more structured in the sense that students are guided in building an understanding of how the structure of their model influences the behavior of the model. In the present study, students were found to be primarily engaged in superficial model fitting behavior, instead of attempting to understand the association between the structure of their model and the phenomenon being modeled. By leading students to examine one model revision at a time, students should be able to discern what types of revisions produce what kinds of output, thereby building their understanding of the system they are modeling (e.g., Hogan & Thomas, 2001).

In order to motivate students to reason more deeply about their model, scaffolds could be offered to enable students to test their model against multiple datasets. This means that they have to compare not only model output to data but also different sets of data with each other. This may be a more fruitful learning experience for students, since multiple datasets trigger students to think of alternative variables or relations in the model revision process. As a result, students may not be primarily engaged in model fit behavior since their model has to be tested against several datasets.

In addition, students should be asked to model phenomena they already have knowledge of. This allows them to dedicate cognitive processing resources to translating their mental model into a system dynamics model instead of having to invest too much effort in identifying relevant variables and relations between them. It is important to free up that capacity for mastering modeling techniques during early stages of learning to model (Hogan & Thomas, 2001). This stance is supported by the finding that students in the present study had difficulties with comprehending the system dynamics modeling formalism of Powersim, even after they received an instruction.

Löhner et al. (2003) investigated the effect of two different external representations used in computer-based dynamic modeling tools on performance of secondary students. They compared a text-based model representation, in which students have to provide a list of equations before the model can be executed, with a graphical representation, in which the model is built by qualitatively linking variables. Results indicated that the different representations support different phases in the modeling process. Löhner et al. (2003) suggested that the graphical representation would be more suitable for the beginning of the modeling process, since this representation enables students to readily identify and implement variables and relations into their model. The text-based representation may be more appropriate for more advanced modeling, since students have to provide complete mathematical expressions to implement model elements.

In the present study students revised a graphical model within a quantitative system dynamics modeling environment. Given the finding that students, in the present study, had difficulties using this tool, it is suggested that students should first be asked to construct a qualitative graphical model. When constructing this model, students can focus on important variables and relations and on mastering the modeling formalism without having to be concerned about the mathematical form of the relations. Subsequently, they can proceed to quantify the relations between variables in a semi-quantitative form. This means that students can choose the qualitative form of each individual relation (such as: 'If A increases, then B also increases'). Finally, when students are more experienced with the modeling tool, relations between variables can be filled in quantitatively. A computer-based learning environment in which this kind of model progression (i.e., from building qualitative, semi-quantitative, to quantitative models) is implemented is Co-lab (Van Joolingen et al., 2005). The modeling tool in Co-Lab resembles Powersim in that the syntax is also based on system dynamics modeling. The modeling tool in Co-Lab enables students to specify variables by selecting pre-defined qualitative relations, drawing graphs or entering mathematical formulas.

Finally, in the present study it was found that less successful students primarily thought about individual variables while revising their model, which implies a bottom-up approach to modeling (Gobert & Discenna, 1997; Hogan, 1999; Hogan & Thomas, 2001). Students who employ a bottom-up approach to modeling do not consider how local model revisions impact the behavior of the model as a whole. More successful students employed a top-down approach to modeling, in contrast, which involves students considering interactions between variables in their model when revising their model. These students elaborate on the dynamics of their model when they change something in their model and, thus, take a more holistic view on their model. Students should be scaffolded to develop these more productive approaches to modeling, in which they learn to reflect on the impact of dependencies between variables on the dynamic behavior of their model. When students are introduced to modeling, it is argued that the top-down approach could be scaffolded by offering an expert model in order for them to productively model a certain phenomenon. When students take a top-down approach to modeling, they enhance their understanding of the workings of their model and ultimately learn more about the phenomenon being modeled.

Important to consider here is that the above mentioned scaffolds are suggestions offered to tackle the difficulties students have during the initial phases of modeling. Further research is needed in order to test these proposals.

CHAPTER 3

RELATING STUDENTS' EPISTEMOLOGICAL UNDERSTANDING OF COMPUTER-BASED MODELS WITH THEIR REASONING DURING MODELING*

While many educators and researchers in science have been arguing that students' epistemological understanding of models and modeling influences how they cognitively process a modeling task, there has been little evidence directly relating the two. Therefore, in this study students' level of epistemological understanding of models and modeling was examined as well as the relation between students' epistemological understanding and the level of their reasoning (i.e., deep versus surface reasoning) during modeling. Twenty-six students, working in dyads, were observed while working on a computer-based modeling task in the domain of physics. Students' epistemological understanding was assessed on four categories (i.e., nature of models, purposes of models, design and revision of models, and evaluation of models). Results indicate a significant relation between students' epistemological understanding and the level of their reasoning. From these results, we emphasize the necessity of considering epistemological issues in research as well as in educational practice.

1. INTRODUCTION

Models of scientific phenomena take a central place in science education (e.g., Gilbert et al., 1998; Raghavan et al., 1998; Schecker, 1993; Stratford, 1997; White & Frederiksen, 1998; Zaraza & Fisher, 1999). Several authors have argued that the activity of building and revising such models is particularly well suited to provide meaningful learning experiences. For instance, some emphasize the opportunity for students to think scientifically about the behavior of complex phenomena (e.g., Bliss, 1994; Jackson et al., 1996; Hestenes, 1997; Spector & Davidson, 1997; Wild, 1996). Others state that if students are allowed to construct models themselves, they can not only reflect upon the science content they are supposed to learn, but also upon the nature of their own knowledge (e.g.,

---

* Sins, P.H.M., Van Joolingen, W.R., Savelsbergh, E.R., & Van Hout-Wolters, B.H.A.M. *(submitted). Relating students' epistemological understanding of computer-based models with reasoning during modeling.*

Hogan & Thomas, 2001; Jonassen et al., 2005; Spector, 2000). According to this argument, models offer students the means to externalize and to test their own mental representations of scientific phenomena (Coon, 1988; Doyle & Ford, 1998; Penner, 2001).

However, these benefits of engaging students in modeling activities may only be realized if students understand the nature and the purpose of models in science, as well as comprehend how scientific models are constructed (e.g., Justi & Gilbert, 2002; 2003; Treagust, Chittleborough, & Mamiala, 2002). Such understanding may help students in using and in producing models, since it is assumed that students' epistemological understanding about models and the process of modeling is related to how they approach the material and ultimately to what they learn (Driver, Leach, Millar, & Scott, 1996; Gobert & Discenna 1997; Hammer, 1994; Schwarz & White, 2005; Van Driel & Verloop, 1999). For instance, a weak epistemological understanding may act to constrain students in using and developing scientific models (Crawford & Cullin, 2004). It seems, therefore, plausible to assume that students' epistemological understanding affects their reasoning during modeling (e.g., Hofer, 2001; Hofer & Pintrich, 1997; Schwarz, 2002a; Windschnitl & André, 1998).

Although there is ample evidence to suggest that students have a poor epistemological understanding of the nature of models, even when they are engaged in building models (e.g., Gobert, Snyder, & Houghton, 2002; Grosslight, Unger, Jay, & Smith, 1991; Roth & Roychoudhury, 1994; Ryder & Leach, 1999; Songer & Linn, 1991), there has been little evidence relating students' epistemology to the reasoning processes they employ during modeling. If such a relation exists, it may be valuable to support students develop a more sophisticated epistemological understanding of models and modeling (c.f. Schwarz & White, 2005). The present study investigates whether a relation between students' epistemological understanding and their reasoning during modeling can be established.

1.1 The epistemological status of models in science

To assess students' epistemological understanding, it is necessary to first portray an expert understanding of models and modeling. Scientists design models with a particular scientific purpose in mind, such as explaining, visualizing, or predicting the behavior of a scientific phenomenon (Giere, 1990; Gilbert, 1991; Gilbert, et al., 1998; Justi & Gilbert, 2002). Models cannot be completely accurate and are almost always tentative, in the sense that they are open to further revision and development (Crawford & Cullin, 2004). In addition, scientists can hold more than one model for the same phenomenon depending on the context, on the purpose of the scientific research and on the perspective of the scientist. An expressed model is available for other scientists to discuss and to reflect upon. Therefore, scientific enterprise can be conceptualized as a process of comparing and testing competing models (Giere, 1990; Hestenes, 1987; 1997; Penner, 2001).

1.2 Students' epistemological understanding of models and modeling

Grosslight et al. (1991) surveyed middle and high school science students and experts' epistemological understanding of models and modeling in science. Answers provided by participants to the interview questions were organized into five categories about models: kinds of models, purpose of models, designing and creating models, multiple models for the same phenomenon, and changing models. Grosslight et al. (1991) identified three general levels of understanding, which emerged from the interviews: A level 1 understanding, which corresponds to a 'naïve realist' epistemology of models and modeling, entails the notion that models are simple copies of reality. Students who hold a level 1 understanding do not recognize an ontological distinction between the observable objects and events of the material world and the entities that are created and defined for the purpose of building scientific models. Students who realize that there is a specific, explicit purpose that determines the way the model is constructed hold a level 2 understanding. Students in level 2 acknowledge that the modeler makes conscious choices about how to achieve the purpose. According to a level 2 understanding, the model no longer must exactly correspond with the real-world phenomenon being modeled. None of the students in Grosslight et al.'s (1991) study reached the highest level of epistemological understanding (i.e., level 3). Experts, however, articulated a clear level 3 understanding, in that they believed that models are constructed in service of developing and testing ideas rather than replicating reality. The modeler takes an active role in constructing the model. Experts acknowledged that models can be manipulated and subjected to tests in the service of informing ideas and generating predictions. The characteristic that distinguishes a level 2 from a level 3 understanding is, that the main focus of a level 2 conception is still on the model and the reality modeled and not on the ideas portrayed.

Employing this characterization, Grosslight et al. (1991) found that the majority (67%) of middle school students were at level 1, 12% were at level 2, and 18% reached an understanding that fell in between the levels 1 and 2 (i.e., mixed level). Of the high school students, only 23% had level 1 scores and the rest were split evenly between the mixed level 1/2 (36%) and level 2 (36%).

Other studies have corroborated Grosslight et al.'s (1991) findings. For instance, Treagust et al. (2002) found that many secondary science students did not understand how models are used in the development of scientific ideas. Most students valued the visual aspect of scientific models, but were not able to reason beyond the descriptive nature of models (cf. Barowy & Roberts, 1999). However, counter to the findings from Grosslight et al.'s study (1991), Treagust et al. (2002) found that students, in general, do have an understanding that it is possible to have multiple models for the same phenomenon and that each model displays a particular perspective or emphasis. Schwarz & White (2005) found that students' epistemological understanding regarding the creation, evaluation and revision of models was moderate (cf. Spitulnik et al., 1999). This was still the case after students had worked through a curriculum on the nature of scientific models and engaged in modeling activities. For instance, when asked about the relative value of alternative models, students responded that all models are of equal value. In addition, more than half of the students responded that model revision occurs when there is new information or evidence available, or even when the model is simply wrong. Studies by Crawford & Cullin (2004), Justi & Gilbert (2003), Harrison (2001), Lederman (1992), Smit & Finegold (1995) and Van Driel & Verloop (1999) show that even science teachers have a less than satisfactory epistemological understanding of the nature of models.

Most of the researchers, who conducted the previously cited studies, have been arguing that students' poor epistemological understanding of models and modeling may negatively affect how they cognitively process a modeling task. Accordingly, Hogan (1998) and Sandoval (2003) argue that analyses of students' epistemological understanding need to be supplemented by examinations of students' cognitive processing during scientific activity.

1.3 Studies relating students' epistemological understanding of models and modeling with their cognitive processing

Although there is ample evidence indicating that students' epistemological understanding of general science influences their cognitive processing during a scientific activity (e.g., Buffler, Allie, Lubben, & Campbell, 2001; Hammer, 1994; Millar, Lubben, Gott, & Duggan, 1994; Purdie, Hattie, & Douglas, 1996; Ryder & Leach, 1999; Schauble, Glaser, Raghavan, & Reiner, 1991; Schommer, 1990; Songer & Linn, 1991), only a small number of studies have attempted to investigate this relation within the area of scientific modeling. For instance, Ryder & Leach (1999) reported that a naïve epistemological understanding of science serves to constrain students' empirical enquiry. Schommer (1990) found that a weak epistemological understanding of learning (i.e., students who reported that successful students learn things quickly or who reported that knowledge is objective) was related to students oversimplifying their (inappropriate) conclusions, poor performance and overconfidence in test performance. Songer & Linn (1991) found that students with a dynamic understanding of scientific knowledge (i.e., scientific knowledge is understandable, interpretive, and integrated) acquired more integrated knowledge than those with a static understanding (i.e., scientific knowledge is static, memorization intensive, and incoherently organized). Finally, Hogan (1998) found that students' epistemological understanding of science was significantly related to the patterns of cognitive engagement that emerged during collaborative scientific knowledge-building tasks.

A problem with generalizing from these studies to the field of scientific modeling, however, is that the epistemological understanding of individuals is highly dependent on both domain and context (Hofer & Pintrich, 1997; Hogan, 2000). When learning science, students invariably use tacit epistemological understanding when doing something in a specific context. Indeed, different contexts make different demands upon students, and some forms of epistemological understanding may be more beneficial to learning in some contexts than in others (Elby & Hammer, 2001; Hogan, 1998; Leach, Millar, Ryder, & Séré, 2000; Schwarz, 2002b).

Within the context of scientific modeling, there has been only little evidence linking students' epistemological understanding and their cognitive processing during modeling with the exception of a few studies. First, Schwarz & White (2005) created and evaluated a model-centered approach to science education, in which middle school students learned about the nature of scientific models and engaged in the process of modeling. Based upon correlational results, including significant correlations of pre-test epistemology scores with post-test scores on an applied physics test, Schwarz & White (2005) argued that developing students' epistemological understanding of models and modeling may play a significant role in the acquisition of modeling skills and physics knowledge. Correspondingly, Gobert & Discenna (1997) found that students who hold a sophisticated epistemology of models were able to make significantly more inferences on the basis of their models compared to those who hold a naïve epistemological understanding. However, both studies did not directly observe the processes via which students had come to a better learning.

Talsma (2000) inferred students' epistemological and strategic understanding (i.e., understanding of students related to how they define, plan, implement and evaluate modeling activities) of models from notes they had made during the construction of computer models in the domain of ecosystems. She found that most students constructed models with a specific purpose focused on reality (i.e., level 2). In evaluating their models, students generally responded that they would add more variables, reflecting an epistemological understanding that is characterized by the addition of new information (i.e., level 2). Only few students showed a focus on ideas or the testing of possible theories (i.e., level 3). Regarding students' strategic understanding, Talsma (2000) found that many students were sufficiently able to define the phenomenon to be addressed by their models. In addition, all students constructed models that could be executed on the computer, but often included redundant relations and variables.

Although Talsma (2000) examined the strategic and epistemological understandings students demonstrate during model building, she did not explicitly relate these constructs. In addition, none of the above reported studies focused on whether and how students' epistemological understanding is related to the level of students' cognitive processing (i.e., deep versus surface processing) during modeling. For instance, high scores on epistemological understanding may be related to the employment of more deep processes during modeling, whereas less surface processes may be employed at higher epistemology scores (cf. Kardash & Howell, 2000; Windschnitl & André, 1998). *Deep cognitive processing*, as described in the work of Marton & Säljö (1976; 1997), Ramsden (1992), and Entwistle (1981; 1988; 2001), involves active learning processes, such as relating ideas, looking for patterns and principles and attempting to integrate new information with prior knowledge and experience, which create the optimal conditions for achievement in a variety of subject-matter areas (Covingtion, 1992). *Surface cognitive processing*, in contrast, entails processes without much reflecting and involves treating the learning material as more or less unrelated bits of information. Surface processing does not implicate elaboration of the learning material and leads to more restricted learning processes.

1.4 The current study

The first purpose of the present study is to investigate the level of secondary students' epistemological understanding of models and modeling. We take students' epistemological understanding of models and modeling to encompass their ideas about: a) the nature of models, b) the purpose of models, c) the evaluation of models, and d) the design and revision of models (cf. Crawford & Cullin, 2004; Grosslight et al., 1991; Schwarz & White, 2005). Students' epistemological understanding was assessed within a particular task context, in order to gain insight in their contextualized understanding as relevant to this task, rather than their generalized opinions. The specific task presented to students was a collaborative computer-based modeling task. The assessment (i.e., an open-ended questionnaire) we employed for capturing students' epistemological understanding, as well as the analyses performed on the answers they provided were framed within this specific context.

The second aim of our study is to examine how students' epistemological understanding is related to students' level of cognitive processing during modeling. In Chapter 2, we analyzed the nature of students' reasoning during modeling (Sins, Savelsbergh, & Van Joolingen, 2005; see Chapter 2). In its application to computer-based scientific modeling, the reasoning processes identified in Chapter 2 may be considered to be the cognitive processes concerned with the drawing of conclusions or inferences that relate to a particular action undertaken by the student (cf. Artzt & Armour-Thomas, 1992; Brickell, Ferry, & Harper, 2002). In this chapter we will focus on the relation between students' epistemological understanding of models and modeling and the level of their reasoning during modeling (i.e., deep and surface reasoning).

Our research questions are:

1) What is the level of secondary students' epistemological understanding of models and modeling?
2) What is the relation between students' epistemological understanding of models and modeling and their reasoning during a computer-based modeling task?

2. METHOD

2.1 Participants

The study involved twenty-six students from eleventh-grade pre-university education, with a major in science. Students had no prior experience with computer models. Students' age ranged between 16-18 years. Data were collected in the context of the study reported in Chapter 2 (Sins et al., 2005). Dyads were composed by having the students choose their own partners from within a group of familiar students.

2.2 Modeling task

Participants were presented with a task asking them to explore and extend a model that described the distance covered by an ice-skater[5]. Since participants had no prior experience with modeling, a completely open modeling task would be too complex for them to be successful within the time constraints of the experiment. Therefore, participants were given an incomplete model as a starting point. Such a model revision task enables students to concentrate on trying to comprehend and improve a model without having to

---

[5] *Adapted from 'Computerondersteund modeleren natuurkunde: Een sportieve beweging' ['Computer-based modeling physics: A sportive movement'] (courtesy of Koos Kortland, Kees Hooyman, and Development Group Dynamic Modeling, University of Utrecht).*

start from scratch. The modeling task was implemented in Powersim (Byrknes & Myrtveit, 1997). Powersim is a modeling tool based on system dynamics (see Figure 1). Powersim uses five model building blocks characteristic for system dynamics modeling: Stocks, rates, auxiliaries, constants, and connectors. Stocks represent a quantity that can increase or decrease from some starting value. A rate connected to a stock decides how quickly the quantity in the stock will change. Quantities can be represented either as constants (i.e., fixed values), or as auxiliaries (i.e., calculated from other quantities). Finally, connectors indicate dependencies between model elements. To insert a modeling element, students can drag and drop the icons on the screen they think are relevant for the phenomenon being modeled, creating a qualitative diagram of the phenomenon. While creating this diagram, students can quantify these elements by entering values and formulas. Once the model is quantified it can be executed. When students run their model, Powersim automatically generates the differential equations required to perform calculations. The results of simulations runs over time can be displayed as graphs or tables.

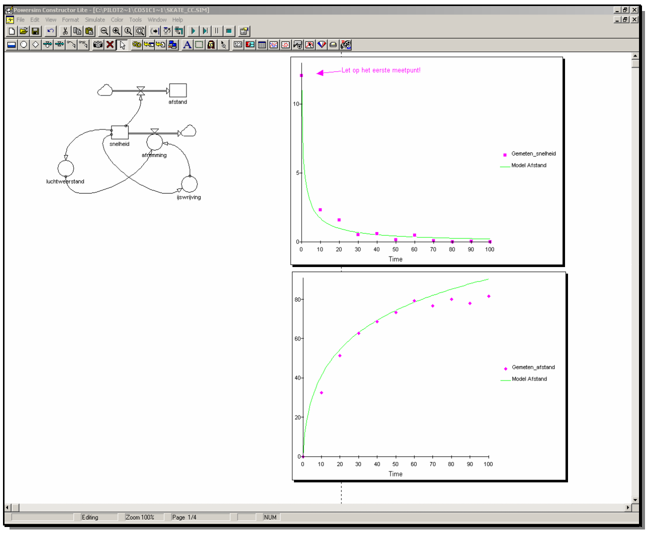


Figure 1. Screenshot of a model in Powersim.

The modeling task was presented in a cover story in which a scientist attempted to construct a model of this phenomenon. Participants were provided with measurements that were obtained by the scientist, which they could use to test the model. The empirical data was presented in two graphs, one for the distance covered by the skater (see upper right hand side of Figure 1) and one graph for the velocity of the skater (lower right-hand side of Figure 1). Participants' task was to extend their model in such a way in that it would provide a good match with the data. Successful completion of the task would require the identification of two friction forces and of a feedback loop which runs from velocity to air resistance. This feedback implies that a skater at a higher velocity experiences more air friction, which consequently leads to a more rapid decrease in velocity.

2.3 Data collection

2.3.1 Level of epistemological understanding

An open-ended questionnaire was developed to measure students' epistemological understanding of the nature of models, the purpose of models, the design and revision of models, and the evaluation of models. The questionnaire (see Appendix B) was patterned after the work of Grosslight et al. (1991). Answers of students to the items in our questionnaire were first analyzed qualitatively to examine whether they would fit in one of the four following categories:

1) Nature of models: How to define a model?
2) Purposes of models: What is the purpose of a model and modeling?
3) Design and revision of models: How are models constructed and when and how are they revised?
4) Evaluation of models: How are models evaluated and what are the criteria for the evaluation of a model?

The answers of every individual student were, first, interpreted in terms of the categories mentioned above. Next, it was investigated within every category whether the provided answers could be categorized according to level of sophistication. Each student was given a level of epistemological understanding rating (strong, moderate, weak) for each of the four categories, employing the three levels articulated in the studies of Carey & Smith (1993) and Grosslight et al. (1991; see Appendix C for scoring and examples). Two independent judges scored half of the total amount of protocols. Interrater reliability was determined by calculating Cohen's kappa, which was considered to be acceptable (n = 52, Cohen's kappa = .70) (Heuvelmans & Sanders, 1993).

2.3.2 Reasoning processes

Voice recordings of the collaborating dyads were recorded using the program Lotus ScreenCam™. Verbal protocols were obtained by transcribing the recordings. Students' reasoning during modeling was measured by analyzing the transcripts employing the protocol analysis scheme of Sins et al. (2005; see Chapter 2). These transcripts were scored employing two categories that were taken from the scheme of Sins et al. (2005; see Chapter 2)[6]: a) students' reasoning processes during modeling and b) type of reference (i.e., argumentation) made by students during reasoning (see Appendix A for coding scheme). Reasoning processes like analyzing or explaining may be considered to be cognitive processes and mostly involve several turns by both partners in a dyad. There-

6 In addition to these two categories, the protocol analysis scheme of Sins et al. (2005; see Chapter 2) also includes the category: 'topic focus of students' reasoning'. We did not include this code in the present analyses, since reasoning processes coupled with the type of reference (i.e., argumentation) students make during process-episodes provide sufficient information concerning students' level of cognitive processing.

fore, the unit of analysis is the process-episode level, an episode being a period of coherent continuous talk on a single issue, rather than single utterances (cf. Chi, 1997).

Reasoning episodes in which students are elaborating on the modeling task and connect to knowledge they have available, either gained from the task at hand or as prior knowledge, were designated as deep reasoning. Episodes in which students employ unelaborated reasoning processes without referring to available knowledge were labeled as surface reasoning. Only episodes that could be clearly marked as either *deep* or *surface* reasoning were counted in this analysis, excluding all others. The following specific codes were operationalized as indications of deep reasoning:

- Evaluating and reference to knowledge
- Explaining and reference to knowledge
- Quantifying and reference to knowledge
- Inductive reasoning and reference to knowledge
- Analyzing and reference to knowledge

The following codes indicated surface reasoning:

- Evaluating and no reference to knowledge
- Quantifying and no reference to knowledge
- Analyzing and no reference knowledge

Since the level of students' epistemological understanding was measured at the individual level and students' reasoning (i.e., deep versus surface reasoning) was measured at the dyad-level, we calculated the proportions of utterances that individual students contributed to the different types of process episodes. Each utterance in the transcript was assigned with the code that was designated to the episode in which that utterance was expressed. Because protocols differ in length, reporting frequencies will result in a skewed image. Therefore, for each individual student within a dyad, frequencies of these codes were converted to proportions of total number of utterances.

Interrater reliability for these codes was determined by comparing the ratings of two independent judges (n = 212; Cohen's kappa = .75). This interrater reliability can be regarded as satisfactory (Heuvelmans & Sanders, 1993).

2.4 Procedure

In order to get acquainted with system dynamics modeling in Powersim, each student individually worked through an instruction manual. This manual was adapted from: 'Computer-based modeling: Manual Powersim'[7], which is developed and disseminated to Dutch schools by the Centre for Science and Mathematics Education at the University of Utrecht. In this instruction manual, students are presented with an example model of a water tank. The simplest system imaginable to illustrate fundamental aspects of modeling and the behavior of dynamic systems is that of a water tank containing a faucet and a drain. The volume of water in the water tank is represented by the stock variable (i.e., reservoir variable), and the flows represent respectively the inflow of water into the water tank via the faucet and the outflow of water from the water tank via a drain. The wa-

7 Translated from Dutch: 'Computerondersteund modeleren: Basishandleiding Powersim' which is available at: http://www.cdbeta.uu.nl/model/literatuur/basishandleiding.pdf ter tank model, its elements (i.e., variables and relations between variables) and how it can be built in Powersim were explained to students in the instruction. Also, students could execute and revise parts of the water tank model in the Powersim environment. The instruction took about one hour. Subsequently, participants were grouped in dyads. Participants were informed that they were going to explore and subsequently revise models working in couples. Next, dyads read the modeling task and were presented with the initial model version and with the data. Dyads were asked to collaboratively revise the model for approximately one and a half hour. Finally, students were asked to complete the epistemology questionnaire.

3. RESULTS

3.1 Level of epistemological understanding

Results from our analysis of students' responses provided on the epistemology questionnaire are given in Table 1. For three of the four categories (i.e., *Nature of models*, *Design and revision of models*, and *Evaluation of models*) more than half of the students were assigned with a level 2 epistemological understanding. For the category *Purposes of models* 46% of the students scored a level 2 epistemological understanding and 42% of the students scored a level 3 epistemological understanding (see Table 1).

Table 1. Classification and examples of students' answers on the epistemology questionnaire

Epistemological Understanding	*Examples*	*Frequency*
1. Nature of Models		
Level 1	'A model is created by means of certain data'	6
Level 2	'A model is a simplified representation of reality'	17
Level 3	'A model predicts events that happen in reality, by using assumptions and calculations'	3
2. Purposes of Models		
Level 1	'With a model you do not have to measure everything and you can include everything in your model'	3
Level 2	'[A model is used] to obtain a clear overview of which variables are important'	12
Level 3	'[A model is used] to examine and discover relations en to make conclusions from the predictions made from a model'	11
3. Design and revision of Models		
Level 1	'Modeling is to construct a model of something'	3
Level 2	'A model can be improved by adding new variables or relations'	21
Level 3	'By constructing simple models, you can use them for multiple purposes'	2
4. Evaluation of Models		
Level 1	'[Models are evaluated by] good looking at them and seeing that you did not left something out'	5
Level 2	[A model is correct] if all known factors have been included and when the model more or less corresponds with the empirical data'	16
Level 3	[A model is correct] of its output corresponds with the mean result of multiple (at least 200) measurements'	5

3.2 Reasoning processes

Table 2 shows the mean proportions and examples of deep and surface reasoning processes students employed during modeling. What is apparent from Table 2 is that students employed much more surface processes compared to deep processes. Most surface processes involve students' *quantifying their model without referring to knowledge*, whereas for deep processes, the process that is employed the most is students' *quantifying and referring to knowledge*. The processes: *explaining and reference to knowledge* and *analyzing and reference to knowledge*, however, are almost missing categories (see Table 2).

Table 2. Mean proportions, standard deviations and examples of deep and surface reasoning processes

Level of students' reasoning	*Examples*	*M*	*SD*
Deep processes			
Evaluating and reference to knowledge	'The velocity is a particular deceleration that disappears [in the model], that is correct'	1.68	2.41
Explaining and reference to knowledge	'Look, the velocity, look, the harder you go, the more the deceleration you experience, so the velocity slows up, so the deceleration also becomes less'	0.22	0.46
Quantifying and reference to knowledge	'The larger the velocity, the larger the air friction, so therefore we have to multiply [these variables]'	7.09	6.69
Inductive reasoning and reference to knowledge	'Air friction depends on your velocity, if you go faster then your air friction is higher'	5.13	4.33
Analyzing and reference to knowledge	'We have to include air resistance in the model, because that is a relevant variable for the ice skater assignment, right?'	0.97	1.10
Surface processes			
Evaluating and no reference to knowledge	'This [the graph] is almost correct, look it runs through all [data] points'	6.02	3.17
Quantifying and no reference to knowledge	'Okay, we set velocity on 9, ok, run [the model]'	22.86	11.49
Analyzing and no reference to knowledge	'I think we have to include weight [to the model], the weight of the ice skater'	4.30	2.76

3.3 Relation between epistemological understanding and reasoning

To investigate whether there is a relation between level of students' reasoning and level of students' epistemological understanding of models and modeling, we aggregated and correlated the scores for both measures. Correlations were analyzed using the non-parametric Spearman rank method. Both the correlations between epistemological understanding and deep reasoning ($r = .40$, $p = .04$) and between epistemological understanding and surface processing ($r = -.51$, $p = .008$) were significant and in the expected direction.

4. CONCLUSION AND DISCUSSION

The purpose of the present study was twofold: 1) to examine secondary students' epistemological understanding of models and modeling, and 2) to investigate whether there is a relation between students' epistemological understanding and the level of their reasoning during modeling (i.e., deep versus surface reasoning). Since students' epistemological understanding is situation dependent, we framed our research questions within a particular context, that of a collaborative computer-based modeling task. We hypothesized that students' epistemological understanding and students' employment of deep processes are positively related, whereas a negative relation was expected between students' epistemology and the employment of surface processes.

Our findings indicate that, in general, the students in this study hold a moderate (i.e., level 2) epistemological understanding of models and modeling. Students reported that models are (simplified) representations of reality, that the design and revision of models involves adding, deleting and specifying variables (e.g., model-fitting) or relations and that models are evaluated by comparing them with data or by deciding whether all relevant variables are included in the model. With respect to students' epistemological understanding concerning the purposes of models, 88% of the students evenly had a level 2 or level 3 score. Students either reported models to be of use in providing an overview of the variables and the relations between them that play a role in describing the behavior of a phenomenon (i.e., level 2) or that models are useful in making predictions (i.e., level 3). Overall, only few students in the present sample held a level 1 epistemological understanding. In most respects, these results are in accordance with the studies cited in the introduction. However, our findings deviate from previous work in that in our study most students did not hold a naïve epistemological understanding of models, but instead indicated that a model should not necessarily has to precisely resemble the phenomenon that is being modeled.

An explanation for the differences found may be that the particular task context employed in the present study, in which a computer model of a particular scientific phenomenon is constructed and extended on the basis of empirical data, may have differently affected students epistemological understanding. The modeling task employed in Schwarz & White's (2005) study was highly structured, in the sense that students were asked to choose among three (or four) computer-modeling rules that most closely resembled their own mental model. Treagust et al. (2002) assessed students' epistemological understanding within the context of a general science curriculum.

The second research question concerned the relation between students' reasoning during modeling and their epistemological understanding of models and modeling. We found a significant positive relation between students' epistemological understanding and the employment of deep processes during modeling. Furthermore, high scores on epistemological understanding were related to students using significantly less surface processes. Thus, within the context of a collaborative computer-based modeling task, there is a significant association between how students conceptualize models and modeling and how they cognitively process a modeling task.

It is important to note, however, that based upon the correlational results reported in the present study, we cannot conclude that there is a causal relation between students' epistemological understanding and the way in which they processed the modeling task.

Although we may assume that having a sophisticated epistemological understanding may help students to process a modeling task more deeply, further research should focus on a possible causal relation. For instance, future studies could measure the development of both students' epistemological beliefs and their reasoning over a longer period of time. Also, there is a place for research on the effects of *instruction* in epistemological issues: what is the best way to raise student's epistemological levels and does such instruction have effect on the level of reasoning in modeling tasks?

We measured students' epistemological understanding of models and modeling at the individual level. In order to assess students' reasoning, we calculated proportions of total utterances an individual student contributed to process episodes. We acknowledge that the nature of students' contributions depends on utterances made by their group mate. Nevertheless, based upon students' contributions to either deep or surface process episodes, we argue that valid inferences about their reasoning can be made (cf. Hogan, 1999; Hogan & Thomas, 2001). For instance, students who are more capable to elaborate upon their model than their group mate, are also likely to be able to articulate a larger number of utterances during process-episodes that are scored as *deep*. In addition, we found a significant positive relation between students' epistemology and the amount of individual contributions made to episodes that were labeled as deep, which may be an important indication that students who made these contributions were processing the modeling task at a deep level. Nevertheless, an area for further research may be to investigate the possibility of assessing students' epistemological understanding on the dyad-level.

CHAPTER 4

EFFECTS OF FACE-TO-FACE VERSUS CHAT COMMUNICATION ON PERFORMANCE WITHIN A COMPUTER-BASED MODELING TASK*

In some collaborative computer-based learning environments, chat is offered as a means by which dispersed students, who are engaged in a co-construction task, can communicate. Still, findings from studies related to the influence of this type of communication on students' performance are rather contradictory. The purpose of the present study was to investigate the differential impact of chat communication versus face-to-face communication on performance within a collaborative computer-based modeling task. We expected that the reduced bandwidth of chat may either hinder students during modeling or may pressure them to increase the efficiency of their modeling. Two alternative hypotheses were considered: 1) Students' in the face-to-face condition will score significantly higher on model quality and will spend significantly less time on surface reasoning and significantly more time on deep reasoning compared to students in the chat condition. No significant difference is expected between the two conditions on the number of modeling activities performed. 2) Students in the face-to-face condition will score significantly lower on model quality and will spend significantly more time on surface reasoning and significantly less time on deep reasoning compared to students in the chat condition. In addition, students in the face-to-face condition will perform significantly more modeling activities than students in the chat condition. Results largely support the second hypothesis. Students in the chat condition compressed their interactions resulting in more efficient modeling, whereas students who communicated face-to-face spent significantly more time on surface processing of the computer-based modeling task.

1. INTRODUCTION

In instructional environments that stress collaborative construction of knowledge, the application of synchronous computer-mediated communication (i.e., chat) can be promoted as a means to support sharing of knowledge among students who are working together on a distance (e.g., Henri, 1992; Selinger, 1998). Although the availability of a chat tool in such environments makes collaboration between dispersed groups possible, the set of modalities by which learners can communicate are reduced, in contrast to face-to-face communication. Computer-mediated communication restricts the exchange of

---

∗ Sins, P.H.M., Van Joolingen, W.R., Savelsbergh, E.R., & Van Hout-Wolters, B.H.A.M. (submitted). *Effects of face-to-face versus chat communication on performance within a computer-based modeling task.*

auditory, visual and nonverbal communication cues which normally help groups to regulate interaction, express information and monitor feedback from others (Straus, 1997). A reduction in these cues may lead students to experience difficulty in following and understanding discussions. This raises the question of what the impact is of implementing computer-mediated chat on performance in co-construction tasks compared to face-to-face conditions. Purpose of the present study is to investigate whether and how groups who communicate through a chat tool differ from face-to-face groups regarding their performance (i.e., learning process and product). The study is performed in the context of a collaborative computer-based modeling task.

1.1 Collaborative computer-based modeling

Computer-based modeling is widely advocated as a way to offer learners an understanding of the process of scientific inquiry (e.g., Bliss, 1994; Gilbert et al., 1998; Hestenes, 1987; Jackson et al., 1996; Schecker, 1993; Stratford, 1997; White & Frederiksen, 1998). When modeling, learners construct external artifacts of the behavior of complex phenomena, such as ecosystems, water management or mechanical oscillations. These models can be constructed and executed with the help of a computer-based modeling tool, such as Powersim (Byrknes & Myrtveit, 1997), STELLA (Steed, 1992), and Modeling-Space (Avouris, Margaritis, Komis, Saez, Melendez, 2003). These tools enable learners to externalize and manipulate their mental representations of phenomena, which supports the reorganization and refinement of the students' conceptual understanding (e.g., Coon, 1988; Doerr, 1996; Doyle & Ford, 1998; Penner, 2001; Raghaven et al., 1998; Spector, 2000; Steed, 1994; Tinker, 1993; Wild, 1996). Also, externalizing ideas opens them up for criticism and discussion (Devi et al., 1996; Hogan & Thomas, 2001; Rouwette et al., 2000; Suthers, 1999; Suthers, Hundhausen, & Girardeau, 2003).

When constructing a model together, learners have to negotiate the meaning of individual contributions in order to establish common ground (i.e., shared knowledge). An important consideration in the design of these computer-based learning environments is whether chat is an appropriate means of communication in order for learners to engage in meaningful modeling; or is it preferable to perform collaborative modeling under face-to-face conditions?

1.2 Chat versus face-to-face communication

Online chat tools permit transmission only of written text, which imposes constraints on communication (Baker & Lund, 1997; Baltes, Dickson, Sherman, Bauer, & LaGanke, 2002; Doerry, 1995; Herring, 1999; McGrath & Hollingshead, 1993; Olson & Olson, 1997). All auditory information is eliminated and students do not have access to nonverbal information concerning whether and/ or how others are responding. In addition, this medium requires greater effort to compose and send messages compared to face-to-face interaction (Condon & Cech, 1996a; 1996b). It is also often impossible for students to respond while messages are being written. This simultaneous feedback plays an important role in signaling listener ship, timing turn-taking effectively, and maintaining continuous interaction (McLaughlin, 1984). However, a notable advantage of chat compared to face-to-face communication is that chat tools offer a history of sent messages. Because a history of students' communication is available, students have the advantage of seeing and revisiting the preceding interaction which is persevered on their computer screens. Face-to-face exchanges, by contrast, do not stay available after they have been uttered. Also, more exchanges, which may be relevant or irrelevant for task performance, are possible within a certain time frame in face-to-face situations compared to chat communication. These features of synchronous computer-mediated communication versus face-to-face communication may differentially affect performance on a co-construction learning task.

A number of studies have compared performance of face-to-face groups to groups using a chat facility. While their results are contradictory in some areas, consistent results have also been reported. For example, computer-mediated discourse requires longer time to complete a task and is reported to be more difficult (e.g., Baltes et al., 2002; Bordia, 1997; Daly, 1993; Mennecke, Valacich, & Wheeler, 2000; Straus & McGrath, 1994; Suthers et al., 2003). In addition, online groups produce fewer and shorter remarks within a given time period than face-to-face groups (e.g., Condon & Cech, 1996a; 1996b; Lebie, Rhoades, & McGrath, 1996; Reid & Reid, 2005; Straus, 1997; Van der Meijden & Veenman, 2005).

Contradictory results have been reported with respect to the effects of communication mode on outcome measures. For instance, although most studies found that students were more satisfied with a face-to-face collaborative process than with a computer-mediated collaboration process (e.g., Carey & Kacmar, 1997; Johnson, Aragon, Shaik, & Palma-Rivas, 2000; Olaniran, Savage, & Sorenson, 1996; Straus, 1996), studies by Jonassen and Kwon (2001) and by Cohen and Scardamalia (1998) show that chat groups were more satisfied compared to students who communicated face-to-face. In addition, some studies found that performance outcome scores of face-to-face groups were significantly higher compared to scores of groups who communicated through a chat tool (e.g., Carey & Kacmar, 1997; Chiu, Wu, & Huang, 2000; Hollingshead, McGrath, & O'Connor, 1993; Mennecke et al., 2000; Straus, 1997; Suthers et al., 2003; Van der Meijden & Veenman, 2005). In contrast, Barile & Durso (2002), Basque & Pudelko (2004), and Dennis (2003) found that the use of a chat tool proved to be as effective as face-to-face communication.

It is important to note, however, that the effects of these modes of communication on task performance are not at all generic, but are dependent on task characteristics and learning objectives (Bordia, 1997; Hollingshead et al., 1993; Rana, Turoff, & Hiltz, 1997; Zigurs & Buckland, 1998). A framework that takes these issues into account in explaining and predicting the effect of communication mode on task performance is the task-media fit hypothesis of McGrath & Hollingshead (1993).

McGrath & Hollingshead (1993) argue that the effectiveness of a communication mode for a given task depends on the degree of fit between the richness of information that can be transmitted through that system's technology and the information richness requirements of that task. Rich information provides a great deal of cues, via nonverbal or paraverbal (e.g., facial expression or vocal pitch) channels, that reduce equivocality by aiding in the interpretation and comprehension of messages (Daft & Lengel, 1984; 1986). They propose four general collaboration task categories (McGrath, 1984): generate tasks (i.e., generating ideas or plans), intellective tasks (i.e., tasks with a demonstrably correct answer), decision-making tasks (i.e., tasks involving attaining consensus because there are no demonstrably correct solutions), and negotiating tasks (i.e., task in which students have to reconcile their conflicts and viewpoints). These tasks represent a theoretical continuum such that each task type requires successively increasing degrees of interdependence between collaborators for successful task performance (Daly, 1993; Hollingshead et al., 1993; Straus, 1996; 1999). At each successive level of interdependence, the group's need for richness of information increases.

According to the task-media fit hypothesis of McGrath & Hollingshead (1993), an appropriate fit between the information richness requirements of the task and the information richness that can be conveyed by the communication medium should result in higher task performance. The task-media fit hypothesis predicts that as the communication demands increase from generate tasks to negotiating tasks, the effectiveness and efficiency of computer-mediated communication decreases (McGrath & Hollingshead, 1993).

However, the taxonomy employed by McGrath & Hollingshead (1993) is limited in its usefulness for predicting the impact of communication mode on performance in the collaborative computer-based modeling task that is the focus of the present study. The difficulty with this classification is its insistence upon mutually exclusive categorization of tasks. According to McGrath & Hollingshead's (1993) task classification, the modeling task can be viewed as an intellective task, in which a demonstrably correct answer needs to be invented, selected, and computed. However, classifying the modeling task as merely an intellective task is insufficient, since students also have to agree upon revisions they implement in their model, which involves negotiation processes. Another limitation is that McGrath & Hollingshead (1993) categorize tasks by their (learning) objective, i.e., what learners are supposed to perform in order to accomplish the task. As Zigurs & Buckland (1998) argue, defining a learning task should not only include *what* must be accomplished in order to meet stated goals, but also *how* those goals are attained by learners, i.e., the processes by which the task is carried out. Finally, Rana et al. (1997) criticize the task-media fit hypothesis of not explicitly recognizing task complexity[8] as an important task dimension. The information richness construct has, thus, to be considered in light of more than the task's objective; the processes employed to complete a learning task and the difficulties encountered by students during task execution. For example, even when students are presented with an apparent simple learning task (e.g., generate task) that, according to the task-media fit hypothesis, fits an information-poor communication mode (i.e., chat communication), students may need a richer communication mode (i.e., face-to-face communication) when they encounter complexities during the process of completing the task.

Consequently, we may hypothesize that modeling tasks may carry on communication channels that enable relatively rich information exchanges, such as face-to-face communication. Constructing a computer model is a complex task for students to be engaged in and requires students to perform activities which are typically associated with experts, such as: identifying and analyzing variables and relations between variables, specifying variables, evaluating model output, and reasoning about how the model calculates its

8 Task complexity is defined here as the subjective mental load a learner experiences during a learning task.

output (Hestenes, 1987; Hogan & Thomas, 2001; Ogborn, 1999; Stratford et al., 1998). Sins et al. (2005; see Chapter 2) found that students encounter a great deal of difficulties during a collaborative computer-based modeling task[9]. For instance, they reported that students have difficulties with relating the computer model to their own conception of the phenomenon being modeled, with considering interactions between variables and with grasping the formalism employed by the modeling tool. Computer-based modeling may, thus, necessitate for a richer communication mode (i.e., face-to-face communication) and using chat communication may hinder students during modeling.

An alternative hypothesis could be that groups who use computer-mediated chat compensate for the communication constraints by being more concise in their interactions, as argued by Condon & Cech, (1996a; 1996b), Jonassen & Kwon (2001) and Newlands, Anderson, & Mullin (2003). Students using a chat tool eliminate unnecessary elaborations and repetitions, and seeking to increase the efficiency of their communication because of the slower pace of interaction. Because exchanging information is more difficult in synchronous computer-mediated communication tools compared to face-to-face communication, students using these tools could be more task oriented and compress their communication resulting in a more efficient interaction. For instance, Ruberg, Moore & Taylor (1996) found that synchronous computer-mediated communication leads to more experimentation, sharing of ideas, increased and more distributed participation compared to face-to-face communication.

1.3 Research question and hypotheses

Main focus of the present study is to investigate the effect of communication mode (i.e., chat versus face-to-face communication) on students' performance in a collaborative computer-based modeling task. We conceptualized modeling performance as: a) students' modeling activities b) students' reasoning during modeling, and c) the quality of the students' models. We will address the following main research question:

What is the effect of chat communication versus face-to-face communication on students' modeling activities, on their reasoning during modeling, and on the quality of their model?

From our review of literature, we deduce two divergent hypotheses with which we approach this question, with no prior expectation that one would dominate:

- *Hypothesis 1*: Face-to-face groups will perform better than groups who communicate through a synchronous computer-mediated communication tool. The complex nature of a computer-based modeling task requires rich information to be communicated in order to attain effective task performance. Communication constraints of the chat tool may hinder students modeling performance.
- *Hypothesis 2*: Groups who communicate through a synchronous computer-mediated communication tool perform better than groups communicating face-to-face. Communication constraints of the chat tool may pressure students to increase the effi-

9 In this study dyads collaborated in a face-to-face setting.

ciency of their interaction by compressing it. Face-to-face communication, in contrast, may involve more interactions in which nonessential information is exchanged. This may result in lack of focus and distraction in students' communication, which may be detrimental to task performance.

Table 1 provides a schematic overview of expectations derived from both hypotheses. Both hypotheses predict that the chat condition will show a lesser amount of total communication, because they will be more constrained in their communication compared to students in the face-to-face condition. The hypotheses will differ regarding their prediction of the impact of this communication constraint on students' reasoning during modeling and on the quality of the model students construct. If the first hypothesis is supported, the face-to-face condition will score significantly higher on model quality than the chat condition. In addition, students in the chat condition will use significantly more surface processes during modeling, whereas students in the face-to-face condition will employ significantly more deep processes compared to the chat condition. Deep reasoning involves elaborate cognitive processing in which students refer to prior knowledge they have available. On the other hand, surface reasoning involves restricted, unelaborated processes without referring to knowledge (cf. Entwistle, 1979; 1988; Marton & Säljö, 1976; 1997). Chat tools are not capable of transmitting the types or amount of communication needed to effectively address the modeling task. Consequently, students who communicate using chat will employ less deep processes and mainly resort to surface reasoning of the modeling task. Hypothesis 1 makes no clear predictions for a difference between conditions on amount of modeling activities.

If the second hypothesis is supported, the chat condition will score higher on model quality compared to the face-to-face condition. Also, the chat condition will employ significantly more deep processes compared to the face-to-face condition, whereas it is expected that students in the face-to-face condition will use significantly more surface processes. In addition, students in the chat condition will perform less modeling activities compared to students in the face-to-face condition, since they will be more precise in revising their model.

Table 1.Overview of the predictions derived from Hypothesis 1 and from Hypothesis 2

Modeling performance	*Hypothesis 1*			*Hypothesis 2*		
1. Modeling activities	Chat	<>	Face-to-face	Chat	<	Face-to-face
2. Reasoning processes						
a. Total amount	Chat	<	Face-to-face	Chat	<	Face-to-face
b. Proportion of time spent on						
b1. Deep processes	Chat	<	Face-to-face	Chat	>	Face-to-face
b2. Surface processes	Chat	>	Face-to-face	Chat	<	Face-to-face
3. Model quality	Chat	<	Face-to-face	Chat	>	Face-to-face

2. METHOD

2.1 Participants

Forty-four students (aged 16-18 years) from eleventh-grade pre-university education, with a major in science participated in our study.

For assigning students to dyads, we preferred a heterogeneous group composition, since in previous studies students with different school grades had been generally more successful working together than homogeneous groups (e.g., Gijlers & de Jong, 2005; Webb 1991; Webb, Welner, & Zuniga, 2001). The reason is that higher achieving students can learn from giving explanations, whereas the lower achieving student can learn from these explanations given (Hooper & Hannafin, 1991; Webb & Farivar, 1994). However, the difference in level between students should not be too large. We used the students' average school grade in science as a measure for group composition. The mean average grade of all students was 6.87 on a scale from 1 to 10, with a standard deviation of 0.98. In order to assure moderately heterogeneous dyads, the group of participants was divided into two equal groups. One group consisted of the top 25% as well as the bottom 25% in average grade for science. The other group consisted of the remaining 50%. Dyads were composed by letting students choose a partner from the other group. This assured a moderate difference between partners in science ability, as well as pairs who had chosen each other to work with. This procedure meant that half of the dyads were low-middle dyads in terms of average grade, whereas the other half was middle-high. Given the low variance in average school grade, as well as the fact that all students were new to the task and the domain, we did not expect any differences from this division.

All participants were familiar with using chat software. Dyads were randomly distributed into two conditions: a chat condition (n = 11) and a face-to-face condition (n = 11). The average grades for the chat condition (M = 6.90, SD = 0.82) and for the face-to-face condition (M = 6.84, SD = 1.10) did not differ significantly (U = 215, p = .53).

2.2 Material

Students performed the modeling task within the Co-Lab environment (van Joolingen et al., 2005). They were asked to revise a simple pre-build model that could give an explanation and prediction of the temperature on earth. The task was simplified to some extent, since the earth in this task was represented by an irradiated black sphere (see Appendix D for the assignment). Participants were given a small initial model as a starting point. Students' task was to extend this model in such a way in that it would ultimately match the data they obtained from experiments. Students could conduct experiments in Co-Lab with a simulation of a black sphere that could be heated (see Figure 1 for a screenshot of the simulation).

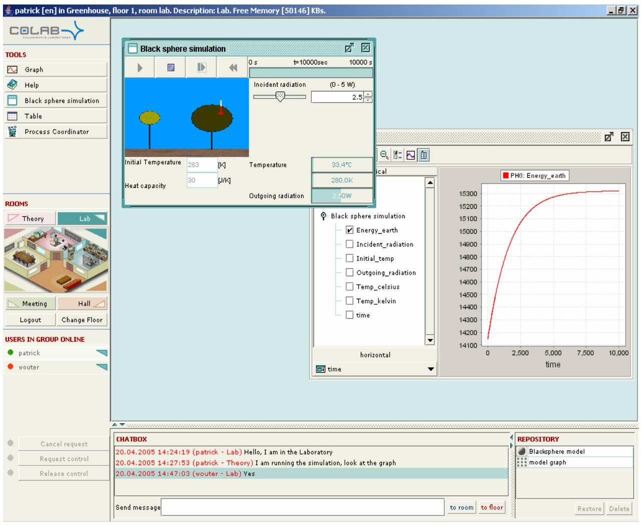


Figure 1. Screenshot of the simulation of the temperature of an irradiated black sphere. Results of the experiments are provided in a graph or table.

Students constructed their models in the model editor tool of Co-Lab (see Figure 2). The model editor in Co-Lab employs the system dynamics formalism (see Forrester, 1961). The model can be executed and results of model runs over time can be displayed as graphs or tables. In order to test their models, students can compare the output of their model with data collected from the simulation. As a result, students may revise their model on the basis of the testing outcomes.

Successful completion of the task would require the identification of the variables heat capacity and temperature, of the relation between these variables, and of a feedback loop which runs from energy earth to temperature (see Figure 2 for the correct model structure). This feedback implies that when the energy of the earth is higher the temperature on earth is also higher, which consequently leads to a more rapid decrease in energy on earth (i.e., a higher energy outflow).

Students in the chat condition could communicate by means of typing their messages in a chat box provided by Co-Lab. Messages typed into an entry box were sent to both participants' shared chat displays once the Enter key was pressed. The face-to-face condition differed only in the absence of the chat facility. Students in this condition sat behind one screen together.

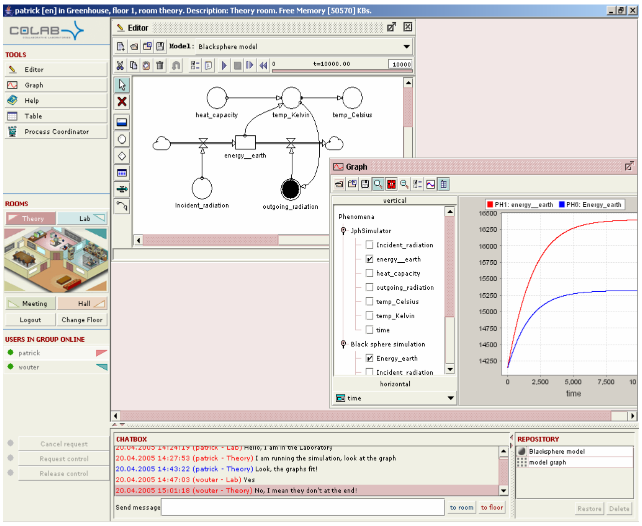


Figure 2. *Screenshot of the model editor in Co-Lab. In this particular example an accurate model diagram is provided of the black sphere simulation. This model shows that the energy of the earth is influenced by the incident radiation from the sun (i.e., energy inflow) and the outgoing radiation (i.e., energy outflow). The outflow is influenced by the temperature of the earth. Finally, the temperature is influenced by the energy of the earth and the heat capacity of the earth.*

2.3 Measurement of the performance variables

2.3.1 Modeling activities

Throughout the Co-Lab activity, all students' computer interactions were recorded. Activities that are directly related to the students' model were obtained from the log-files. Modeling activities discerned in this study were: Specifying quantities, adding model elements (i.e., quantities, relations, or flows), deleting model elements, saving models, opening saved models, and running models. Frequencies of students' modeling activities were used in the analyses.

2.3.2 Reasoning processes

To obtain verbal protocols, voice recordings of dyads working in the face-to-face condition were transcribed, whereas protocols of the chat messages were obtained from the software logs. Reasoning processes students employed during the modeling task were measured by scoring the transcripts on two categories that were taken from the protocol analysis scheme of Sins et al. (2005; see Chapter 2)[10]: a) students' reasoning processes during modeling and b) type of reference (i.e., argumentation) made by students during reasoning (see Appendix A). Reasoning processes like analyzing or explaining may involve several turns by both partners in a dyad. Therefore, the unit of analysis is the 'process-episode' level, an episode being a period of coherent continuous talk on a single issue, rather than single utterances (cf. Chi, 1997).

Reasoning episodes in which students are elaborating on the modeling task and connect to knowledge they have available, either gained from the task at hand or as prior knowledge, were designated as deep reasoning. Episodes in which students employ unelaborated reasoning processes without referring to available knowledge were labeled as surface reasoning. Only episodes that could be clearly marked as either *deep* or *surface* reasoning were counted in this analysis, excluding all others. The following specific codes were operationalized as indications of deep reasoning:

- Evaluating and reference to knowledge
- Explaining and reference to knowledge
- Quantifying and reference to knowledge
- Inductive reasoning and reference to model components[11]
- Inductive reasoning and reference to knowledge
- Analyzing and reference to knowledge

The following codes indicated surface reasoning:

- Evaluating and no reference to knowledge
- Quantifying and no reference to knowledge
- Analyzing and no reference knowledge

Transcripts were segmented into episodes and were scored with the help of the coding scheme. Because protocols differ in length, and because protocol episodes are of different length as well, frequencies of each code were converted to proportions of total time for each dyad and further analyses are based on these proportions.

2.3.3 Model quality

The quality of the models students constructed were assessed with the help of a model score sheet. Students' final model was awarded with points for each variable which name and specification was correct. In addition, points were awarded for correct links between variables and for correct specifications for these relations. Finally, for each incorrect relation between quantities a point was subtracted from the total score.

---

10 In addition to these two categories, the protocol analysis scheme of Sins et al. (2005; see Chapter 2) also includes the category: 'topic focus of students' reasoning'. We did not include this code in the present analyses, since reasoning processes coupled with the type of reference (i.e., argumentation) students make during process-episodes provide sufficient information concerning students' level of processing.

11 We included this code as an indication of deep processing, since students who reason inductively and refer to model components are elaborating on their model. In addition, students who use chat may also communicate more via the components present in the representation (cf. Suthers et al., 2003).

2.4 Procedure

The full experiment took two sessions, each of about two and a half hours on two separate days. In the first session, students were first presented with a short plenary introduction to the Co-Lab environment. Then, in both conditions, students were provided with an instruction manual to get acquainted with the system dynamics modeling in Co-Lab. In this instruction manual, students were presented with an example model of a water tank. The water tank model, its elements (i.e., variables and relations between variables) and how it can be built in Co-Lab were explained and students could execute the model in Co-Lab. The instruction took about one and a half hour. Subsequently, students were asked to revise the water tank model themselves for the remaining time. In the second session, dyads were presented with the black sphere modeling task.

3. RESULTS

3.1 Modeling activities

Table 2 shows the mean frequencies of modeling activities dyads performed during the modeling task. The standard deviations of these frequencies are high, indicating large differences between dyads. Because the sample size was relatively small, between-group differences were analyzed using non-parametric Mann-Whitney U test.

Dyads in the face-to-face condition saved their models significantly more often than dyads in the chat condition ($U = 30.50$, $p = .048$). In addition, dyads in the face-to-face condition executed their model significantly more often than dyads in the chat condition ($U = 28$, $p = .03$).

Table 2.Mean frequencies and standard deviations of modeling activities and the results of a Mann-Whitney U test for differences between the chat condition and the face-to-face condition

	Chat condition		*Face-to-Face condition*		*Mann-Whitney U*	
Modeling activities	*M*	*SD*	*M*	*SD*	*U*	*p*
Specifying quantities	25.52	24.04	42.01	29.65	41.5	.21
Adding model elements	18.89	12.65	15.20	11.63	49.0	.45
Deleting model elements	6.28	5.28	7.52	6.11	56.0	.77
Saving models	8.30	9.62	14.26	9.24	30.5*	.048
Opening saved models	22.01	22.03	16.88	8.77	57.0	.82
Running models	8.07	11.82	15.22	9.42	28.0*	.03

U and p scores were obtained with a non-parametric Mann-Whitney U-test.
** $p < 0.05$.*

3.2 Reasoning processes

The total amount of process-episodes was strongly different between the two conditions, with dyads in the face-to-face condition having significantly more episodes (M = 85.64, SD = 16.78) compared to dyads in the chat condition (M = 20.82, SD = 9.78; U = 0.0, p < .001).

Table 3 shows the mean proportions of time the groups spent on the deep and surface processes. Dyads largely differed in the proportion of time they spent on these cognitive processes, as indicated by the high standard deviations.

When looking at the differences between the two conditions in proportion of time spent on deep reasoning, one significant difference is found. Dyads in the chat condition spent significantly more time inductively reasoning about their model with reference to components than dyads in the face-to-face condition (U = 31.0, p = .03).

The pattern of results on surface processes shows that dyads in the face-to-face condition spent significantly more proportion of time than dyads in the chat condition on the following processes: evaluating with no reference to knowledge (U = 22.0, p = .01), quantifying with no reference to knowledge (U = 18.0, p = .005), and analyzing with no reference to knowledge (U = 25.5, p = .02).

Table 3. Mean proportion and standard deviations of time spent on deep and surface reasoning processes and the results of a Mann-Whitney U test for differences between the chat condition and the face-to-face condition

	Chat condition		*Face-to-Face condition*		*Mann-Whitney U*	
Reasoning processes	*M*	*SD*	*M*	*SD*	*U*	*p*
Deep processes						
Evaluating and reference to knowledge	0.62	1.10	1.10	1.20	37.0	.11
Explaining and reference to knowledge	0.23	0.48	0.00	0.00	49.5	.15
Quantifying and reference to knowledge	5.44	6.78	4.60	5.44	56.5	.79
Inductive reasoning and reference to knowledge	0.64	1.45	2.49	3.65	38.5	.10
Inductive reasoning and reference to components	7.29	24.18	2.21	3.70	31.0*	.03
Analyzing and reference to knowledge	1.20	1.80	1.72	2.24	49.0	.42
Surface processes						
Evaluating and no reference to knowledge	3.50	6.60	9.71	3.94	22.0*	.01
Quantifying and no reference to knowledge	3.10	5.83	8.00	4.96	18.0*	.005
Analyzing and no reference to knowledge	3.17	3.79	8.11	4.42	25.5*	.02

* $p < 0.05$.

3.3 Model quality and relation with modeling activities and reasoning processes

Model quality scores did not differ significantly between the chat condition (M = 6.97, SD = 1.41) and the face-to-face condition (M = 7.02, SD = 1.74; U = 58, p = .90).

To investigate whether particular types of modeling activities and reasoning processes are related to the quality of the model students constructed in both conditions, we calculated non-parametric Spearman rank correlations. Between-condition differences for these correlations were analyzed more closely by conducting the Fisher Z-test.

3.3.1 Concerning modeling activities

Within the chat condition significant positive correlations can be observed between the frequencies of dyads specifying quantities (r = .69, p = .02) and adding model elements (r = .67, p = .02) on the one hand and model quality score on the other (see Table 4).

The correlation between the frequency of dyads specifying quantities and model quality score (Z' = 2.37, p = .02) and the correlation between the frequency of students running models and model quality score (Z' = 2.39, p = .02) differ significantly. These differences between conditions mean that the more often dyads in the chat condition perform these modeling activities (i.e., specifying quantities and running models) the higher their model quality score will be, whereas the relation between those variables is significantly different for dyads in the face-to-face condition.

Table 4. Spearman rank correlations between frequency of modeling activities and model quality scores and the results of a Fisher Z-test for differences between the chat condition and the face-to-face condition

	Chat condition	*Face-to-face condition*	*Fisher Z*	
Modeling activities			*Z'*	*p*
Specifying quantities	.69*	-.34	2.37*	.02
Adding model elements	.67*	.16	1.30	.19
Deleting model elements	.20	.04	0.32	.75
Saving models	.55	-.15	1.54	.12
Opening saved models	.59	-.27	1.90	.06
Running models	.49	-.58	2.39*	.02

* $p < 0.05$.

3.3.2 Concerning reasoning processes

There is no correlation between the total amount of process-episodes and model quality score within the chat condition (r = -.01, p = .98), whereas the negative correlation found within the face-to-face condition is significant (r = - .67, *p* = .02). These correlations do not differ significantly (Z' = 1.60, *p* = .11).

The correlations for both conditions between the proportion of time dyads spent on processes (i.e., deep and surface processes) at the one hand and model quality score at the other are provided in Table 5. Within the chat condition, there is a significant positive correlation between the proportion of time spent on quantifying with reference to knowledge (i.e., deep process) and model quality score ($r = .63$, $p = .04$). One significant negative correlation is found within the face-to-face group between the proportion of time spent on quantifying with no reference to knowledge (i.e., surface process) and model quality score ($r = -.76$, $p = .007$).

The Fisher *z*-test reveals that the correlations between the proportion of time dyads spent on quantifying with no reference to knowledge (i.e., surface process) and model quality score differ significantly between the two conditions ($Z' = 2.28$, $p = .02$).

Table 5. Spearman rank correlations between proportion of time spent on deep and surface reasoning processes at the one hand and model quality scores at the other together with the results of a Fisher Z-test for differences between the chat condition and the face-to-face condition

	Chat condition	*Face-to-face condition*	*Fisher Z*	
Reasoning processes			*Z'*	*p*
Deep processes				
Evaluating and reference to knowledge	-.27	-.32	0.11	.92
Explaining and reference to knowledge	.36	.00	0.75	.45
Quantifying and reference to knowledge	.63*	.48	0.45	.65
Inductive reasoning and reference to knowledge	.15	.34	-0.40	.69
Inductive reasoning with reference to components	.49	-.32	1.74	.08
Analyzing and reference to knowledge	.50	-.12	1.35	.18
Surface processes				
Evaluating and no reference to knowledge	-.34	-.26	-0.16	.87
Quantifying and no reference to knowledge	.15	-.76**	2.28*	.02
Analyzing and no reference to knowledge	-.08	-.26	0.37	.71

* $p < 0.05$. ** $p < 0.01$.

4. CONCLUSION AND DISCUSSION

The present study addressed the impact of chat communication versus face-to-face communication on students' performance within a collaborative modeling task. We operationalized performance by measuring: a) students' activities during modeling, b) students' reasoning during modeling, and c) the quality of the students' models. Two hypotheses were considered: 1) Collaborative computer-based modeling is a complex task for students to be engaged in and requires the transmission of maximally rich information, as in face-to-face situations. Synchronous computer-mediated communication tools provide less information richness than the task requires. As a result, groups who communicate using chat will perform poorer on the modeling task compared to face-to-face groups. 2) Because of communication constraints, students who communicate through a chat tool compress their communication by being more task focused and more concise. Face-to-face exchanges may contain more nonessential information which may act as a distraction during modeling. Consequently, groups who communicate using chat will perform better compared to students who communicate face-to-face.

In general, results were more in line with the predictions of hypothesis 2. First, it was found that students in the chat condition saved and executed their models significantly less often compared to students in the face-to-face condition. Nonetheless, for the other modeling activities (i.e., specifying quantities, adding model elements, deleting model elements, and opening saved models) no significant differences were found. Second, the amount of process-episodes was significantly lower for the chat condition compared to the face-to-face condition, a finding which was also predicted by hypothesis 1. Third, students in the chat condition spent significantly more time on inductive reasoning with reference to model components (i.e., a deep process) compared to students in the face-to-face condition. For the remaining deep processes (i.e., evaluating and reference to knowledge, explaining and reference to model components, quantifying and reference to knowledge, inductive reasoning and reference to knowledge, and analyzing and reference to knowledge) no significant differences were found between the conditions. Finally, students in the face-to-face condition spent significantly more time on surface processes (i.e., evaluating and no reference to knowledge, quantifying and no reference to knowledge, and analyzing and no reference to knowledge) than students in the chat condition. Although no significant difference was found between conditions on model quality, results conflict with the main prediction of hypothesis 1, that chat communication will negatively affect modeling performance compared to face-to-face communication.

The expectations derived from hypothesis 2, that students in the chat condition will compress their communication resulting in more efficient modeling, whereas students who communicate face-to-face will exchange more irrelevant information which may hinder modeling performance, was further supported in the additional analyses we performed. These analyses revealed that, within the chat condition, a significant positive relation was found between the proportion of time students spent on episodes in which they quantified and referred to knowledge (i.e., a deep process) and model quality score. Also, it was found that students who communicated face-to-face and who spent more time on quantifying without referring to knowledge (i.e., a surface process) scored significantly less on model quality. Furthermore, a significant negative relation was found between the amount of process-episodes expressed by students in the face-to-face condition and model quality, indicating the negative impact (overly) rich communication may have on modeling performance.

Correlations calculated between students' modeling activities and model quality showed that the more often students in the chat condition specified quantities or added model elements, the higher they scored on model quality. In addition, the correlations found between specifying quantities and model quality and the correlations between running models and model quality differed significantly between conditions. Thus, the more often students in the chat condition performed these activities the higher they scored on model quality, whereas this was not found for students who communicated face-to-face. As Suthers et al. (2003) argue, chat groups rely more on the external model representation to compensate for the absence of face-to-face modalities and that communicative work also happens through the components present in the representation. This argument was supported in the present study, since students in the chat condition spent a significantly greater amount of time on episodes in which they inductively reasoned with referring to model components (i.e., a deep process) compared to students who communicated face-to-face.

Thus, communication constraints in the chat condition pressured students to compress their communication during modeling and to rely more on the modeling representation, ultimately leading to more efficient modeling. We expect that if the available time had been prolonged for both groups, the chat condition might have outperformed the face-to-face condition, since, in general, it is found that students who communicate by means of chat need more time to complete a given task, because of typing requirements (Baltes et al., 2002; Bordia, 1997; Carey & Kacmar, 1997).

Our results are in agreement with the findings from Barile & Durso (2002), Basque & Pudelko (2004), Condon & Cech (1996a; 1996b), Dennis (2003), Jonassen & Kwon (2001), and Newlands et al. (2003). Nevertheless, the reported findings seem to contrast with studies that show that face-to-face groups outperform chat groups (e.g., Carey & Kacmar, 1997; Chung et al., 2000; Hollingshead, et al., 1993; Mennecke et al., 2000; Straus, 1997; Suthers et al., 2003; Van der Meijden & Veenman, 2005). Two factors which may explain this discrepancy in findings are: a) that the groups in the present study consisted of dyads (instead of groups consisting of more than two students) and b) that the students used in our study had a great deal of experience with using chat as a communication mode (cf. Baltes et al., 2002; Hollingshead et al., 1993; Van der Meijden & Veenman, 2005). It may be that when groups consist of more than two students, the modeling task necessitates for more information richness, since it becomes more difficult for students to keep track of the discussion and of each other when communicating using chat. And secondly, if students are less skilled in typing and using chat software they may be less concise and less task focused than the students in our study, since they first have to learn to use the chat tool.

Although the collaborative computer-based modeling task we used in our study can be classified as a complex task, students in the chat condition compensated for this by being more selective and more thorough in their interactions than students in the face-to-face condition resulting in more efficient modeling in the former group. However, an important consideration to take into account is that the reported findings are applicable only within the context of the present study. It remains an open issue in how far these findings can be generalized to other task settings.

CHAPTER 5

MOTIVATION AND PERFORMANCE WITHIN A COMPUTER-BASED MODELING TASK: RELATIONS BETWEEN STUDENTS' ACHIEVEMENT GOAL ORIENTATION, SELF-EFFICACY, REASONING AND MODEL QUALITY*

Purpose of the present study was to test a conceptual model of relations among achievement goal orientation, self-efficacy, reasoning and achievement of students working within a particular collaborative task context. The task involved a collaborative computer-based modeling task. In order to test the model, group measures of mastery-approach goal orientation, performance-avoidance goal orientation, self-efficacy and achievement (i.e., model quality) were employed. Students' reasoning was assessed using an online log file measure. As predicted, both self-efficacy and mastery-approach goal orientation had a significant positive effect on model quality, which was mediated through students' use of deep processes. No significant relations could be found between performance-avoidance goal orientation and surface reasoning and between surface reasoning and model quality. Results are discussed with respect to general theoretical implications and lead to suggestions for the design of appropriate scaffolds.

1. INTRODUCTION

Current models of self-regulated learning integrate motivational and cognitive elements of learning, showing how achievement goals and learning expectancies influence students' use of cognitive processes (Covington, 2000; Nicholls, Patashnick, Cheung, Thorkildsen, & Lauer, 1989; Pintrich, 2003; Schunk, 2005; Zimmerman & Schunk, 2001). The basic assumption of these models is that the achievement motives and the intentions that guide students' academic behavior determine to a great extent the types of cognitive processes they employ in various learning situations (Harackiewicz, Barron & Elliot, 1998; Meece, Blumenfeld & Hoyle, 1988). The learning outcome (i.e., achieve-

* Sins, P.H.M., Van Joolingen, W.R., Savelsbergh, E.R., & Van Hout-Wolters, B.H.A.M. (submitted). *Motivation and performance within a computer-based modeling task: Relations between students' achievement goal orientation, self-efficacy, reasoning, and model quality.*

ment) is dependent on how deep students process information (Craik & Lockhart, 1972; Entwistle, 1979; 1988; Graham & Golan, 1991).

The majority of studies that have examined the consequences of students' achievement goal orientation and self-efficacy on their cognitive processing of the learning material, concern the *self-reported* use of cognitive processes of *individual* students over a *complete course or curriculum* (e.g., Greene & Miller, 1996; Nolen, 1988; Pintrich, 2000a). However, a great deal of learning takes place within collaborative task contexts, in which students construct knowledge through mutual communication and shared use of representations (Pintrich, Conley, & Kempler, 2003). Goal of the present study is to investigate the relations among students' achievement goal orientation, self-efficacy, reasoning and achievement within a particular collaborative task setting. The variables will be measured at the level of collaborating students (i.e., dyads) rather than at the individual level. In addition, instead of using self-report questionnaires which are traditionally employed within the field of achievement motivation, students' reasoning will be assessed by means of an online measure, based on inter-student communication.

1.1 Achievement goal orientation and self-efficacy as predictors of cognitive processing

Many authors in the domain of motivation research have argued that the type and the level of motivation influences students' employment of particular cognitive processes in learning situations (e.g., Atkinson, 1964; Covington, 2000; Graham & Golan, 1991; Jagacinsky, 1992; Meece et al., 1988; Nicholls et al., 1989; Pintrich & Schrauben, 1992; Wolters, 1996). Two motivational factors that are presumed to be important determinants of students' cognitive processing are: (a) achievement goal orientation and (b) self-efficacy.

Broadly defined, achievement goal orientation reflects the reasons and the purposes of students to engage in achievement tasks. Two distinct types of achievement goal orientations are traditionally distinguished (e.g., Butler, 1991; Dweck & Leggett, 1988; Elliot, 1999; Elliot & Dweck, 1988; Nicholls et al., 1989): *mastery goal orientation* and *performance goal orientation*. Mastery goal orientation involves the belief that effort leads to improvement in performance and that competence is malleable. Students who are mastery goal oriented focus on the development of new skills and knowledge, try to elaborate on the task they are given and attempt to reach their own learning goals. Performance goal orientation, in contrast, involves the belief that competence can be demonstrated by performing better compared to peers. Students who are performance goal oriented tend to focus on attaining normative learning goals.

Recent research on achievement goal orientation has questioned the utility and validity of this two-goal model and proposes instead that besides the mastery-performance distinction, another dimension to consider is whether achievement goal orientations lead students to approach or avoid a task (Elliot, 1997, 1999; Elliot & Church, 1997; Harackiewicz, Barron, & Elliot, 1998; Harackiewicz, Barron, Pintrich, Elliot & Thrash, 2002; Pintrich, 2000b). The performance goal orientation construct is bifurcated into a performance-approach goal orientation and a performance-avoidance goal orientation. Students who are performance-approach oriented are focused on achieving higher levels compared to their peers and aim to demonstrate high ability, whereas students who are performance-avoidance oriented are concerned with avoiding failure and with avoiding the demonstration of low ability. Following this logic, mastery goal orientation can also be separated into approach and avoidance goal orientations. Whereas a mastery-approach goal orientation involves striving to develop one's skills and abilities, to advance one's learning, to understand the material, or to complete a task, a mastery-avoidance goal orientation entails focusing on avoiding misunderstandings or not learning the material.

An important issue to consider at this point is that it is ineffective for students to be striving to master a task, if they are less convinced that they have the necessary ability and competence to do so. Thus, the influential role of self-efficacy on task performance must be taken into account. Self-efficacy has been defined as students' belief regarding their performance capabilities in a particular domain (Bandura, 1982; 1986).

Achievement goal theorists hypothesize that students who are mastery-approach goal oriented attempt to gain rich insight in the given learning material and will therefore employ more *deep* cognitive processes to increase their comprehension (e.g., Dweck, 1985; Graham & Golan, 1991; Nicholls et al., 1989; Pintrich & DeGroot, 1990). Deep cognitive processing, as described in the work of Marton & Säljö (1976; 1997), Ramsden (1992), and Entwistle (1981; 1988; 2001), involves active learning processes, such as relating ideas, looking for patterns and principles and attempting to integrate new information with prior knowledge and experience. Surface cognitive processing, in contrast, entails processes without much reflecting and involves treating the learning material as more or less unrelated bits of information. Surface processing does not implicate elaboration of the learning material and leads to more restricted learning processes. Because mastery-approach goal oriented students tend to attribute learning success to invested effort and attempt to understand the learning material, they may be more likely to employ and value processes that stress understanding, even if these processes require more effort then less effective processes. In addition, self-efficacy theory and other theories on self-perception state that self-efficacy beliefs are positively related to the use of deep processes (Bandura, 1986; Pintrich & Schrauben, 1992; Schunk, 1991).

Numerous studies have demonstrated that mastery-approach goal orientation and self-efficacy are positive predictors of reported use of deep processes (e.g., Ames & Archer, 1988; Al-Emadi, 2001; Bruinsma, 2004; Butler, 1991; Dupeyrat & Mariné, 2005; Elliot, McGregor, & Gable, 1999; Elliot & McGregor, 2001; Fisher & Ford, 1998; Ford, Smith, Weissbein, Gully & Salas, 1998; Garcia, McCann, Turner & Roska, 1998; Graham & Golan, 1991; Greene & Miller, 1996; McWhaw & Abrami, 2001; Meece, Blumenfeld & Hoyle, 1988; Meece et al., 1988; Meece & Holt, 1993; Miller, Greene, Montalvo, Ravindran, & Nichols, 1996; Nolen, 1988; Pintrich, 200a; Pintrich & Garcia, 1991; Pintrich & Schrauben, 1993; Schraw, Horn, Thorndike-Christ & Bruning, 1995; Valle, Cabanach, Nunez, Gonzalez-Pienda, Rodriguez, & Pineiro, 2003; Vermetten, Lodewijks & Vermunt, 2001; Wolters, 1996; Wolters & Yu, 1996). In addition, some studies found that both self-efficacy and mastery-approach goal orientation are indirectly related to achievement via a direct relation with the employment of deep processing strategies (e.g., Elliot et al., 1999; Greene & Miller, 1996; Pintrich & DeGroot, 1990). Moreover, Greene and Miller (1996), Meece et al. (1988), and Vrugt et al. (1999; 2002) found an additional positive relation between self-efficacy and mastery-approach goal orientation. Performance-avoidance oriented students, in contrast, focus on the avoidance of demonstration of incompetence relative to peers. As a result, these students resort to the employment of surface reasoning processes, which is linked to decreases in achievement (Al-Emadi, 2001; Elliot et al., 1999; Elliot & McGregor, 2001; Greene & Miller, 1996; Meece et al., 1988; Midgley, Kaplan & Middleton, 2001; Nolen, 1988; Pintrich & Garcia, 1991; Pintrich & Schrauben, 1992; Schraw et al., 1995).

For performance-approach and mastery-avoidance goals, more variable and complex patterns in terms of deep and surface processing can be expected, when compared to mastery-approach goals and performance-avoidance goals (Elliot, 1999). Mastery-approach goal orientation and performance-avoidance goal orientation typically represent pure approach and pure avoidance motivation respectively (Elliot, 1997). In contrast, mastery-avoidance goal orientation and performance-approach goal orientation may involve both approach and avoidance motivational concerns (i.e., respectively need for achievement and fear of failure; see Elliot & Church, 1997). When, for instance, performance-approach goals are the result of a need for achievement (i.e., congruent), the pursuit of these goals may prompt the use of deep processes (Harackiewicz et al., 1998; 2002). When performance–approach goals are incongruent with their underlying motivational foundation, the pursuit of these goals represents approach in order to avoid something aversive. This may lead to surface processing of the learning material. Like performance-approach goals, predictions for mastery-avoidance goals are somewhat difficult to generate, since the two components of mastery-avoidance goal orientation are likely to evoke a rather divergent set of processes (Elliot, 1999). That is, the mastery component of this type of goal orientation may facilitate deep processing (Ames, 1992), whereas the avoidance component may impel surface processing (Elliot, 1997). It is, thus, difficult to predict the exact nature of the processes that will be evoked by these two types of achievement goal orientation, as this is dependent on their motivational foundation.

Although, empirical data regarding mastery-avoidance goals are not yet available (see Pintrich et al., 2003), the association between performance-approach goals and students' cognitive processing of the learning material has indeed been shown to be contradictory. Some studies found that performance-approach goals are associated with surface processing (Al-Emadi, 2001; Dupeyrat & Mariné, 2005; Elliot & McGregor, 2001; Elliot et al., 1999; Greene & Miller, 1996; Middleton & Midgley, 1997; Nolen, 1988 Vermetten, et al., 2001), whereas other studies found a positive association between performance-approach goals and deep processing (Bouffard, Boisvet, Vezeau & Larouche, 1995; Meece et al., 1988; Pintrich, 200a; Wolters, 1996; Wolters & Yu, 1996). Furthermore, some studies found that performance-approach goals are positively related to high performance outcomes (Barron & Harackiewicz, 2001; Elliot & Church, 1997; Harackiewicz, Barron, Elliot, Carter, & Lehto, 1997; Wolters, 1996). As the goal of our study is to investigate whether the consistent relations found in literature reproduce at the level of actually observed processes, performance-approach goal orientation and mastery-avoidance goal orientation were excluded. The present study was framed within the context of a computer-based scientific modeling task.

1.2 Computer-based modeling

Computer-based models are executable external representations of the behavior of complex scientific phenomena, such as ecosystems, water management, and weather (Bliss, 1994; Penner, 2001; Spector, 2000; Stratford, 1997). The act of modeling is the activity in which models are constructed, evaluated and revised with the help of a computer-based modeling tool, such as STELLA (Steed, 1994) and Powersim (Byrknes & Myrtveit, 1997). The construction of models is particularly well suited to provide the basis for meaningful learning experiences (e.g., Doerr, 1996; Gilbert et al., 1998; Hestenes, 1997; Jackson et al., 1996; Raghaven et al., 1998; Schecker, 1993; Tinker, 1993). Models focus on the continuous relations among variables that are part of a phenomenon and provide a platform for understanding how these variables interact (e.g., Mandinach, 1988; Steed, 1994). Computer-based modeling tools, thus, enable students to express and manipulate their mental representation of a phenomenon, which supports the reorganization and refinement of their conceptual understanding (e.g., Jonassen et al., 2005; White & Frederiksen, 1998; Wild, 1996). Moreover, since modeling tools help students to externalize their ideas, they are accessible to criticism and discussion, which is an important prerequisite for collaborative learning (Devi et al., 1996; Hogan & Thomas, 2001).

1.3 The current study

The purpose of the present study is to test a conceptual model of relations among motivation, reasoning and achievement (i.e., model quality) of students working within a computer-based modeling task (see Figure 1). More specifically, we investigated whether students' achievement goal orientation and self-efficacy influences their reasoning during modeling, and whether the relation between motivation and model quality is mediated by students' reasoning

In contrast to most studies in the field of motivation, we did not assess students' processing based on self-report measures, but rather we base our assessments on the process observations of students' reasoning as described in the previous chapters. This is first because, the validity of self-report on these issues has become under doubt (Hogan & Thomas, 2001; Pintrich, 2000a), and furthermore because in the present context, where it comes to interpreting the detailed processes of students in a collaborative setting, more fine-grained measures are needed and the interaction between the collaborating partners must be taken into account.

As the modeling task required students to work in dyads, students' reasoning (i.e., deep and surface reasoning) can be naturally measured using this online measure that includes inter-student communication through a chat. In addition, achievement can be operationalized as the quality of model dyads constructed. This also requires measurement of all other variables on the dyad level, in order to test the conceptual model in Figure 1. This means that we have to relate individual student characteristics, such as self-efficacy, to the same characteristics of a dyad.

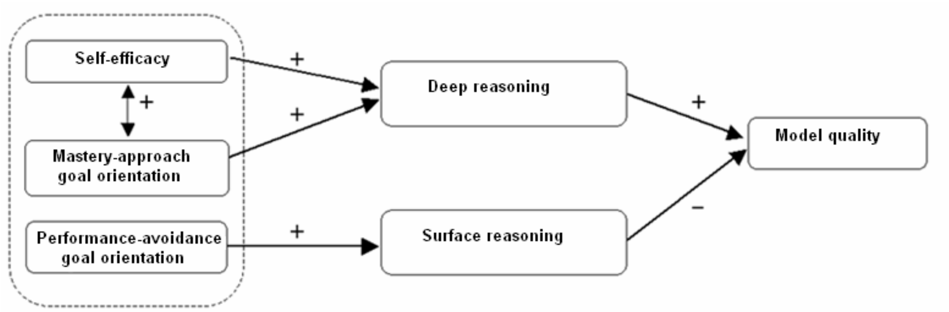


Figure 1. Conceptual model indicating the hypothesized relations between students' motivation (i.e., mastery-approach goal orientation, performance-avoidance goal orientation and self-efficacy), reasoning (i.e., deep and surface reasoning) and model quality. Signs indicate the direction of the hypothesized relation.

2. METHOD

2.1 Participants

Sixty students (i.e., thirty dyads) from eleventh-grade pre-university education, with a major in science participated in our study. Participants had no prior experience with a modeling task as the one used in this study, therefore it is safe to assume they had no task-specific prior knowledge. Students' age ranged between 16-18 years. Participants were awarded € 20 for their participation.

As dyad composition was not a variable in our model, we preferred a heterogeneous group composition, since in previous studies students with different school grades had been generally more successful working together than homogeneous groups (e.g., Gijlers & de Jong, 2005; Webb 1991; Webb et al., 2001). The reason is that higher achieving students can learn from giving explanations, whereas the lower achieving student can learn from these explanations given (Hooper & Hannafin, 1991; Webb & Farivar, 1994). However, the difference in level between students should not be too large. As task and domain-specific prior knowledge was absent, we used the students' average school grade in science as a measure for group composition. The mean average grade of all students was 7.10 on a scale from 1 to 10, with a standard deviation of 1.03. In order to assure moderately heterogeneous dyads, the group of participants was divided into two equal groups. One group consisted of the top 25% as well as the bottom 25% in average grade for science. The other group consisted of the remaining 50%. Dyads were composed by letting students choose a partner from the other group. This assured a moderate difference between partners in science ability, as well as pairs who had chosen each other to work with. This procedure meant that half of the dyads were low-middle dyads in terms of average grade, whereas the other half was middle-high. Given the low variance in average school grade, as well as the fact that all students were new to the task and the domain, we did not expect any differences from this division.

2.2 Material

Students performed the modeling task within the computer-based learning environment Co-Lab (Van Joolingen et al., 2005). In Co-Lab students can collaborate online using a synchronous chat facility on inquiry assignments for the science courses.

Participants were asked to extend a simple pre-build model that could give an explanation and prediction of the temperature on earth. The task was simplified to some extent, since the earth in this task was represented by a irradiated black sphere (see Appendix D for the assignment). Because participants had no prior experience with modeling, a completely open modeling task was assumed to be too complex for them to be successful within the time constraints of the modeling task. Therefore, participants were given a model skeleton as a starting point. Such a model revision task enables students to concentrate on trying to comprehend and improve a model without having to start from scratch. Students constructed their models in the model editor tool of Co-Lab (see Figure 2).

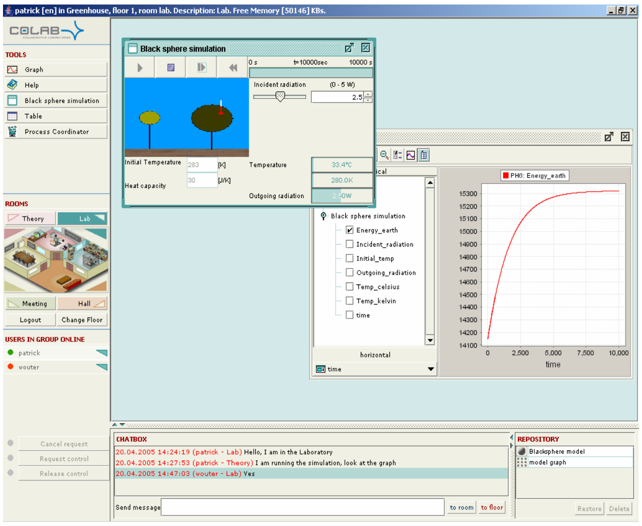


Figure 2. Screenshot of the simulation of the temperature of an irradiated black sphere. Results of the experiments are provided in a graph or table.

The model editor in Co-Lab uses five model building blocks characteristic for system dynamics modeling: Stocks, rates, auxiliaries, constants, and connectors. Stocks represent a quantity that can increase or decrease from some starting value. A rate connected to a stock decides how quickly the quantity in the stock will change. Quantities can be represented either as constants (i.e., fixed values), or as auxiliaries (i.e., calculated from other quantities). Finally, connectors indicate dependencies between individual model elements. To insert a modeling element, students can drag and drop the icons on the screen they think are relevant for the phenomenon being modeled, creating a qualitative diagram of the phenomenon. After creating this diagram, students have to quantify these elements by entering values and formulas. Once the model is quantified it can be executed. When students run their model, the model editor automatically generates the differential equations required to perform calculations. The results of model runs over time can be displayed as graphs or tables. In order to test their models, students can compare the output of their model with data they can obtain from running a simulation of a black sphere (see Figure 3 for a screenshot of the simulation). Consequently, students may revise their model on the basis of the testing outcomes.

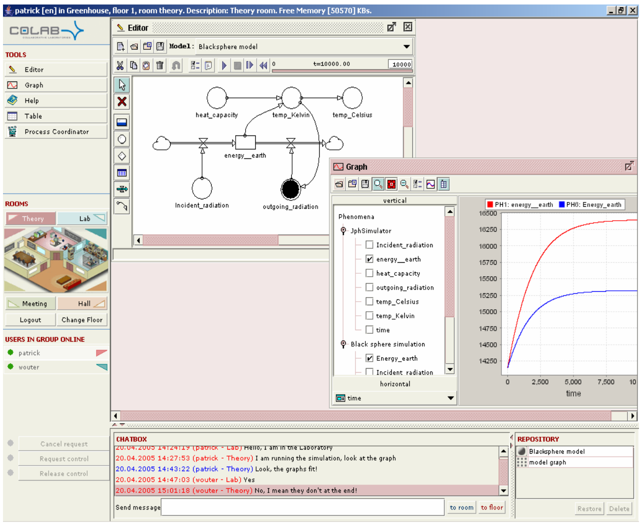


Figure 3. Screenshot of the model editor in Co-Lab. In this particular example an accurate model diagram is provided of the black sphere simulation. This model shows that the energy of the earth is influenced by the incident radiation from the sun (i.e., energy inflow) and the outgoing radiation (i.e., energy outflow). The outflow is influenced by the temperature of the earth. Finally, the temperature is influenced by the energy of the earth and the heat capacity of the earth.

2.3 Instruments

2.3.1 Achievement goal orientation

Achievement goal orientation (i.e., mastery-approach goal orientation and performance-avoidance goal orientation) was measured on the individual level as well as on the dyad-level, employing questionnaires. For measuring mastery-approach goal orientation and performance-avoidance goal orientation of the individual student, items from corresponding subscales of the Goal-Orientation Questionnaire of Seegers and Boeckaerts (1993) were adapted. The questionnaire of Seegers and Boeckaerts (1993) was originally based on the one that Nicholls et al. (1989) describe and which has also been adapted and used in the studies conducted by Duda and Nicholls (1992), Nolen (1988), and Vrugt et al. (1999). For the present study, items from the Goal-Orientation Questionnaire were contextualized by rephrasing them into statements directed at assessing students' goal orientation regarding physics. The first scale consists of five items formulated to express mastery-approach goal orientation (e.g., "I like to work hard on a physics task"), and the second scale involves six items for performance-avoidance orientation (e.g., "I want to avoid doing poorly in physics class"). Students indicated on a 4-point scale to what extent they usually react in the described manner. Response alternatives ranged from "never" to "always". Principal-components analysis with varimax rotation supported the presence of two factors, with items loading on its designated factor. The two primary factors accounted for 71% of the total variance. Coefficient alphas were calculated for the two subscales that are focus of this study. Alpha was .80 for mastery-approach goal orientation and alpha was .79 for performance-avoidance goal orientation.

To measure mastery-approach goal orientation and performance-avoidance goal orientation on the dyad-level we employed the same questionnaire, with items being rephrased to fit the group, replacing "I" with "We" . Dyads answered this questionnaire together and were prompted to discuss the question before answering it. They needed to agree on the answer. Principal-components analysis with varimax rotation showed that the two factor solution accounted for 69% of the total variance. Coefficient alphas for the two subscales that are focus of this study are .69 and .72, for mastery-approach goal orientation and for performance-avoidance goal orientation respectively.

2.3.2 Self-efficacy

As with achievement goal orientation, self-efficacy was also measured on both the individual level and the dyad-level. Self-efficacy was captured with the translated General Self-Efficacy questionnaire of Schwarzer (1992). Students were asked to indicate how adequate they estimated their ability with respect to the modeling task (e.g., "I can solve most problems if I invest the necessary effort"). The scale consisted of ten statements and students were asked to indicate on a 4-point scale to what extent they agreed with the statement. Response alternatives ranged from "Not at all true" to "Exactly true". Group self-efficacy was measured in a similar fashion as with the group measures of achievement goal orientation. Internal consistencies were found to be .70 and .68 for the individual self-efficacy questionnaire and the group self-efficacy questionnaire, respectively.

2.3.3 Reasoning processes

Students' reasoning was measured by analyzing the inter-student chat, taken from the log files of the students' sessions, employing the protocol analysis scheme of Sins et al. (2005; see Chapter 2). The chat logs were scored employing two categories that were taken from the scheme of Sins et al. (2005)[12]: a) students' reasoning processes during modeling and b) type of reference (i.e., argumentation) made by students during reason-

---

12 In addition to these two categories, the protocol analysis scheme of Sins et al. (2005; see Chapter 2) also includes the category: 'topic focus of students' reasoning'. We did not include this code in the present analyses, since reasoning processes coupled with the type of reference (i.e., argumentation) students make during process-episodes provide sufficient information concerning students' level of processing.

ing (see Appendix A for coding scheme). Reasoning processes like analyzing or explaining may be considered to be cognitive processes and mostly involve several turns by both partners in a dyad. Therefore, the unit of analysis is the process-episode level, an episode being a period of coherent continuous talk on a single issue, rather than single utterances (cf. Chi, 1997).

Reasoning episodes in which students are elaborating on the modeling task and connect to knowledge they have available, either gained from the task at hand or as prior knowledge, were designated as deep reasoning. Episodes in which students employ unelaborated reasoning processes without referring to available knowledge were labeled as surface reasoning. Only episodes that could be clearly marked as either *deep* or *surface* reasoning were counted in this analysis, excluding all others. The following specific codes were operationalized as indications of deep reasoning:

- Evaluating and reference to knowledge
- Explaining and reference to knowledge
- Quantifying and reference to knowledge
- Inductive reasoning and reference to model components
- Inductive reasoning and reference to knowledge
- Analyzing and reference to knowledge

The following codes indicated surface reasoning:

- Evaluating and no reference to knowledge
- Quantifying and no reference to knowledge
- Analyzing and no reference to knowledge

Chat utterances were segmented into episodes and were scored with the help of our coding scheme. Because protocols differ in length, and because protocol episodes are of different length as well, frequencies of each code were converted to proportions of total time for each dyad, and analyses are based on these proportions. The total proportion of time dyads spent on deep reasoning was calculated by summing up the proportions of time for the six codes indicating deep reasoning processes. The same procedure was performed for surface reasoning.

The use of deep versus surface processes was also individually measured using a *self-report questionnaire* in order to investigate the correspondence between results from such questionnaires and the online log file measure. Items from this questionnaire were based on the available literature (e.g., Greene & Miller, 1996; Entwistle, 1981; 1988; Marton & Säljö, 1997; Nolen, 1988; Pintrich & Garcia, 1991; Valle et al., 2003) and based on codes from the protocol analysis scheme of Sins et al. (2005; see Chapter 2). Thirteen items were constructed which were specifically tailored for measuring students' reasoning during modeling. The deep reasoning subscale consisted of seven statements (e.g., "I used my own knowledge during modeling") and the surface reasoning subscale consisted of six statements (e.g., "I tried to improve our model by changing values most of the time"). Students were asked to indicate the extent to which they agreed with the statements on a 5-point scale. Response alternatives ranged from "Totally agree" to "Totally disagree". Principal-components analysis with varimax rotation supported the presence of two factors, with items loading on its designated factor. The two primary factors accounted for 42% of the total variance. Coefficient alpha was .46 for surface reasoning and alpha was .59 for deep reasoning.

2.3.4 Model quality

The models dyads constructed were judged with the help of a model scoring template. The students' final model was awarded with points for each variable which name and specification was correct. In addition, students were awarded with a point for correct links between variables and with a point for correct specifications for these relations. Finally, for each incorrect relation between quantities a point was subtracted from the total score.

2.4 Procedure

The study consisted of two sessions each of about two and a half hours on two separate days. In the first session students were given a plenary introduction to the Co-Lab environment and were provided with an individual modeling tutorial. Subsequently, students were asked to complete the individual self-efficacy questionnaire and the individual achievement goal orientation questionnaire. Afterwards, dyads were composed and a training task was given, in which they were asked to collaboratively extend a pre-build model involving the inflow and outflow of water from a water tank. Students worked with a simulation of a water tank and could investigate background information about this task using the Co-Lab help tool. On the basis of data obtained from this simulation and on the basis of information gathered, students could extend and revise their model. At last, dyads were asked to complete the group measures of self-efficacy and achievement goal orientation.

In the second session dyads were presented with a modeling task in which they were asked to extend a given model. Students' goal was to extend the model so that it could give an explanation and prediction of the temperature on earth (see Appendix A for the modeling task). Co-Lab provided support for students in order to complete this modeling task: Students could consult background information and could work with a simulation of a black sphere. Students worked for two hours on the modeling task. When working within Co-Lab members of a dyad each worked on one computer. The Co-Lab environment was shared between students and students communicated through a chat channel. Finally, the individual self-report questionnaire of students' reasoning was administered and completed.

3. RESULTS

In order to check whether there were any differences between low-middle and middle-high dyads in terms of performance, t-tests were done on all measures with group composition as independent variable. No significant differences were found, indicating that we can treat the sample as one group.

3.1 *Validity of group measures for mastery-approach goal orientation, performance-avoidance goal orientation and self-efficacy*

While students worked in dyads on the modeling task, mastery-approach goal orientation, performance-avoidance goal orientation and self-efficacy were measured on the dyad-level. Table 1 shows the means, standard deviations, and ranges for these variables.

Table 1.Means, standard deviations and minimum and maximum scores for the motivation variables

	Mean	*SD*	*Observed range*
Group mastery-approach goal orientation	3.02	0.43	1.00 – 4.00
Group performance -avoidance goal orientation	1.41	0.36	1.00 – 4.00
Group Self-efficacy	3.07	0.29	1.00 – 4.00

Since, to our knowledge, no studies on group measures of these constructs have been conducted, we attempted to obtain some insight into the validity of these measures. We averaged for each dyad the total individual scores on the self-efficacy questionnaire, mastery-approach goal orientation subscale, and performance-avoidance goal orientation subscale and correlated these figures with the corresponding scores of the dyads on the group measures. The zero-order correlations for mastery-approach goal orientation ($r = .56$; $p < .01$), performance-avoidance orientation ($r = .77$; $p < .01$) and self-efficacy ($r = .68$; $p < .01$) are significant and positive. Also, the averaged individual measure for self-efficacy per dyad and the group measure for mastery-approach goal orientation are significantly related ($r = .38$; $p < .05$).

3.1.1 *Correspondence between the online log file measure and the self-report measure of reasoning processes*

The total individual scores on the subscales deep and surface reasoning of the self-report questionnaire were averaged per dyad and correlated with the group scores for deep and surface reasoning obtained from protocol analysis. The correlations for deep reasoning ($r = .18$; $p = .32$) and surface reasoning ($r = .06$; $p = .74$) are low and not significant.

Table 2 shows the proportions of time dyads spent on deep versus surface reasoning, which were obtained from the protocol analysis. This table shows that a small proportion of time was spent on either surface reasoning ($M = 12.93\%$) or on deep reasoning ($M = 15.84\%$). Dyads spent the remaining time on talking about modeling actions ($M = 43\%$), on reading the learning material in Co-Lab ($M = 8.58\%$), on off-task communication ($M = 7.69\%$), and on other processes that did not fall under our conceptualization of deep versus surface reasoning ($M = 11.96\%$).

Table 2.Percentage of time spent on deep and surface reasoning

Reasoning processes	*Percentage of total time*
Deep processes	
Evaluating and reference to knowledge	0.59
Explaining and reference to knowledge	0.39
Quantifying and reference to knowledge	4.45
Inductive reasoning and reference to knowledge	5.14
Inductive reasoning and reference to components	3.12
Analyzing and reference to knowledge	2.15
Surface processes	
Evaluating and no reference to knowledge	3.99
Quantifying and no reference to knowledge	5.60
Analyzing and no reference to knowledge	3.34

3.1.2 *Testing the conceptual model*

We used the group measures of all variables and the online log file measure of students' reasoning for testing the conceptual model. Relations between the variables of our conceptual model (see Figure 1) were first examined with Pearson product-moment correlations between variables (see Table 3). The relation between self-efficacy and mastery-approach goal orientation is significant. In addition, both self-efficacy and mastery-approach goal orientation are significantly correlated with deep reasoning. The correlation between performance-avoidance goal orientation and surface reasoning is not significant. Deep reasoning is significantly positive related to model quality on the modeling task. Also, the correlations between mastery-approach goal orientation and model quality and between self-efficacy and model quality are significant.

Table 3. Correlations among motivational variables, deep and surface reasoning, and model quality

	1.	*2.*	*3.*	*4.*	*5.*
1. Group mastery-approach goal	–				
2. Group performance-avoidance goal	.03	–			
3. Group Self-efficacy	.59**	-.03	–		
4. Deep reasoning	.50**	-.10	.48**	–	
5. Surface reasoning	-.09	.18	-.16	-.02	–
6. Model quality	.38*	-.19	.35*	.46**	-.29

** $p < 0.01$.* $p < 0.05$.

Second, a series of multiple hierarchical regression analyses were performed (cf. Dupeyrat & Mariné, 2005; Greene & Miller, 1996). Variables were entered into the regression equation based on their temporal sequencing in the conceptual model. Each variable was regressed on the variables that had causal paths leading to it. In the first set of regression analyses deep reasoning and surface reasoning were the dependent variables. The predictors for each of these equations were self-efficacy, mastery-approach goal orientation, and performance-avoidance orientation. The second set of analyses investigated the effects of the motivational variables and students level of reasoning on model quality. The dependent variable, in these analyses, was model quality with the three motivational variables entered into the regression equation on the first step, and the two measures for students' reasoning entered on the second step. The results of these analyses are presented in Table 4.

Table 4. Multiple regression outcomes

Dependent variable	*Predictor*	R^2	ΔR^2	*β on step*	*Final β*
Deep reasoning	Mastery-approach goal orientation	.248**	.248**	.498**	.328*
	Self-efficacy	.354**	.106*	.288*	.288*
Surface reasoning					
Model quality	Mastery-approach goal orientation	.144*	.144*	.380*	.130
	Deep reasoning	.417**	.273**	.334*	.334*

*** $p < 0.01$. * $p < 0.05$.*

The first set of regression analyses supports the hypothesized positive associations between mastery-approach goal orientation and self-efficacy on the one hand and deep reasoning on the other. None of the motivational variables significantly predicts surface reasoning. The second set of regression analyses, with model quality as dependent variable, show that mastery-approach goal orientation and deep reasoning are significant predictors of model quality. However, the positive influence of mastery-approach goal orientation on model quality is not significant after controlling for the mediating influence of deep reasoning. The resulting path model is presented in Figure 4. For surface reasoning, no such path from performance-avoidance goal orientation can be found. However, there is a correlation between surface reasoning and model quality ($r = -.29$, $p = .12$) which is not significant, which may be caused by the relatively low power of the sample.

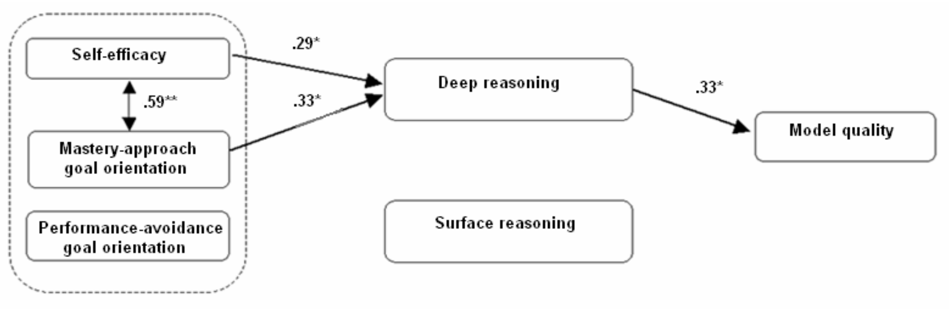


Figure 4. Results of the path analysis.

4. DISCUSSION

In the present study we tested a conceptual model linking achievement goal orientation, self-efficacy, reasoning and model quality of dyads working on a computer-based modeling task. In support of our conceptual model, we found that mastery-approach goal orientation and self-efficacy were both positively related to model quality and that these relations were mediated by dyads' employment of deep reasoning processes. However, in contrast to our predictions, the path model in Figure 4 shows that performance-avoidance goal orientation was not significantly related to dyads' use of surface processes. In addition, no significant relation was found between surface reasoning and model quality.

Our conceptual model was based on findings from the available achievement motivation literature (e.g., Covington, 2000; Dupeyrat & Mariné, 2005; Elliot et al., 1999; Greene & Miller, 1996; Middleton & Midgley, 1997). Traditionally, studies conducted within this field of research focus on the self-reported learning of individual students over a whole course. We tested our conceptual model for students who worked in dyads within a specific task context. Therefore, group measures of the variables in our model were employed. In addition, we used an online log file measure to capture collaborating students' reasoning processes.

An indication of the validity of the group measures for mastery-approach goal orientation, performance-avoidance goal orientation and self-efficacy was reflected in the finding that these measures were significantly positively related to the corresponding scores on the individual questionnaires, aggregated per dyad. This correlation shows that the measures on the individual and group levels are similar. In contrast, it was found that the group scores on the self-report questionnaire of reasoning were not significantly related to scores obtained from the online log file measure. If we accept that the log file measure is closer to actual observation than the post-hoc self report (Pintrich, 2000a), this result implies that post-hoc administered self-report measures do not indicate students' factual behavior during a given task. However, this finding should be taken with care, since the self-report questionnaire was low on reliability. We found that deep reasoning positively contributed to model quality, whereas surface reasoning showed a negative, but non-significant correlation with model quality. This finding may indicate that the online log file measure is a valid indication of processing quality.

The results of our study are consistent with previous research within the field of achievement motivation, showing that self-efficacy and mastery-approach goal orientation are significantly positively related to students' use of deep cognitive processes (see Figure 4). In addition, the hypothesis that self-efficacy is related to mastery-approach goal orientation was supported. In the path model of Figure 4, model quality is positively affected by self-efficacy and mastery-approach goal orientation, but the effects of these variables are indirect, operating through the observed use of deep processes (i.e., mediation).

The paths between the variables performance-avoidance orientation and surface reasoning and between surface reasoning and model quality were not significant. However, the correlational analysis (see Table 3) showed that for both relations a trend was visible in the hypothesized direction. With a larger sample size these coefficients could turn out to be significant. On the other hand, recent studies also found that the paths between performance-avoidance goal orientation and surface processing (e.g., Dupeyrat & Mariné, 2005) and between surface processing and achievement were not significant.

Dupeyrat & Mariné (2005) and Elliot et al. (1999) found a significant negative relation between performance-avoidance orientation and students' use of deep processes, which implies that students high on performance-avoidance report that they employ less deep processes. However, this finding is not unequivocally replicated as our study, and the studies of Al Emadi (2001), Wolters (1996) and Midgley and Middleton (1997) have shown.

The present study shows that it is important to consider not only cognition as an important determinant of collaborative computer-based learning, but also to take into account the important impact of motivational factors, such as students' achievement goal orientation and their self-efficacy. In addition, our study shows that the conceptual model is also applicable to particular collaborative tasks using an online log file measure of students' reasoning.

An educational implication of this study may be that strategies that promote a mastery-approach goal orientation and advance students' self-efficacy lead to deep reasoning during modeling and ultimately to a higher achievement. For instance, a mastery-approach orientation can be encouraged when knowledge development is emphasized instead of evaluation of learning. In addition, mastery-approach goal orientation may be stimulated when an assignment provided to students is made interesting and challenging for them (Jagacinski, 1992; Meece et al., 1988). Finally, although one of the strongest ways for students to build self-efficacy is to experience success in accomplishing tasks (Schunk, 1991), external support and encouragement can also be provided. Relating this to collaborative computer-based modeling tasks, mastery-approach goal orientation and self-efficacy may be promoted by presenting students with modeling tasks that interests them, by avoiding normative comparisons with other students, and by having students exchange their model and ideas with other dyads (e.g., Hogan & Thomas, 2001). Moreover, collaborative learning activities, such as the assignment employed in our study, have also been found to promote students' achievement goal motivation (e.g., Jagacinski, 1992; Meece et al., 1988). Further research is needed in order to test these ideas.

CHAPTER 6

CONCLUSION AND DISCUSSION

The main purpose of this dissertation was to examine how students' reasoning during computer-based scientific modeling was influenced by characteristics of the students and of the modeling environment. In order to accomplish this aim we first developed an analysis scheme to assess reasoning processes students employ during modeling (Chapter 2). Then, we investigated the impact of students' epistemology (Chapter 3), of communication mode (Chapter 4), and of students' motivation (Chapter 5) on their reasoning. Four empirical studies were conducted in order to address the following research questions:

1) Which reasoning processes, employed by students during a computer-based modeling task, need support?
2) What is the relation between students' epistemological understanding of models and modeling and their reasoning during a computer-based modeling task?
3) What is the effect of chat communication versus face-to-face communication on students' modeling activities, on their reasoning during modeling, and on the quality of their model?
4) What is the relation between students' motivation and their reasoning during a computer-based modeling task and is the relation between students' motivation and model quality mediated by their reasoning?

These studies were reported in Chapters 2 through 5 and involved secondary students (age ranged between 16-18) working collaboratively in dyads on a computer-based dynamic modeling task. In Chapters 2 and 3 modeling was performed within the systems-oriented modeling tool Powersim (Byrkness & Myrtveit, 1997), and in Chapters 4 and 5 students modeled within the environment Co-Lab (Van Joolingen, et al., 2005). In this chapter, the main findings from these studies are summarized, discussed and integrated.

1. OVERVIEW AND DISCUSSION OF THE STUDIES

1.1 Which reasoning processes, employed by students during a computer-based modeling task, need support?

The study reported in Chapter 2 concerned the identification of features that are required to describe students' reasoning during scientific computer-based modeling. This issue was addressed by the development of a framework of analysis. We started with an overview of global reasoning processes identified in the available research literature and then refined our framework on the basis of a qualitative analysis of protocol data obtained from collaborating dyads. This resulted in a coding scheme with the following dimensions: a) type of reasoning process, b) topic focus, and c) type of argumentation. The coding scheme proved a useful tool to analyze occurrences of reasoning processes, as well as to assess their quality.

The protocols of three dyads were analyzed more closely in order to investigate how successful and less successful students can be characterized with respect to the reasoning processes they employ. We found that the former group referred more often to both experiential knowledge and physics knowledge during reasoning. In addition, the more successful students focused more often on relations between variables or on their model as a whole during modeling. The less successful students, in contrast, tended to limit their focus to specifying individual quantities to fit the empirical data (i.e., model fitting behavior). Quantitative analyses of all obtained protocols confirmed this picture. Correlation analysis revealed that students who spent much time on model fitting without argumentation arrived at models of lower quality. Students who spent their time on inductive reasoning with reference to prior knowledge scored higher on model quality. Although these correlations do not prove causation, based upon our theoretical framework we may assume that some reasoning processes may lead students to build models of higher quality and that other processes may result in models of lower quality.

Quantitative analyses also showed that students encountered particular difficulties with regard to the task perception, the content addressed and the tool used. At the level of task perception, we found that students were mainly engaged in superficial model fitting behavior, indicating that they were not able to use the model as a means to comprehend the behavior of a complex phenomenon. At the content level, students had problems with considering relations between quantities during modeling and with connecting the computer model with their own prior knowledge of the phenomenon being modeled. At the level of the tool, we found that students had a great deal of problems with grasping the system dynamics formalism, even after they had been presented with a modeling tutorial.

The degree to which students' reasoning and achievement within a computer-based modeling task is related to characteristics of the students and to features of the modeling environment was examined in Chapters 3 through 5.

1.2 What is the relation between students' epistemological understanding of models and modeling and their reasoning during a computer-based modeling task?

The purpose of the study reported in Chapter 3 was twofold: 1) to examine students' epistemological understanding of models and modeling, and 2) to investigate whether there is a relation between students' epistemological understanding and the level of their reasoning during computer-based modeling. We expected that a sophisticated epistemology is related to the employment of more deep reasoning processes and to the use of less surface reasoning processes during modeling. On the basis of the analysis scheme that was described in Chapter 2, we distinguished between deep versus surface reasoning in the modeling process by the criterion that deep reasoning processes involve integration with existing knowledge.

To assess students' epistemological understanding of models and modeling, we employed the three general levels articulated in the studies of Carey & Smith (1993) and Grosslight et al. (1991). In a level 1 epistemological understanding, models are viewed as simple copies of reality. A level 2 understanding involves an understanding that models are created for a specific purpose and that the model no longer must exactly correspond with the phenomenon being modeled. A level 3 understanding is characterized by the view that models are constructed in service of developing and testing ideas, rather than replicating reality. We determined the level of students' epistemological understanding of models and modeling for each of the following categories: a) nature of models, b) purposes of models, c) design and revision of models, and d) evaluation of models.

On average, students were found to hold a moderate (i.e. level 2) epistemological understanding of the nature of models, of the evaluation of models, and of the design and revision of models. With respect to students' epistemological understanding of the purposes of models, most students evenly had a level 2 or a level 3 score.

Regarding the relation between students' epistemological understanding of models and modeling and the level of their reasoning during modeling, we found two significant correlations: A positive correlation between students' epistemological understanding and deep reasoning and a negative correlation between students' epistemology and surface reasoning. Thus, we found a significant relation between how students conceptualize models and the nature of reasoning processes they employ during modeling. Consequently, appropriate scaffolds could be developed focusing on enhancing students' epistemological understanding of models and modeling (see Justi & Van Driel, 2005; Schwarz & White, 2005).

We measured students' epistemological understanding *after* they had completed the modeling task. Reason was that some authors have argued that students' epistemology is enacted during task performance and has to be measured within the specific context of the task (e.g., Hammer & Elby, 2002; Niessen, Abma, Widdershoven, & Van der Vleuten, 2004). Accordingly, we chose to contextualize our epistemology questionnaire and to administer it directly after the modeling task.

1.3 What is the effect of chat communication versus face-to-face communication on students' modeling activities, on their reasoning during modeling, and on the quality of their model?

In Chapter 4, we investigated the differential impact of face-to-face communication versus chat communication on: a) students' activities during modeling, b) students' reasoning during modeling, and c) the quality of the model students constructed.

In other studies, students who use chat are found to be constrained in their interactions, because of a lack of communication cues (e.g., Condon & Cech, 1996a; 1996b; Herring, 1999; Straus, 1997). We expected that this constraint of chat communication may either hinder students during modeling or may pressure them to increase the efficiency of their modeling. Two alternative hypotheses were considered regarding the consequences of face-to-face communication versus chat communication on students' performance: 1) Students' in the face-to-face condition will score significantly higher on model quality and will spend significantly less time on surface reasoning and significantly more time on deep reasoning compared to students in the chat condition. No significant difference is expected between the two conditions on the number of modeling activities performed. 2) Students in the face-to-face condition will score significantly lower on model quality and will spend significantly more time on surface reasoning and significantly less time on deep reasoning compared to students in the chat condition. In addition, students in the face-to-face condition will perform significantly more modeling activities than students in the chat condition.

We found that dyads in the face-to-face condition spent significantly more time on surface reasoning processes compared to students in the chat condition. In addition, students in the former group spent significantly less time on one of the six processes that were categorized as deep reasoning (i.e., inductive reasoning and reference to components). Finally, we found that dyads in the face-to-face condition saved and executed their model more often compared to students who communicated using chat. Although students in both conditions scored similar on model quality, we concluded that, in general, results were more in accordance with hypothesis 2.

The expectations derived from hypothesis 2, that students in the chat condition will compress their communication resulting in more efficient modeling, whereas students in the face-to-face condition will exchange more irrelevant information, were supported in additional analyses we performed. For instance, we found a significant negative correlation between the total number of reasoning episodes and model quality for the face-to-face group. Within the chat group, a significant positive correlation was found between the time students spent on episodes in which they quantified and referred to knowledge (i.e., deep reasoning) and model quality.

Our results conflict with some studies that show that face-to-face groups outperform groups who communicate using chat (e.g., Carey & Kacmar, 1997; Mennecke et al., 2000; Straus, 1997; Van der Meijden & Veenman, 2005). Two factors may explain this discrepancy in findings: a) the groups in the study reported in Chapter 4 consisted of dyads (instead of groups consisting of more than two students) and b) the participants in our study had more experience with using chat. It may be that if groups consist of more than two students, the modeling task necessitates for more communication cues, since it becomes more difficult for students to keep track of the discussion and of each other when they communicate through a chat facility. In addition, when students are less skilled in typing and using computer-mediated chat they may be less concise and less task focused than the students in our study, since they first have to learn to use the chat tool.

1.4 What is the relation between students' motivation and their reasoning during a computer-based modeling task and is the relation between students' motivation and model quality mediated by their reasoning?

Goal of the study reported in Chapter 5 was to investigate whether students' achievement goal orientation and self-efficacy (i.e., students' motivation) influences their reasoning during modeling and whether students' reasoning mediates the relation between motivation and quality of the model students constructed. Broadly defined, students' achievement goal orientation involves the reasons and the purposes to engage in achievement tasks. In the study reported in Chapter 5, we focused on two divergent types of achievement goal orientations: mastery-approach goal orientation and performance-avoidance goal orientation. A mastery-approach goal orientation involves attempting to develop one's skills and abilities, to advance one's learning, to understand the material, or to complete a task. Performance-avoidance goal orientation involves the avoidance of failure and the avoidance of the demonstration of low ability. Self-efficacy was defined as the learner's belief regarding their performance capabilities in a particular domain.

In contrast to most studies in the field of motivation, we did not assess students' reasoning based on self-report measures, but rather we based our assessments on the process observations as described in the previous chapters. In addition, students' achievement goal orientation and their self-efficacy were measured on the dyad-level.

The group measures for the motivation variables were found to be valid, in that the aggregated scores on the individual questionnaires correlated significantly with their corresponding group scores. As expected, we found that self-efficacy and mastery-approach goal orientation were significantly positively related to students' employment of deep reasoning processes. In addition, self-efficacy was related to mastery-approach goal orientation. Self-efficacy and mastery-approach goal orientation were both positively related to model quality, but the effects of these variables were indirect, operating through the use of deep reasoning processes (i.e., mediation). Unfortunately, we did not find a significant relation between performance-avoidance goal orientation and surface reasoning. In addition, the association between surface reasoning and model quality was not significant.

An educational implication of this study may be that strategies that promote a mastery-approach goal orientation and advance students' self-efficacy lead to deep reasoning during modeling and ultimately to a higher achievement.

2. GENERAL DISCUSSION

In this section we will discuss some issues that can be raised with respect to the research methodology chosen and regarding the scope of the research presented in Chapters 2 through 5 of this dissertation.

2.1 *The measurement of students' reasoning*

In the studies reported in this dissertation, we employed an online method for the measurement of students' reasoning during computer-based scientific modeling. Verbal protocols of collaborating dyads were transcribed (in Chapters 2 and 3) or obtained from log files (in Chapters 4 and 5) and analyzed according to a coding scheme, which was described in Chapter 2. This coding scheme was developed to assess the occurrence as well as the quality of students' reasoning. Although students' verbalizations do not necessarily reveal their covert reasoning, the assumption of employing this method was that it allowed us to build valid inferences about cognition on the basis of what students articulate during modeling activities (cf. Hogan & Thomas, 2001).

Except for the study reported in Chapter 3, students' reasoning was assessed at the level of the dyad. A problem related to measuring at this level may be that the quality of students' reasoning is over–or underestimated in those cases where the contributions of students within a dyad are unevenly distributed. For instance, it is possible that the amount of time a dyad spent on deep reasoning was the result of one student contributing consistently more than his/ her group mate. This risk remained largely hypothetical however, because in most dyads students were found to contribute in an even way

In Chapters 3 through 5 we differentiated between surface and deep reasoning processes. We conceptualized deep reasoning as episodes in which students were elaborating on the modeling task and connect to knowledge they have available. Episodes in which students employ unelaborated processes without referring to available knowledge were labeled as surface reasoning. The assumption underlying our choice for examining students' level of reasoning in these chapters was that the degree to which students obtain a learning benefit from computer-based modeling, highly depends on the nature of their reasoning (i.e., deep versus surface reasoning) during task performance. A review conducted by Covington (2000) supports this contention, showing that deep processing creates the optimal conditions for achievement in a variety of subject-matter areas, including science, whereas surface processing is linked to decreases in achievement.

In addition, our aim in these chapters was to elaborate on the findings of previous studies within the fields of students' motivation and epistemology that have also employed this dichotomy. In these studies the link between these constructs and level of students' processing during task performance is often made but only rarely investigated.

2.2 *The modeling task*

In most studies on computer-based scientific modeling students are asked to explore a given pre-built model or to construct a model from scratch, without having them to explicitly test their model against given or collected data. To make students' assignment more authentic, in the studies reported in this dissertation we asked students to extend an incomplete model that would ultimately match empirical data. In Chapters 2 and 3 (i.e., the ice-skater assignment) the data were provided to them in graphs and in Chapters 4 and 5 (i.e., the black sphere assignment) the data were generated from running a simulation.

It could be contended that the modeling task we presented to participants in our studies may have promoted model fitting behavior, which implies the occurrence of surface reasoning. That is, the assignment may have led students to adopt an engineering approach to the modeling task (cf. Schauble et al., 1991). However, to obtain an acceptable fit between the model output and the empirical data, employment of deep reasoning processes is, to some degree, necessary. For instance, in order to accomplish the modeling task used in Chapters 2 and 3, students had to recognize the presence of a feedback between the velocity of the ice-skater and the ice friction the ice-skater experiences. Similarly, the modeling task employed in the studies reported in Chapters 4 and 5 also required the use of deep reasoning processes in order to fit the model output with the data from the simulation.

2.3 Scope of our research

We conjecture that the findings obtained from the studies reported in this dissertation can be generalized to other forms of collaborative modeling which takes place within a shared computer-based environment. This also implies that the applicability of the conclusions and the implications of this dissertation are restricted to this context and it is not in our intention to generalize beyond the settings we had chosen for our research. Our measures for students' reasoning, motivation and epistemology were developed for this specific task context. In addition, as mentioned in Chapter 4 (see section Conclusion and Discussion), the impact of communication mode on students' reasoning is highly dependent on task characteristics, making it difficult to generalize to other task contexts.

In the studies reported in Chapters 3 and 5, we assessed students' motivation and epistemology at one point in time. However, the degree to which students were motivated to perform on the modeling task could have changed over time. In addition, it is possible that students' altered their epistemological understanding of models during modeling. Ultimately, this may have affected their reasoning and quality of the model they constructed (e.g., Järvelä, Rahikainen, & Lehtinen, 2001; Pintrich, 2003). For instance, Bandura (1982; 1986) argues that students' self-efficacy is highly influenced by the amount of success experienced during task performance. Even though we found significant relations between students' motivation, epistemology and the level of their reasoning, these results must hence be interpreted with care. Future research should investigate this issue of whether and how students' motivation and epistemology may change during modeling and how this may influence students reasoning.

A final remark is that even though we found significant positive correlations between some deep reasoning processes and model quality in the studies reported in Chapter 2, Chapter 4, and Chapter 5, the question what exactly students learn as a result of their reasoning, remains unanswered. That is, we do not know whether students have gained more conceptual knowledge of the phenomenon being modeled or of system dynamics in general, or whether they have acquired more modeling skills. Purpose of the present dissertation was to investigate students' reasoning during the initial phase(s) of computer-based modeling in order to provide suggestions for scaffolding. Future research should address the issue of whether and how students' reasoning affects what they learn.

3. IMPLICATIONS FOR SCIENCE EDUCATION

In this dissertation we focused on students' reasoning during modeling and attempted to examine the conditions that may foster deep reasoning and which may ultimately lead to a higher achievement. Based on the findings reported in this dissertation, we will provide some proposals for support which could help to improve students' reasoning during modeling. It has to be considered that the suggestions for scaffolding we present relate to novice modelers who, working in dyads, start using computer-based scientific models of complex and dynamic phenomena.

Based on the findings reported in Chapter 2 we suggest that appropriate support should be designed to scaffold students' reasoning on all three levels (i.e., level of task perception, content level, level of the tool) at which they are found to encounter difficulties. For instance, specific modeling (sub) goals can be presented which support students in comprehending how the structure of their model influences its behavior. This scaffold may support students' reasoning at the level of *task perception*. At the *content* level, scaffolds should encourage students to activate their prior knowledge not only during modeling but also before engaging in any modeling activities. When students are initially prompted to think about the phenomenon they are going to model, activating whatever prior knowledge they have available, can serve as an anchor for the further modeling process (e.g., Clement, et al., 1989; Hammer, 2000). Ideally, this knowledge activation takes place within a collaborative setting in which dyads exchange and discuss their models with other groups (e.g., Rouwette et al., 2000). Finally, at the level of *the tool*, students should start with the construction of a qualitative model of a phenomenon they are familiar with, before they are to quantify parameters in their model. In this way, students can first concentrate on grasping the system dynamics formalism, before they are asked to model more complex phenomena. In addition, a thorough instruction, provided by the teacher, to the modeling tool is necessary in order for students to become familiar with the formalism used. This instruction should also include the top-down approach to modeling, in which students learn to consider the dynamic behavior of their model (e.g., Hogan & Thomas, 2001).

For improving educational practice it is important not only to consider the pedagogical design of the modeling environment and the task(s) used, but also to take into account the characteristics of the students. In the studies reported in Chapters 3 and 5 we found that students' epistemology as well as their motivation was significantly related to the level of their reasoning during modeling. These findings offer a strong indication that attempts to foster students' deep reasoning during modeling will have a more powerful effect when these characteristics are also considered.

The study reported in Chapter 3 suggested several contexts for fostering a productive epistemological understanding of models, including reflecting on the nature, purpose and evaluation of models and on the process of scientific modeling. In addition, students should be confronted with strong examples of dynamic computer-based models. Finally, we argued that students should have the opportunity to build and to revise multiple models (e.g., Schwarz, 2002; Schwarz & White, 2005). These scaffolds could lead students to think about models in a more sophisticated manner ultimately leading them to employ more deep reasoning processes during modeling.

An implication that can be drawn from the findings of the study reported in Chapter 5 is, that it could be beneficial for students' reasoning and achievement to advance a mastery-approach goal orientation within the classroom and to foster students' confidence in their own ability (i.e., self-efficacy). A mastery-approach goal orientation can be promoted by avoiding normative comparisons and by developing an interest in the modeling task (e.g., Harackiewicz et al., 1998; Jagacinski, 1992; Pintrich & Schrauben, 1992). For instance, a mastery-approach goal orientation may be advanced when the teacher provides personalized comments, by emphasizing the importance of understanding the phenomena being modeled, or by providing feedback on students' reasoning during modeling. In addition, giving students the opportunity to obtain data from experiments they have conducted themselves could increase their motivation to perform (White & Frederiksen, 1998). Students' self-efficacy can be enhanced when the teacher offers support and encouragement (Schunk, 1991).

Finally, in Chapter 4 we found that students who communicated through a chat tool during modeling showed to be more efficient during modeling, compared to students who interacted face-to-face. The chat facility constrained students' communication leading them to compress their communication by being more concise compared to students who communicated face-to-face. This affordance of computer-mediated chat may lead students to be more task focused and more precise in their articulations. Although, more research in this area is needed, we concur that chat can be employed as an effective mode of communication in a complex computer-based modeling task.

The scaffolds we have suggested may enhance students' reasoning and achievement to a some extent and they could be implemented within the context of a curriculum based on computer-based modeling. However, it has to be noted that the studies reported in this dissertation were of relatively short duration and focused on students who had no prior experience with computer-based modeling. Therefore, we can only state that these scaffolds should be implemented in the phase where students start to use and construct models. We recommend that future research should examine the impact of students' epistemology, motivation, and communication mode on students' reasoning over a whole curriculum to decide what support is needed over the long run.

REFERENCES

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*, 131-152.

Al-Emadi, A.A. (2001). The relations among achievement, goal orientation, and study strategies. *Social Behavior and Personality, 29*(8), 823-832.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84*, 261-271.

Ames, C. & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*(3), 260-267.

Artzt, A.F. & Armour-Thomas, E. (1992). Development of a cognitive-metacognitive framework for protocol analysis of mathematical problem solving in small groups. *Cognition and Instruction, 9*(2), 137-175.

Atkinson, J. (1964). *An Introduction to Motivation*. Princeton, NJ: Van Nostrand.

Avouris, N., Margaritis, M., Komis, V., Melendez, R., & Saez A. (2003). ModelingSpace: Interaction Design and Architecture of a collaborative modeling environment. *Paper presented at the Sixth International Conference on Computer Based Learning in Sciences*, Nicosia, Cyprus.

Baker, M.J. & Lund, K. (1997). Promoting reflective interactions in a CSCL environment. *Journal of Computer Assisted Learning, 13*, 175-193.

Baltes, B.B., Dickson, M.W., Sherman, M.P., Bauer, C.C., & LaGanke, J.S. (2002). Computer-mediated communication and group decision making: A meta-analysis. *Organizational behavior and Human Decision Processes, 87*(1), 156-179.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*, 122-147.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Barile, A.L. & Durso, F.T. (2002). Computer-mediated communication in collaborative writing. *Computers in Human Behavior, 18*, 173-190.

Barowy, W., & Roberts, N. (1999). Modeling as an inquiry activity in school science: What's the point? In N. Roberts, W. Feurzeig, & B. Hunter (Eds.), *Modeling and Simulation in Science and Mathematics Education* (pp. 197-225). Berlin: Springer Verlag.

Barron, K.E. & Harackiewicz, J.M. (2001). Achievement goals and optimal motivation: Testing multiple goal models. *Journal of Personality and Social Psychology, 80*(5), 706-722.

Basque, J. & Pudelko, B. (2004). The effect of collaborative knowledge modeling at distance on performance and on learning. *Paper presented at the First International Conference on Concept Mapping*, Pamplona, Spain.

Bliss, J. (1994). From mental models to modeling. In H. Mellar, J. Bliss, R. Boohan, J. Ogborn, & C. Thompsett (Eds.), *Learning with Artificial Worlds: Computer Based Modeling in the Curriculum* (pp. 27-33). London: The Falmer Press.

Bordia, P. (1997). Face-to-face versus computer-mediated communication: A synthesis of the experimental literature. *The Journal of Business Communication, 34*(1), 99-120.

Bouffard, T., Boisvet, J., Vezeau, C., & Larouche, C. (1995). The impact of goal orientation or self-regulation and performance among college students. *British Journal of Educational Psychology*, *65*(3), 317-329.

Brickell, G., Ferry, B., & Harper, B. (2002). Developing informal reasoning skills in ill-structured environments. A case study into problem-solving strategies. In D. Williamson, C. Gunn, A. Young, & T. Clear (Eds.), *Proceedings of the 19th Annual Conference of the Australasian Society for Computers in Learning and Tertiary Education* (pp.65-74). Auckland, New Zealand: UNITEC Institute of Technology.

Bruinsma, M. (2004). Motivation, cognitive processing and achievement in higher education. *Learning and Instruction, 14*, 549-568.

Buffler, A., Allie, S., Lubben, F., & Campbell, B. (2001). The development of first year physics students' ideas about measurement in terms of point and set paradigms. *International Journal of Science Education, 23*(11), 1137-1156.

Butler, R. (1993). Effects of task- and ego-achievement goals on information seeking during task engagement. *Journal of Personality and Social Psychology, 65*(1), 18-31.

Byrknes, A. & Myrtveit, M. (1997). *Learning dynamic modeling*. Goteberg, Sweden: Powersim Press.

Carey, J.M. & Kacmar, C.J. (1997). The impact of communication mode and task complexity on small group performance and member satisfaction. *Computers in Human Behavior, 13*(1), 23-49.

Carey, S. & Smith, C.L. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist, 28*(3), 235-251.

Chi, M.T.H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, *6*(3), 271-315.

Chinn, C.A. & Brewer, W.F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction*. Review of Educational Research, 63*, 1-49.

Chiu, C.H., Wu, W.S., & Huang, C.C. (2000). Collaborative concept mapping processes mediated by computer. In G. Davies & C.B. Owen (Eds.), *Proceedings of WebNet 2000 - World Conference on the WWW and Internet* (pp.95-100). San Antonio, Texas: AACE 2000.

Clement, J., Brown, D., & Zeitsman, A (1989). Not all preconceptions are misconceptions: Finding 'anchoring conceptions' for grounding instruction on students' intuitions*. International Journal of Science Education, 11*, 554-565.

Cohen, A. L., & Scardamalia, M. (1998). Discourse about ideas: Monitoring and regulation in face-to-face and computer-mediated learning environments. *Interactive Learning Environments, 6*(1-2), 114-142.

Condon, S.L., & Cech, C.G. (1996a). Functional comparison of face-to-face and computer-mediated decision-making interactions. In Herring, S.C. (Ed.), *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives* (pp. 65-80). Philadelphia: John Benjamin,.

Condon, S.L., & Cech, C.G. (1996b). Discourse management in face-to-face and computer-mediated decision-making interactions. *Electronic Journal of Communication/La Revue Electronique de Communication, 6*(3).

Coon, T. (1988). Using STELLA simulation software in life science education. Computers in Life Science Education, *5*(9), 56-71.

Covington, M.V. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual Review of Psychology, 51*, 171-200.

Cox, M & Webb, M. (1994). In H. Mellar, J. Bliss, R. Boohan, J. Ogborn, & C. Thompsett (Eds.), *Learning with Artificial Worlds: Computer Based Modeling in the Curriculum* (pp. 188-198). London: The Falmer Press.

Craik, F.I. & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 971-684.

Crawford, B.A. & Cullin, M.J. (2004). Supporting prospective teachers' conceptions of modeling in science. *International Journal of Science Education, 26*(11), 1379-1401.

Daft, R.L., & Lengel, R.H. (1984). Information richness: A new approach to managerial behavior and organizational design. *Research in Organizational Behavior, 19,* 528-563.

Daft, R.L. & Lengel, R.H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*(5), 554-571.

Daly, B.L. (1993). The influence of face-to-face versus computer-mediated communication channels on collective induction. *Accounting, Management, and Information Technologies, 3*(1), 1-22.

Dennis, J.K. (2003). Problem-based learning in online vs. face-to-face environments. *Education for Health, 16*(2), 198-209.

Devi, R., Tiberghien, A., Baker, M.J. & Brna, P. (1996). Modeling students' construction of energy models in physics*. Instructional Science, 24*, 259-293.

Doerr, H.M. (1995*). An integrated approach to mathematical modeling: A classroom study.* Paper presented at the Annual Meeting of the American Educational Research Association, San Fransisco, CA.

Doerr, H.M. (1996). STELLA ten years later: A review of the literature. *International Journal of Computers for Mathematical Learning, 1*, 201-224.

Doerry, E. (1995). *An Empirical Comparison of Co-present and Technologically-Mediated Interaction based on Communicative Breakdown.* Unpublished doctoral dissertation. University of Oregon.

Doyle, J.K. & Ford, D.N. (1998). Mental models concepts for system dynamics research. *System Dynamics Review, 14*(1), 3-29.

Driel, J.H. van & Verloop, N. (1999). Teachers' knowledge of models and modeling in science. *International Journal of Science Education, 21*(11), 1141-1153.

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young People's Images of Science*. Buckingham, UK: Open University Press.

Duda, J.L. & Nicholls, J.G. (1992). Dimensions of achievement motivation in school work and sport. *Journal of Educational Psychology, 84,* 290-299.

Dupeyrat, C. & Mariné, C. (2005). Implicit theories of intelligence, goal orientation, cognitive engagement, and achievement: A test of Dweck's model with returning to school adults. *Contemporary Educational Psychology, 30,* 43-59.

Dweck, C.S. (1985). Intrinsic motivation, perceived control, and self-evaluation maintenance: An achievement goal analysis. In C. Ames & R. Ames (Eds.), *Research on Motivation in Education: Vol. 2. The Classroom Milieu* (pp. 289-305). New York: Academic.

Dweck, C.S. & Leggett, E.L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273.

Elby, A. & Hammer, D. (2001). On the substance of a sophisticated epistemology. *Science Education*, *85*(5), 554-567.

Elliot, A.J. (1997). Integrating the "classic" and "contemporary" approaches to achievement motivation: A hierarchical model of approach and avoidance achievement motivation. In M.L. Maehr & P.R. Pintrich (Eds.), *Advances in Motivation and Achievement* (Vol. 10, pp.143-179). Greenwich, CT: JAI Press.

Elliot, A.J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*(3*)*, 169-189.

Elliot, A.J. & Church, M.A. (1997). A hierarchical model of approach and avoidance achievement motivation*. Journal of Personality and Social Psychology, 72*(1), 218-232.

Elliot, A.J., McGregor, H.A., & Gable, S. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. *Journal of Educational Psychology, 91*(3), 549-563.

Elliot, A.J. & McGregor, H.A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501-519.

Elliott, E.S. & Dweck, C.S. (1988). Goals: An approach to motivation and achievement*. Journal of Personality and Social Psychology, 54*(1), 5-12.

Entwistle, N.J. (1979). Stages, levels, styles, or strategies: Dilemmas in the description of thinking. *Educational Review, 31,* 123-132.

Entwistle, N.J. (1988). *Styles of Learning and Teaching*. David Fulton: London.

Entwistle, N.J. (2001). Styles of learning and approaches to studying in higher education. *Learning Styles in Higher Education, 30*(5/6), 593-602.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis; verbal reports as data.* Cambridge: MIT Press.

Erkens, G. (1998). *Multiple Episode Protocol Analysis (MEPA 4.0)* (Internal publication). Utrecht, The Netherlands: Department of Educational Sciences, Utrecht University.

Fisher, S.L. & Ford, J.K. (1998). Differential effects of learner effort and goal orientation on two learning outcomes. *Personnel Psychology, 51*, 397-420.

Ford, J.K., Smith, E.M., Weissbein, D.A., Gully, S.M., & Salas, E. (1998). Relations of goal orientation, metareasoning activity, and practice strategies with learning outcome and transfer. *Journal of Applied Psychology, 83*(2), 218-233.

Forrester, J. (1961). *Industrial Dynamics*. Portland, OR: Productivity Press.

Fretz, E.B., Wu, H., Zhang, B., Davis, E.A., Krajcik, J.S., & Soloway, E. (2003). An investigation of software supporting modeling practices. *Research in Science Education, 32,* 567-589.

Garcia, T., McCann, E.J., Turner, J.E., & Roska, L. (1998). Modeling the mediating role of volition in the learning process. *Contemporary Educational Psychology, 23*, 392-418.

Garner, R. (1988). Verbal-report data on cognitive and metacognitive strategies. In C.E. Weinstein, E.T. Goetz, & P.A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 63-76). San Diego: Academic Press.

Giere, R.N. (1990). *Explaining science*. Chicago: University of Chicago Press.

Gijlers, H. & Jong, T. de (2005). The relation between prior knowledge and students' collaborative discovery learning processes. *Journal of Research in Science Teaching, 42*(3), 264-282.

Gilbert, S.W. (1991). Model building and a definition of science. *Journal of Research in Science Teaching, 28,* 73-79.

Gilbert, J. K., Boulter, C. J., Rutherford, M. (1998). Models in Explanations, Part 1: Horses for Courses?. *International Journal of Science Education*, *20*(1), 83-97.

Gobert, J. & Discenna, J. (1997). *The relation between students' epistemologies and model-based reasoning*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Gobert, J., Snyder, J., & Houghton, C. (2002). *The Influence of Students' Understanding of Models on Model-based Reasoning.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Graham, S. & Golan, S. (1991). Motivational influences on cognition: Task involvement, ego involvement, and depth of information processing. *Journal of Educational Psychology, 83*(2), 187-194.

Greene, B.A. & Miller, R.B. (1996). Influences on achievement: Goals, perceived ability, and reasoning engagement. *Contemporary Educational Psychology, 21*, 181-192.

Grosslight, L., Unger, C., Jay, E., & Smith, C.L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching, 28*(9), 799-822.

Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction, 12*(2), 151-183.

Hammer, D. (2000). Students' resources for learning introductory physics. *American Journal of Physics, Physics Education Research Supplement, 68*(SI), S52-S59.

Hammer, D. & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer, & P. R. Pintrich (Ed.), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing* (pp. 169-190). Mahwah, N.J.: Lawrence Erlbaum.

Harackiewicz, J.M., Barron, K.E., & Elliot, A.J. (1998). Rethinking achievement goals: When are they adaptive for college students and why? *Educational Psychologist, 33*(1), 1-21.

Harackiewicz, J.M., Barron, K.E., Elliot, A.J., Carter, S.M., & Lehto, A.T. (1997). Predictors and consequences of achievement goals in the college classroom: Maintaining interest and making the grade. *Journal of Personality and Social Psychology, 73*(6), 1284-1295.

Harackiewicz, J.M., Barron, K.E., Pintrich, P.R., Elliot, A.J. & Thrash, T.M. (2002). Revision of achievement goal theory: Necessary and illuminating. *Journal of Educational Psychology, 94*(3), 638-645.

Harrison, A.G. (2001). How do teachers and textbook writers model scientific ideas for students? *Research in Science Education, 31*, 401-435.

Henri, F. (1992). Computer conference and content analysis. In A. R. Kaye (Ed.), *Collaborative learning through computer conferencing: The Najaden papers* (pp. 117-136). London: Springer Verlag.

Herring, S.C. (1999). Interactional coherence in CMC. *Journal of Computer-Mediated Communication, 4*(4).

Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics, 55*(5), 440-454.

Hestenes, D. (1997). Modeling methodology for physics teachers. In E. Redish & J. Rigden (Eds.), The Changing Role of the Physics Department in Modern Universities (pp. 935-957). American Institute of Physics Part II.

Heuvelmans, A.P.J.M. & Sanders, P.F. (1993). Beoordelaarsoverstemming [Inter-judgement reliability measurement]. In T.J.H.M. Eggen & P.F. Sanders (Eds.), *Psychometrie in de praktijk* [*psychometrics in practice*] (pp. 443–469). Arnhem, The Netherlands: CITO.

Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review, 13,* 353–382.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67,* 88–140.

Hogan, K. (1999). Relating students' personal frameworks for science learning to their cognition in collaborative contexts. *Science Education, 83*(1), 1-32.

Hogan, K. (2000). Exploring a process view of students' knowledge about the nature of science. *Science Education,* 84, 51-70.

Hogan, K. & Thomas, D. (2001). Cognitive comparisons of students' systems modeling in ecology. *Journal of Science Education and Technology, 10*(4), 319-345.

Hollingshead, A.B., McGrath, J.E., & O'Connor, K.M. (1993). Group task performance and communication technology: A longitudinal study of computer-mediated versus face-to-face work groups. *Small Group Research, 24*(3), 307-333.

Hooper, S. & Hannafin, M.J. (1991). The effects of group composition on achievement, interaction, and learning efficiency during computer-based cooperative instruction. *Educational Technology Research and Development, 39,* 27-40.

Jackson, S.L., Stratford, S.J., Krajcik, J.S., & Soloway, E. (1996). Making dynamic modeling accessible to precollege science students. *Interactive Learning Environments*, *4*(3), 233-257.

Jagacinski, C.M. (1992). The effects of task involvement and ego involvement on achievement-related cognitions and behaviors. *Student perceptions in the classroom* (pp. 307-322). Greenwich, CT: JAI press.

Järvelä, S., Rahikainen, M., & Lehtinen, E. (2001). *Meaningful participation in a process of computer-based inquiry: The analyses of students' motivational involvement.* Paper presented at the 9th European conference for Research on Learning and Instruction, Freiburg, Switzerland.

Johnson, S. D., Aragon, S. R., Shaik, N., & Palma-Rivas, N. (2000). Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments. *Journal of Interactive Learning Research, 11*, 29-49.

Jonassen, D.H. & Kwon, H.I. (2001). Communication patterns in computer-mediated versus face-to-face group problem solving. *Educational Technology Research and Development, 49*(1), 35-51.

Jonassen, D.H., Strobel, J., & Gottdenker, J. (2005). Modeling for meaningful learning. In Learning Sciences and Technologies Group (Ed.), *Engaged Learning with Emerging Technologies* (pp. 1-28). Dordrecht, NL: Springer Verlag.

Jong, T. de & Joolingen, W.R. van (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, *68*(2), 179-291.

Jong, T. de, Joolingen, W.R. van, Lazonder, A.W., Ootes, S., Savelsbergh, E.R., & Wilhelm, P. (2002). *Co-Lab Specifications; Part 1 Theoretical background (Technical Report)*. Enschede, The Netherlands: University of Twente.

Joolingen, W.R. van, Jong, T. de, Lazonder, A.W., Savelsbergh, E.R. & Manlove, S. (2005). Co-Lab: Research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior, 21*, 671-688.

Justi, R.S. & Gilbert, J.K. (2002). Modeling, teachers' views on the nature of modeling, and implications for the education of modelers. *International Journal of Science Education, 24*(4), 369-387.

Justi, R.S. & Gilbert, J.K. (2003). Teachers' views on the nature of models. *International Journal of Science Education, 25*(11), 1369-1386.

Justi, R.S. & Driel, J.H. van (2005). The development of science teachers' knowledge on models and modeling: Promoting, characterising, and understanding the process. *International Journal of Science Education, 27*(5), 549-573.

Kainz, D. & Ossimitz, G. (2002). *Can students learn stock-flow-thinking? An empirical investigation.* Paper presented at the 2002 System Dynamics Conference, Palermo, Italy.

Kardash, C.M. & Howell, K.L. (2000). Effects of epistemological beliefs and topic-specific beliefs on undergraduates' cognitive and strategic processing of dual-positional text. *Journal of Educational Psychology, 92*(3), 524-535.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review, 96*(4), 674-689.

Kurtz dos Santos, A.C., & Ogborn, J. (1994). Sixth form students' ability to engage in computational modeling. *Journal of Computer Assisted Learning, 10*, 182-200.

Leach, J., Millar, R., Ryder, J., & Séré, M.G. (2000). Epistemological understanding in science learning: The consistency of representations across contexts. *Learning and Instruction, 10*, 497-527.

Lebie, L., Rhoades, J.A., & McGrath, J.E. (1996). Interaction process in computer-mediated and face-to-face groups*. Computer Supported Cooperative Work, 4*, 127-152.

Lederman, N.G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 26*(9), 771-783.

Löhner, S. & Joolingen, W. R. van, & Savelsbergh, E.R. (2003). The effect of external representation on constructing computer models of complex phenomena. *Instructional Science, 31*(6), 395-418 .

Löhner S., Joolingen, W. R. van, & Savelsbergh, E.R., & Hout-Wolters, B.H.A.M. van (2005). Students' reasoning during modeling in an inquiry learning environment. *Computers in Human Behavior, 21*(3), 441-461.

Mandinach, E.B. (1988). *The cognitive effects of simulation-modeling software and systems thinking on learning and achievement.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Marton, F. & Säljö, R. (1976). On qualitative differences in learning. I – outcome and process. *British Journal of Educational Psychology, 46*, 4-11.

Marton, F. & Säljö, R. (1997). Approaches to learning. In F. Marton, D.J. Hounsell, & N.J. Entwistle (Eds.), *The Experience of Learning, 2nd Edition (*pp. 39-58). Edinburgh: Scottish Academic Press.

McGrath, J.E. (1984). *Groups: Interaction and Performance.* Englewood Cliffs, NJ: Prentice-Hall.

McGrath, J.E. & Hollingshead, A.B. (1993). Putting the "group" back in group support systems: Some theoretical issues about dynamic processes in groups with technological enhancements. In Jessup, L.M. (Ed.), *Group Support Systems: New Perspectives* (pp. 78-96). New York: MacMillan Publishing.

McLaughlin, M. L. (1984). *Conversation: How talk is organized*. Beverly Hills, CA: Sage.

McWhaw, K. & Abrami, P.C. (2001). Student goal orientation and interest: Effects on students' use of self-regulated learning strategies. *Contemporary Educational Psychology, 26,* 311-329.

Meece, J.L., Blumenfeld, P.C., & Hoyle, R.H. (1988). Students' goal orientations and reasoning engagement in classroom activities. *Journal of Educational Psychology, 80*(4), 514-523.

Meece, J.L. & Holt, K. (1993). A pattern analysis of students' achievement goals. *Journal of Educational Psychology, 85*(4), 582-590.

Meijden, H. van der & Veenman, S. (2005). Face-to-face versus computer-mediated communication in a primary school setting. *Computers in Human Behavior, 21*, 831-859.

Mennecke, B.E., Valacich, J.S., & Wheeler, B.C. (2000). The effects of media and task on user performance: A test of the task-media fit hypothesis. *Group Decision and Negotiation, 9,* 507-529.

Middleton, M.J. & Midgley, C. (1997). Avoiding the demonstration of lack of ability: An underexplored aspect of goal theory. *Journal of Educational Psychology, 80*, 514-523.

Midgley, C., Kaplan, A., & Middleton, M.J. (2001). Performance-approach goals: Good for what, for whom, under what circumstances, and at what costs? *Journal of Educational Psychology, 93*(1), 77-86.

Millar, R., Lubben, F., Gott, R., & Duggan, S. (1994). Investigating in the school science laboratory: Conceptual and procedural knowledge and their influence on performance. *Research Papers in Education, 9*(2), 207-248.

Miller, R.B., Greene, B.A., Montalvo, G.P., Ravindran, B., & Nichols, J.D. (1996). *Contemporary Educational Psychology, 21*, 388-422.

Milrad, M., Spector, J.M., & Davidsen, P.I. (2002). Model facilitated learning. In S. Naidu (Ed*.), Learning and Teaching with technology: Principles and Practice* (pp. 13-27). London: Kogan Page.

Newlands, A., Anderson, A., & Mullin, J. (2003). Adapting communicative strategies to computer-mediated communication: An analysis of task performance and dialogue structure. *Applied Cognitive Psychology, 17*, 325-348.

Nicholls, J.G., Patashnick, Cheung, P.C., Thorkildsen, T.A., & Lauer, J.M. (1989). Can achievement motivation theory succeed with only one conception of success? In F. Halisch & J.H. van den Bercken (Eds.), *Interactional Perspectives on Achievement and Task Motivation* (pp. 198-208). Amsterdam/ Lisse: Swets & Zeitlinger.

Niessen, T., Abma, T., Widdershoven, G., Vleuten, C. van der (2004). *Contemporary epistemological research: the need for a reconceptualization*. Paper presented at the Onderwijs Research Dagen, Utrecht, The Netherlands.

Nolen, S.B. (1988). Reasons for studying: Motivational orientations and study strategies. *Cognition and Instruction, 5*(4), 269-287.

Ogborn, J. (1999). Modeling Clay for Thinking and Learning. In Roberts, N., Feurzeig, W. & Hunter, B. (Eds.), *Computer Modeling and Simulation in Science Education* (pp. 5-37). Berlin: Springer Verlag.

Ogborn, J. & Wong, D. (1984). A microcomputer dynamical modeling system. *Physics Education, 19*, 138-142.

Olaniran, B. A., Savage, G. T. & Sorenson, R. L. (1996). Experimental and experiential approaches to teaching face-to-face and computer-mediated group discussion. C*ommunication Education, 45*, 244-257.

Olson, G.M. & Olson, J.S. (1997). Research on computer supported cooperative work. In Helander, M., Landauer, T.K. & Prabhu, P. (Eds.), *Handbook of Human-Computer Interaction* (pp. 1433-1456). Amsterdam: Elsevier Science.

Ossimitz, G. (2002). *Stock-flow-thinking and reading stock-flow-related graphs: An empirical investigation in dynamic thinking abilities.* Paper presented at the 2002 System Dynamics Conference, Palermo, Italy.

Penner, D. E. (2001). Cognition, computers, and synthetic science: Building knowledge and meaning through modeling. *Review of Research in Education*, *25*, 1-35.

Pintrich, P.R. (2000a). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*(3), 544-555.

Pintrich, P.R. (2000b). An achievement goal theory perspective on issues in motivation terminology, theory, and research. *Contemporary Educational Psychology, 25*, 92-104.

Pintrich, P.R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*(4), 667-686.

Pintrich, P.R., Conley, A.M., & Kempler, T.M. (2003). Current issues in achievement goal theory and research. *International Journal of Educational Research, 39,* 319-337.

Pintrich, P.R. & DeGroot, E.V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33-40.

Pintrich, P.R. & Garcia, T. (1991). Student goal orientation and self-regulation in the college classroom. In M.L. Maehr & P.R. Pintrich (Eds.), *Advances in Motivation and achievement*, *vol. 7* (pp. 371-401). Greenwich, CT: JAI Press.

Pintrich, P.R. & Schrauben, B. (1992). Students' motivational beliefs and their reasoning engagement in classroom academic tasks. In D.H. Schunk & J.L. Meece (Eds.), *Student perceptions in the classroom* (pp. 149-183). Greenwich, CT: JAI press.

Purdie, N., Hattie, J. and Douglas, G. (1996) Student conceptions of learning and their use of self-regulated learning strategies: a cross cultural comparison. *Journal of Educational Psychology, 88*(1), 87-100.

Raghavan, K., Sartoris, M.L & Glaser, R. (1998). Impact of the MARS curriculum: The mass unit. *Science Education, 82*, 53-91.

Ramsden, P. (1992). *Learning to Teach in Higher Education*. Routledge: London.

Ramsden, P. & Entwistle, N.J. (1981). Effects of academic departments on students' approaches to studying. *British Journal of Educational Psychology, 51,* 151-164.

Rana, A.R., Turoff, M., & Hiltz, S.R. (1997). Task and technology interaction (TTI): A theory of technological support for group tasks. *Proceedings of the Thirtieth Annual Hawaii International Conference on System Sciences – Volume 2* (pp.66-76). Washington DC: IEEE Computer Society.

Reid, D.J. & Reid, F.J.M. (2005). An in-depth comparison of computer-mediated and conventional focus group discussions. *International Journal of Market Research, 47*(2), 131-162.

Riley, D. (1990). Learning about systems by making models. *Computers and Education, 15*(1-3), 255-263.

Roth, W. & Roychoudhury, A. (1994). Physics students' epistemologies and views about knowing and learning. *Journal of Research in Science Teaching, 31*(1), 5-30.

Rouwette, E.A.J.A., Vennix, J.A.M., & Thijssen, C.M. (2000). Group model building: A decision room approach. *Simulation & Gaming, 31*(3), 359-379.

Ryder, J. & Leach, J. (1999). University science students' experiences of investigative project work and their images of science. *International Journal of Science Education, 21*(9), 945-956.

Sandoval, W.A. (2003). The inquiry paradox: Why doing science doesn't necessarily change ideas about science. In C.P. Constantinou & Z.C. Zacharia (Eds.), *Proceedings of the Sixth International Computer-Based Learning in Science Conference* (pp. 825-834). Nicosia, Cyprus.

Schauble, L., Glaser, R., Raghavan, K., & Reiner, M. (1991). Causal models and experimentation strategies in scientific reasoning. *The Journal of the Learning Sciences, 1*(2)*,* 201-238.

Schecker, H.P. (1993). Learning physics by making models. *Physics Education, 28*, 102-106.

Schecker, H.P. (1998). *Physik – Modellieren, grafikorientierte modelbildungssysteme im physikunterricht* [Physics – Modeling, graphically oriented model building systems in physics instruction]. Stuttgart, Germany: Ernst Klett Verlag GmbH.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*(3), 498-504.

Schraw, G., Horn, C., Thorndike-Christ, T., & Bruning, R. (1995). Academic goal orientations and student classroom achievement. *Contemporary Educational Psychology, 20,* 359-368.

Schunk, D.H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*, 207-231.

Schunk, D.H. (2005). Self-regulated learning: The educational legacy of Paul P. Pintrich. *Educational Psychologist*, *40*(2), 85-94.

Schunk, D. H., & Zimmerman, B. J. (2001). *Self-regulated learning and academic achievement: Theoretical perspectives (2nd ed.)*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc..

Schwarz, C.V. (2002a). *Using Model-centered Science Instruction to foster Students' Epistemologies in Learning with Models*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Schwarz, C.V. (2002b). The role of epistemic forms and games: Perspectives on the role of meta-modeling knowledge in learning with models. In P. Bell, R. Stevens & T Satwicz (Eds.), *Keeping Learning Complex: The proceedings of the Fifth International Conference of the Learning Sciences (ICLS)* (pp. 414-420). Mahwah, NJ: Earlbaum.

Schwarz, C.V. & White, B.Y. (2005). Metamodeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction, 23*(2), 165-205.

Seegers, G. & Boeckaerts, M. (1993). Task motivation and mathematics achievement in actual task situations. *Learning and Instruction, 3*, 133-150.

Selinger, M. (1998). Forming a critical community through telematics. *Computers Education 30*(1/2), 23-30.

Sins, P.H.M., Savelsbergh, E.R., Joolingen, W.R. van (2005). The difficult process of scientific modeling: An analysis of novices' reasoning during computer-based modeling. *International Journal of Science Education, 27*(14), 1695-1721.

Smit, J.J., & Finegold, M. (1995). Models in physics: Perceptions held by final-year prospective physical science teachers studying at South African Universities. *International Journal of Science Education, 19,* 621-634.

Songer, N.B. & Linn, M.C. (1991). How do students' views of science influence knowledge integration? *Journal of Research in Science Teaching, 28*(9), 761-784.

Spector, J.M. (2000). System dynamics and interactive learning environments: Lessons learned and implications for the future. *Simulation & Gaming, 31*(4), 528-535.

Spector, J. M., & Davidsen, P. I. (1997). Creating engaging courseware using system dynamics. *Computers in Human Behavior*, 13(2), 127-155.

Spitulnik, M.W., Krajcik, J.S., & Soloway, E. (1999). Construction of models to promote scientific understanding. In N. Roberts, W. Feurzeig, & B. Hunter (Eds.), *Modeling and Simulation in Science and Mathematics Education* (pp. 71-94). Berlin: Springer Verlag.

Steed, M.B. (1994). *Effects of computer simulation construction on shifts in cognitive representation: A case study using STELLA*. Unpublished doctoral dissertation. University of Massachusetts, Amherst, MA.

Stratford, S.J. (1997). A review of computer-based model research in pre-college science classrooms. *Journal of Computers in Mathematics and Science Teaching, 16*(1), 3-23.

Stratford, S.J., Krajcik, J.S., & 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.

Straus, S.G. (1996). Getting a clue: The effects of communication media and information distribution on participation and performance in computer-mediated and face-to-face groups*. Small Group Research, 27*(1), 115-142.

Straus, S.G. (1997). Technology, group process, and group outcomes: Testing the connections in computer-mediated and face-to-face groups. *Human-Computer Interaction, 12*, 227-266.

Straus, S.G. (1999). Testing a typology of tasks: An empirical validation of McGrath's (1984) group task circumplex. *Small Group Research, 30*(2), 166-187.

Straus, S.G. & McGrath, J.E. (1994). Does the medium matter? The interaction of task type and technology on group performance and member reactions. *Journal of Applied Psychology, 79*(1), 87-97.

Suthers, D.D. (1999). Effects of Alternate Representations of Evidential Relations on Collaborative Learning Discourse. In C. M. Hoadley and J. Roschelle (Eds.), *Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 Conference* (pp. 611-620). Palo Alto, CA: Stanford University.

Suthers, D.D., Hundhausen, C.D. & Girardeau, L.E. (2003). Comparing the roles of representations in face-to-face and online computer supported collaborative learning. *Computers and Education, 41,* 335-351.

Sweeney, L. B. & Sterman, J. (2000). Bathtub dynamics: Initial results of a systems thinking inventory. *System Dynamics Review, 16*(4), 249-286.

Talsma, V.L. (2000). Scientific understandings revealed by students' computer models of a stream: A trickle or a flood? In B. Fishman & S. O'Conner-Divelbiss (Eds.), *Fourth International Conference of the Learning Sciences* (pp. 98-105). Mahwah, NJ: Erlbaum.

Tinker, R. (1993). Modeling and theory building: Technology is support of student theorizing. In D.L. Ferguson (Ed.), *Advanced Educational Technologies for Mathematics and Science* (pp. 91-113). Berlin: Springer-Verlag.

Treagust, D.F., Chittleborough, G., & Mamiala, T.L. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education, 24*(4), 357-368.

Valanides, N. & Angeli, C. (2005). Effects of instruction on changes in epistemological beliefs. *Contemporary Educational Psychology, 30,* 314-330.

Valle, A., Cabanach, R.G., Nunez, J.C., Gonzalez-Pienda, J., Rodriguez, S., & Pineiro, I. (2003). Cognitive, motivational, and volitional dimensions of learning: An empirical test of a hypothetical model. *Research in Higher Education, 44*(5), 557-580.

Vermetten, Y.J., Lodewijks, H.G., & Vermunt, J.D. (2001). The role of personality traits and goal orientations in strategy use. *Contemporary Educational Psychology, 26*, 149-170.

Vrugt, A., Hoogstraten, J., Oort, F.J. (1999). Doelorientaties, waargenomen eigen competentie en studieresultaten (Goal orientations, perceived self-efficacy and study results). *Tijdschrift voor Onderwijs Research, 23*, 210-223.

Vrugt, A., Oort, F.J., & Zeeberg, C. (2002). Goal orientations, perceived self-efficacy and study results amongst beginners and advanced students. *British Journal of Educational Psychology, 72,* 385-397.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13,* 21-39.

Webb, N. M. (1991). Task related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education, 22*, 366-389.

Webb, N.M. & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal, 31*(2), 369-395.

Webb, N.M., Welner, K., & Zuniga, S. (2001). Short circuits or superconductors? The effects of group composition on high-ability students' science assessment performance. *American Educational Research Journal, 39,* 943–989.

White, B.Y. & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1)*,* 3-117.

Whitfield, A. (1988). STELLA and its impact on the teaching of mathematical modeling. In F. Lovis & E.D. Tagg (Eds.), *Computers in Education, Proceedings of the IFIP TC3 First European Conference on Computers in Education – ECCE 88, International Federation for Information Processing* (pp. 299-304). Amsterdam, Netherlands: Elsevier Science Publishers.

Wild, M. (1996). Mental models and computer modeling. *Journal of Computer Assisted Reasoning, 12(1),* 10-21.

Windschnitl, M. & André, T. (1998). Using computer simulations to enhance conceptual change: The roles of constructivist instruction and student epistemological beliefs. *Journal of Research in Science Teaching, 35*(2), 145-160.

Wolters, C.A. (1996). Advancing achievement goal theory: Using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *Journal of Educational Psychology, 96*(2), 236-250.

Wolters, C.A. & Yu, S.L. (1996). The relation between goal orientation and students' motivational beliefs and self-regulated learning. *Learning & Individual Differences, 8*(3), 211-238.

Zaraza, R. & Fisher, D.M. (1999). Training system modelers: The NSF CC-STADUS and CCSYSTAIN projects. In N. Roberts, W. Feurzeig, & B. Hunter (Eds.), *Modeling and Simulation in Science and Mathematics Education* (pp. 38-69). Berlin: Springer Verlag.

Zhang, B., Wu, H., Fretz, E.B., Krajcik, J.S., Marx, R., Davis, E.A., & Soloway, E. (2002). *Comparison of modeling practices of experts and novice students using a dynamic, student-centered modeling tool.* Paper presented at the annual meeting of the National Association of Research in Science Teaching, New Orleans, LA.

Zigurs, I. & Buckland, B.K. (1998). A theory of task/technology fit and group support systems effectiveness. *Management Information Systems Quarterly, 22*(3), 313-334.

Zohar, A. (1995). Reasoning about interactions between variables. *Journal of Research in Science Teaching, 32*(10), 1039-1063.

AUTHOR INDEX

APPENDIX A

Coding scheme for Reasoning processes.

Type of reasoning process (i.e., what are they doing?)	
Guiding by experimenter	The experimenter provides guidance to the students. The scoring starts with a question of one of the students that is addressed directly to the experimenter or by a spontaneous utterance from the experimenter
Evaluate	Students positively/ negatively evaluate an element(s) in relation to their model. Students make a (elaborate) value judgment on a modeling element
Explain	Students explain to each other how elements within their model work or why they were included. An explanation must be preceded by a clear-cut question of one of the students
Quantify	Students talk about quantifying or specifying a quantity or relation within their model
Inductive Reasoning	Students elaborate upon/ about elements within or with respect to their model (involves mainly qualitative reasoning)
Analyze	Students interpret modeling elements. Or identify factors that may be relevant/ included in their model
Other processes	
Read & para-phrase	Students read or paraphrase model elements
Off task	Students talk about subjects unrelated to the assignment at hand
Other	Other categories that are not included in the 'type of process' analysis scheme (only use this code when no other process-code can be applied!). Also inedible murmur is coded as other

Focus (i.e., what are they talking about?)	
Quantity	Quantity (i.e., constant or stock). In case students are specifying an auxiliary without talking about the relation that is implied
Relation	(Not yet implemented) relation/ interaction between quantities. In case students are specifying an auxiliary and talk about the relation that is implied
Model fit/ model output	Fit between the model output and experimental data. Students have to explicitly mention (the extent of) model fit or output the model generates (i.e., the model graphs or table)
Data points	Data points/ data graph. Students have to explicitly talk about the data.
Model structure	Structure of the model at hand, how the quantities are (visually) linked (i.e., visual structure). How the quantities in the constructed model are causally linked to each other (i.e., causal structure). Or how their constructed model works over time (i.e., dynamics). *Students have to explicitly talk about their model at hand:* When students talk about more than one relation in their model
Modeling actions/ the tool	Talk about/ mentioning modeling actions: what the students are doing. Or the students are trying to find out how Powersim (i.e., tools, buttons, formalism etc.) works
The assignment	Students talk about the assignment

Argumentation (where do they refer to during modeling?)	
None	No reference to modeling element or knowledge type
Knowledge	
Physics knowledge	Use of terminology, concepts (i.e., units, quantities), formula's common in physics
Mathematics knowledge	Use of terminology, concepts, formula's common in mathematics
Experiential knowledge	Knowledge from everyday experience is used
Model components	
Correspondence model graph and data	Students refer to (the extent of) correspondence between model output and experimental data
Experimental data	Experimental data (i.e., data points/ graph)
Html-documents[13]	Information about the black sphere assignment provided in the html-documents

13 This code was added for analyzing the protocols obtained in the studies reported in Chapters 4 and 5. In these studies students worked on the black sphere task within the Co-Lab environment.

APPENDIX B

Questionnaire used to assess students' epistemological understanding of models and modeling.

In the next statements we ask you for your view on models and modeling (for example the modeling of friction forces). At every statement you have to indicate whether you agree or disagree with this statement and provide also a short explanation of why you agree or disagree. At some points, you will be presented with an open question, try to provide a clear answer.

1) How do you define a model?
   Try to mention everything you can think of when you hear the words 'model' and 'modeling'. Use whole sentences in your answer.

2) Why are models used in science?

3) The model of the scientist in the assignment is wrong
   Agree/ Disagree, because:

4) Only a small part of reality can be described with this model
   Agree/ Disagree, because:

5) How do you decide if a model is correct?

6) Scientist should test their model(s)
   Agree/ Disagree, because:

7) It is impossible to decide which model is the best
   Agree/ Disagree, because:

8) The scientist in the assignment has to perform more measurements/ more experiments in order to construct a better model
   Agree/ Disagree, because:

9) Another scientist (Dr. Schintaler) thinks there is no way to decide which model can describe the given data the best. What do you think?

10) Models are useful in order to better understand phenomena in physics (for example, friction forces or resonance)
    Agree/ Disagree, because:

APPENDIX C

Scoring of students' epistemological understanding of models and modeling.

Nature of models

Level	*Description*	*Examples*
1.	Students view a model as a creation of an object on the basis of data	'A model is created by means of the data you have measured'
	Students view a model as a (vague) scheme, concept map or drawing of variables. Students focus on visual aspects of models	'[A model is] an overview of a situation where a lot of things influence other things' '[A model is a] scheme where you can change things and where you can see what the output is if you add or change something'
2.	Students view a model as a (simplistic) representation of reality. Students acknowledge that a model represents something else (i.e., reality)	'With a model you can represent reality' 'A model is a simplified representation of reality in which you can make estimates'
	Students view a model as a simulation (i.e., representation of reality), formula, or as a computational tool. The concern is with how the model works.	'[A model is] a simulation of reality' 'You can simulate different things, such as kinetics, heating of objects, etc.' 'A model is an extended formula with which you can predict measurements' 'By employing several variables and formulas, we can build a simulation of a particular situation or event'
3.	Students acknowledge that a model draws on assumptions, is a theory of reality. Students view a model as a depiction of an idea of a researcher.	'A model is an image of reality, measurements can be explained with a model […] everything is set up from a theoretical and fundamental (and simplistic) basis'

Purposes of models

Level	*Description*	*Examples*
1.	Students think models are used to measure things, find out answers, collect data (students conflate data with model output)	'[With a model] you do not have to measure everything and you can include more' 'In science models are used to better notice and discover things'
2.	Students view models to be of use for showing how something works (i.e., provide an overview of a phenomenon) or to show dependencies between variables	'Models are used to gain a better insight into difficult and complex issues' 'A model quickly provides relations between different variables'
	Students think models are useful to make research more efficient or to analyze data	'[Models are used] to conduct efficient research. You can easily add relations and change variables' 'Models can be used for the analysis of measurements'
3.	Students think models are useful for making predictions about a phenomenon (in service of improving the design of an artifact)	'[Models are used to] make predictions of particular scientific processes' '[Models are used to] make predictions of what happens when you change something in the design of an object/ apparatus'

Design and revision of models

Level	*Description*	*Examples*
1.	Students view model construction/ revision simply as building/ revising a model	'From a situation you make a model' 'Modeling is to construct and change a model'
2.	Students define model revision as model-fitting	'Modeling = playing with values and formulas until the output of the model matches the measured values' '[Modeling is] changing values and formulas on the computer with more accurateness'
	Students view model construction/ revision as changing or adding variables or relations to a model	'[Modeling is] deciding which factors are involved' 'Modeling is when you add several important factors to predict something'
	Students view modeling as a process of specifying variables (and monitoring the output of their model)	'[Modeling is] putting formulas together, by which you enable more factors to play a role in the output' '[Modeling is] examining what the effect is of changing a variable'
	Students view modeling as a process of simulating a phenomenon	'[Modeling is] to simulate a scientific phenomena'
3.	Students try to keep their model as simple as possible in order to obtain more predictive validity (i.e., simplistic elegance in constructing a model)	'By using very simple models, you may use them for multiple goals'
	Students think that model revision occurs by rethinking one's data and their implication as well as the purpose of the models	'If the researcher uses new measurement methods, he may come to other conclusions regarding his/ her model and may ultimately change his/ here model'

Evaluation of models		
Level	*Description*	*Examples*
1.	Students report that they vaguely check their model (at face value or compare the model with reality)	'If the model corresponds with reality' 'If the model is reliable'
2.	Students think that models are evaluated by comparing it with empirical data	'If it matches the measurements' 'You have to compare the data from the model with data you have obtained and the one that fits the best is correct'
	Students view model evaluation as examining whether all relevant variables are included in the model	'Checking whether you have taken all factors into account'
3.	Students report that they evaluate models by comparing it with data obtained from multiple measurements or with the opinion of other researchers	'You have to check whether the output of the model corresponds with data obtained from several measurements'

APPENDIX D

Task Co-Lab Black sphere.

There has been a lot of publicity about the earth's changing climate. Scientists all around the world are trying to understand what is going on, in order to predict what will happen next, or maybe more importantly, to give advice on what to do about it. The earth's climate is a very complex system, however, and even with all those scientists working on it, uncertainties remain. In such a situation scientists usually begin by making all kinds of simplifications. They first try to understand this simplified system, for instance by making a computer model. Then they use the computer model of the simplified system to make predictions about the real earth. Then they compare their predictions to empirical data, and consider what refinements are most needed.

In this module, you'll take a similar approach. We have made a very simplified small scale version of the earth and the sun: In our laboratory we have ignored the differences between oceans, forests and deserts. All that remains of the earth is a black sphere and at some distance you'll find a strong light, which takes the function of the sun. Not too similar to the world we live in, you'll say, and you are right. Nevertheless, you can investigate how the earth's temperature responds to changes of solar activity, and what the effects will be if the earth's surface changes color, for instance because it becomes covered with ice. Once you have got a computer model to make proper predictions about this simplified situation, you'll have discovered the basic model structure that underlies even the most advanced climate models today.

To summarize, your goal in this module will be to build a model that can predict the temperature of a black sphere (the barren earth) after being exposed to a source of light (the sun) for some while. To assist you, we provide you with an initial but still incomplete model. This model shows that the energy content of the earth is influenced by an energy inflow (incident radiation from the sun) and an energy outflow (outgoing radiation). Your goal is to extend this model in such a way in that it will provide a good match with the data you obtained from the simulation of the black sphere. In order to fulfill this goal, you'll need to find out first which factors play a role, and how they depend on each other.

Good luck!

SUMMARY

In secondary science education, many complex real world phenomena, such as ecosystems, climate change or mechanical oscillations, can only be discussed at a very basic level. Describing and predicting how these complex phenomena behave requires a level of formal reasoning and mathematical skill that is beyond the reach of most of secondary school students. Computer models overcome these problems, since the computer takes over the task of solving differential equations, allowing students to experiment with and to visualize the behavior of the phenomenon.

Several claims have been put forward with respect to the educational value of computer-based modeling. First, computer models are tools that enable students to externalize and to reflect on their own understanding of the phenomenon being modeled. Second, the act of modeling offers students with the opportunity to think scientifically about the behavior of phenomena. Finally, computer models can serve to make ideas accessible to criticism from peers, thus supporting collaboration between students. Despite these optimistic expectations, the process of computer-based modeling poses a complex task for secondary students, and without proper support they are unlikely to succeed. In order to shed light on students' needs and to eventually provide suggestions for scaffolding, a thorough understanding of the reasoning processes students employ during computer-based modeling is needed.

Main purpose of this dissertation was to investigate the nature of students' reasoning during computer-based modeling. In addition, we aimed at examining how characteristics of the students and of the modeling environment affect students' reasoning. More specifically, we investigated whether and how students' reasoning is influenced by: a) their epistemological understanding of models, b) the mode they use to communicate (i.e., chat versus face-to-face communication), and c) their motivation.

In the studies reported in Chapters 2 through 5 the following four research questions were addressed:

1) Which reasoning processes, employed by students during a computer-based modeling task, need support?
2) What is the relation between students' epistemological understanding of models and modeling and their reasoning during a computer-based modeling task?
3) What is the effect of chat communication versus face-to-face communication on students' modeling activities, on their reasoning during modeling, and on the quality of their model?
4) What is the relation between students' motivation and their reasoning during a computer-based modeling task and is the relation between students' motivation and model quality mediated by their reasoning?

These studies involved students from eleventh-grade pre-university education, with a major in science (age range 16-18). They had no prior experience with computer models. Students worked in dyads on a task to extend an incomplete computer model. Students could test their model against given empirical data. In Chapters 2 and 3 students worked on a model of the distance covered by an ice-skater using the systems-oriented modeling tool Powersim (Byrkness & Myrtveit, 1997). In Chapters 4 and 5 students' worked on a model of the temperature of an irradiated black sphere using a similar modeling tool implemented in the Co-Lab environment (Van Joolingen, et al., 2005).

CHAPTER 1

Chapter 1 reviews the claims put forward by educational researchers with regard to the educational qualities of computer-based modeling, and the difficulties found in previous studies. Despite high expectations, the process of constructing and testing computer models appears to involve complex thinking and the expected benefits are unlikely to occur without appropriate support. We argue that in order to identify students' needs and difficulties, it is important to investigate the nature of their reasoning during modeling. In addition, the degree to which computer-based modeling leads to deep reasoning and ultimately to better achievement may be dependent on characteristics of the student as well as on features of the learning environment. Our choice to investigate the influence of students' epistemology, communication mode and students' motivation on students' reasoning is explained. Finally, an overview is presented of the four research questions that were addressed in the studies reported in this dissertation.

CHAPTER 2

Purpose of Chapter 2 is to determine the nature of secondary students' reasoning during a scientific computer-based modeling task. We argue that in order to gain insight into these processes, we need to investigate the occurrences as well as the quality of the relevant types of reasoning processes. In this chapter we addressed the following main research question:

Which reasoning processes, employed by students during a computer-based modeling task, need support?

This question encompasses the following sub questions:

1) What features are relevant to describe students' reasoning processes during computer-based modeling?
2) What distinguishes successful from less successful students?
3) Which reasoning processes are difficult for students to perform?

To answer our first sub question, we started from a framework of global reasoning processes identified in the literature, and then refined this framework on the basis of a qualitative analysis of verbal recordings of the collaborating dyads. This resulted in an analysis scheme with the following dimensions: a) type of reasoning process, b) topic focus, and c) type of argumentation. Reasoning processes like analyzing or explaining mostly involve several turns by both partners in a dyad. Therefore, the unit of analysis was the process-episode level, an episode being a period of coherent continuous talk on a single issue, rather than single utterances. Verbal transcripts were segmented into episodes and scored using the analysis scheme. The analysis scheme proved a useful tool to assess the occurrence as well as the quality of students' reasoning processes.

The second sub question, about differences between more and less successful modelers, was first answered qualitatively on the basis of three case studies, after which these answers were corroborated in a quantitative analysis. In the qualitative analysis, we found that students who had built models of high quality (i.e., the more successful students) tended to justify their reasoning in terms of both experiential and physics prior knowledge. The less successful students, in contrast, spent much of their time on model fitting behavior (that is, they were manipulating individual quantities to make their model fit the empirical data). The quantitative analysis confirmed a significant positive correlation between model quality and inductive reasoning with reference to prior knowledge; and a significant negative correlation between model quality and quantifying quantities without argumentation (indicative of model fitting behavior).

With respect to the third sub question, about which are the most difficult reasoning processes, we identified difficulties at several levels. At *the level of task perception*, we found frequent evidence of model fitting behavior, indicating that students had difficulties with using the model as a means to describe and predict the behavior of a complex phenomenon. At *the content level*, students had difficulties with comprehending interaction between variables and with relating their own prior knowledge to the phenomenon being modeled. Finally, at *the level of the tool*, we found that, even though students had received a tutorial in system dynamics modeling, they still had difficulties in grasping the formalism used.

From these findings we conclude that appropriate support should be designed to scaffold students' reasoning during modeling on all three levels at which they were found to encounter difficulties. Scaffolds we suggest are a) encourage students to activate their prior knowledge, b) articulate specific modeling (sub-) goals, c) enable students to test their model against multiple datasets, d) ask students to build models of phenomena they already have knowledge of, e) let students first start to model qualitatively, and f) provide a top-down instruction in which students learn to comprehend the dynamic behavior of their model.

CHAPTER 3

The study reported in Chapter 3 serves two purposes. First, we investigated the level of secondary students' epistemological understanding of models and modeling. We took students' epistemological understanding of models and modeling to encompass their views regarding: a) the nature of models, b) the purpose of models, c) the evaluation of models, and d) the design and revision of models. To assess students' epistemological understanding, we employed the three general levels defined in the studies of Carey & Smith (1993) and Grosslight et al. (1991). At level 1, models are viewed as simple copies of reality. Level 2 involves the view that models are created for a specific purpose and that the model no longer must exactly correspond with the phenomenon being modeled. Level 3 is characterized by the idea that models are constructed in service of developing and testing ideas, rather than replicating reality. We determined the level of students' epistemological understanding of models and modeling for each of the four categories mentioned. In contrast to previous studies, the assessment of students' epistemological understanding was framed within our particular task context (i.e., a computer-based modeling task), in order to gain insight in their contextualized understanding as relevant to this task, rather than their generalized opinions.

The second aim of this study was to examine how the level of students' epistemological understanding is related to the level of their reasoning. We expected that a sophisticated epistemology (i.e., a level 3 understanding) is related to the employment of more deep reasoning processes and fewer surface reasoning processes during modeling. On the basis of the analysis scheme that was described in Chapter 2, we conceptualized deep reasoning as process-episodes in which students are elaborating on the modeling task and connect to knowledge they have available. Episodes in which students employ unelaborated processes without referring to available knowledge were labeled as surface reasoning.

Our findings indicate that, in general, the students held a moderate (i.e., level 2) epistemological understanding of models and modeling. Students reported that models are (simplified) representations of reality and that the processes of designing and revising models involve adding, deleting and specifying variables or relations. In addition, they stated that models are evaluated by comparing them with empirical data or by deciding whether all relevant variables are included in the model. With respect to students' epistemological understanding concerning the purpose of models, most students evenly had a level 2 or level 3 score. Students either reported models to be of use in providing an overview of the variables and the relations between them (i.e., level 2) or that models are useful in making predictions (i.e., level 3). In general, only few students held a level 1 epistemological understanding.

In answering the second research question we found a significant positive relation between students' epistemological understanding and deep reasoning during modeling. Furthermore, we found a significant negative relation between students' epistemological understanding and surface reasoning during modeling.

Based upon these findings, we hypothesized that students' epistemological understanding of models will have an impact on the nature of their reasoning during modeling. Consequently, appropriate scaffolds should be developed focusing on enhancing students' epistemological understanding of models and modeling.

CHAPTER 4

In the study reported in Chapter 4, we investigated the differential impact of face-to-face communication versus chat communication on: a) students' activities during modeling, b) students reasoning (i.e., deep and surface reasoning) during modeling, and c) the quality of the models students constructed.

Although the availability of chat makes collaboration between dispersed groups possible, when compared to direct face to face communication, the set of modalities by which learners can communicate is reduced. This raises the question of what the impact is of communication through a chat tool on students' performance during modeling compared to face-to-face communication. In this study, we addressed the following main research question:

What is the effect of chat communication versus face-to-face communication on students' modeling activities, on their reasoning during modeling, and on the quality of their model?

In other studies, students who use chat are found to be constrained in their interactions, because of a lack of communication cues. We expected that this constraint of chat communication may either hinder students during modeling or may pressure them to increase the efficiency of their modeling. Two alternative hypotheses were considered regarding the consequences of face-to-face communication versus chat communication on students' performance: 1) Students' in the face-to-face condition will score higher on model quality and will relatively spend less time on surface reasoning and relatively more time on deep reasoning compared to students in the chat condition. No significant difference is expected between the two conditions on the number of modeling activities performed. 2) Students in the face-to-face condition will score lower on model quality and will relatively spend more time on surface reasoning and relatively less time on deep reasoning compared to students in the chat condition. In addition, students in the face-to-face condition will perform significantly more modeling activities than students in the chat condition.

We found that students in the face-to-face condition spent significantly more of their time on surface reasoning and relatively less time on one of the six processes that were categorized as deep reasoning (i.e., inductive reasoning and reference to components) compared to students in the chat condition. Finally, we found that the face-to-face group performed more modeling activities (such as saving or executing a model) compared to the chat group. Although students in both conditions had similar scores on model quality, we conclude that, in general, results were more in accordance with hypothesis 2. The expectation of hypothesis 2, that students in the chat condition will compress their communication resulting in more efficient modeling compared to the face-to-face condition was further supported in additional analyses.

In addition, we found that students who communicated using chat relied more on the external model representation to compensate for the absence of face-to-face modalities and that communicative work, thus, also happens via the components present in the shared representation.

CHAPTER 5

The purpose of the study reported in Chapter 5 was to test a conceptual model of relations among motivation, reasoning and model quality of students working within a computer-based modeling task. More specifically, we investigated whether students' achievement goal orientation and self-efficacy influences their reasoning during modeling, and whether the relation between motivation and model quality is mediated by students' reasoning. Students' achievement goal orientation involves their beliefs regarding learning success and their motives to engage in learning activities, such as modeling. In this study, we focused on two divergent types of achievement goal orientations: mastery-approach goal orientation and performance-avoidance goal orientation. A mastery-approach goal orientation involves striving to develop one's skills and abilities, to advance one's learning, to understand the material, or to complete a task. Performance-avoidance goal orientation involves the avoidance of failure and the avoidance of the demonstration of low ability. Self-efficacy was defined as the learner's belief regarding their performance capabilities in a particular domain.

We expected that mastery-approach goal orientation and self-efficacy are positively related to deep reasoning during modeling, whereas performance-avoidance goal orientation is positively related to the employment of surface reasoning processes. In addition, the relation between students' motivation and model quality was expected to be mediated by the level of students' reasoning.

In contrast to most studies in the field of motivation, we did not assess students' reasoning based on self-report measures, but rather we based our assessments on the process observations as described in the previous chapters.

Path analyses revealed that self-efficacy and mastery-approach goal orientation were significantly positively related with each other and with students' employment of deep reasoning processes during modeling. In addition, both self-efficacy and mastery-approach goal orientation were positively related to model quality, but the relation between these variables was mediated by students' use of deep reasoning processes. No significant relation was found between performance-avoidance goal orientation and surface reasoning. In addition, the relation between surface reasoning and model quality was not significant.

An educational implication of this study may be that strategies that promote a mastery-approach goal orientation and advance students' self-efficacy lead to deep reasoning during modeling and ultimately to a higher achievement.

CHAPTER 6

In Chapter 6 we present and discuss the main results obtained from the studies reported in this dissertation. In addition, some general issues that can be raised with respect to the scope of our research and the research methodology chosen for our studies are discussed. Finally, implications for science education are suggested.

With respect to our *measurement of students' reasoning processes* we argue that, although students' verbalizations do not necessarily reveal their covert reasoning processes, the assumption of our method of analysis was that it allowed us to build valid inferences about students cognitive processing based on what students articulate during modeling. In addition, we discuss the risk of over- or underestimating the quality of students' reasoning in those cases where the contributions of students within a dyad are unevenly distributed. This risk remained largely hypothetical however, because in most dyads students were found to contribute in an even way. Finally, we discuss the assumption underlying our choice for examining students' level of reasoning (i.e., deep and surface reasoning) in the studies reported in Chapters 3 through 5. We state that deep processing creates the optimal conditions for achievement in a variety of subject-matter areas, whereas surface processing is often related to decreases in achievement. In addition, our aim in these chapters was to elaborate on the findings of previous studies within the fields of students' motivation and epistemology that have also employed this dichotomy. In these studies the link between these constructs and level of students' processing during task performance is often made but only rarely investigated.

We discuss *the modeling tasks* used in Chapters 2 through 5. In contrast to most studies in the field of computer-based modeling, students in these studies were provided with the opportunity to test their computer model against empirical data. Also, we argue that the modeling task necessitated students to employ deep reasoning processes in order to complete the task to a satisfactory end.

Regarding *the scope of our research*, we maintain that the findings that were obtained from the studies reported in this dissertation can be generalized to all other forms of collaborative modeling which takes place within a shared computer-based environment. It has to be noted, though, that because our measures for students' motivation and epistemology were contextualized, conclusions and implications of this dissertation are restricted to this context. In addition, the impact of communication mode on students' reasoning was also reported to be highly dependent on task characteristics, making it difficult to generalize to other task contexts.

Also, we discuss the validity of assessing students' motivation and epistemology at only one point in time instead of at different moments during modeling. We suggest that future research should examine whether and how changes in students' motivation and/ or epistemology affects their reasoning during computer-based modeling.

We stress that it was not possible to determine whether students have gained more conceptual knowledge of the phenomenon being modeled or of system dynamics in general, or whether they have acquired more modeling skills. Future research should address the issue of whether and how students' reasoning affects what they learn.

Finally, we present suggestions for scaffolding students' reasoning based on the findings reported in this dissertation. Firstly, appropriate support should be designed to scaffold students' reasoning on all three levels (i.e., level of task perception, content level, and level of the tool) at which they were found to encounter problems. Secondly, students should be prompted to reflect on the nature, purpose and evaluation of models and on the design and revision of models in order to foster a productive epistemology. Thirdly, during science class a mastery-approach orientation can be promoted when, for instance, the teachers provides personalized comments or when the importance of understanding the phenomenon being modeled is emphasized. In addition, students' self-efficacy may be enhanced when the teacher offers support and encouragement. Finally, based on the findings from the study reported in Chapter 4, we hypothesize that the constraints imposed by chat on communication may have led students to be more concise and more task focused compared to students who communicated face-to-face during modeling. Although, more research in this area is needed, we concur that chat can be employed as an effective mode of communication in a complex computer-based modeling task.

The scaffolds we have suggested may enhance students' reasoning to some extent and they could be implemented within the context of a curriculum based on computer-based modeling. Still, we recommend that future research should investigate the influence of students' epistemology, motivation, and communication mode on students' reasoning during a whole curriculum to decide what scaffolding is needed in the long term.

SAMENVATTING

Veel complexe verschijnselen in de wereld om ons heen, zoals ecosystemen, klimaatsveranderingen of mechanische oscillaties, kunnen slechts op een sterk vereenvoudigd niveau worden behandeld in de exacte vakken van het secundair onderwijs. Het beschrijven en voorspellen van het gedrag van deze complexe verschijnselen vereist een niveau van rekenkundige en wiskundige vaardigheid dat veelal buiten het bereik van leerlingen ligt. Computermodellen kunnen deze beperkingen gedeeltelijk opheffen, doordat het oplossen van differentiaalvergelijkingen overgenomen wordt door de computer. Met behulp van modelleersoftware kunnen leerlingen experimenteren met hun model en kunnen ze het gedrag van het verschijnsel observeren door hun model te laten doorrekenen.

Met betrekking tot de onderwijskundige waarde van computermodelleren worden in de literatuur verschillende voordelen genoemd. Ten eerste kunnen leerlingen reflecteren op hun ideeën over het verschijnsel dat gemodelleerd wordt, aangezien zij zelf modellen kunnen maken en simuleren op de computer. Ten tweede leren leerlingen op een wetenschappelijke wijze te redeneren over het gedrag van een systeem. Tenslotte kan de samenwerking tussen leerlingen bevorderd worden doordat de modellen, die leerlingen hebben gemaakt, zichtbaar blijven op het computerscherm en als zodanig toegankelijk zijn voor kritiek van klasgenoten. Ondanks deze optimistische verwachtingen is computermodelleren een complexe taak die veel ondersteuning vergt. Om meer inzicht te verkrijgen in de moeilijkheden en behoeften van leerlingen en om uiteindelijk suggesties te kunnen doen voor dergelijke ondersteuning is een grondig begrip van de redeneerprocessen die leerlingen tijdens modelleren gebruiken noodzakelijk.

Het hoofddoel van dit proefschrift is om de aard van het redeneren van leerlingen tijdens computermodelleren te beschrijven. Daarnaast trachten wij na te gaan hoe kenmerken van de leerling en van de modelleertool het redeneren van leerlingen beïnvloeden. Meer specifiek onderzoeken wij of en hoe het redeneren van leerlingen beïnvloed wordt door: a) hun epistemologisch begrip van modellen, b) het medium waarmee ze communiceren (chat versus face-to-face communicatie) en c) hun motivatie.

In de hoofdstukken 2 tot en met 5 worden studies beschreven waarin de volgende vier onderzoeksvragen centraal staan:

1) Welke redeneerprocessen, die leerlingen gebruiken tijdens een computergestuurde modelleertaak, vereisen ondersteuning?
2) Wat is de relatie tussen het epistemologisch begrip van modellen en modelleren van leerlingen en de redeneerprocessen die zij gebruiken tijdens een computergestuurde modelleertaak?
3) Wat is het effect van chatcommunicatie versus face-to-face-communicatie op de modelleeractiviteiten van leerlingen, op de redeneerprocessen die zij gebruiken tijdens een computergestuurde modelleertaak en op de kwaliteit van hun model?

4) Wat is de relatie tussen de motivatie van leerlingen en de redeneerprocessen die zij gebruiken tijdens een computergestuurde modelleertaak en wordt de relatie tussen motivatie en modelkwaliteit gemedieerd door redeneerprocessen?

Deelnemers in deze studies waren leerlingen 5 VWO met het profiel Natuur & Techniek (leeftijd tussen 16 en 18 jaar). Leerlingen hadden geen ervaring met computermodellen. Ze werkten in tweetallen aan een taak waarin ze gevraagd werden een incompleet computermodel uit te breiden met behulp van empirische data die ze aangeboden kregen. In hoofdstukken 2 en 3 werkten leerlingen aan een modelleertaak waarin ze de door een uitrijdende schaatser afgelegde afstand moesten beschrijven met behulp van de modelleertool Powersim (Byrkness & Myrtveit, 1997). In hoofdstukken 4 en 5 werkten leerlingen aan een modelleertaak waarin ze de temperatuur van een zwarte bol moesten beschrijven, onder invloed van invallende warmtestraling. Deze taak voerden ze uit binnen de Co-Lab omgeving (Van Joolingen, et al., 2005).

HOOFDSTUK 1

Hoofdstuk 1 geeft een overzicht van de claims van onderzoekers ten aanzien van de onderwijskundige kwaliteiten van computermodellen. Daarnaast worden de moeilijkheden die in andere modelleerstudies zijn gevonden besproken. Ondanks alle hoge verwachtingen, vereist het construeren en testen van computermodellen complexe denkprocessen en kunnen de voordelen van modelleren alleen worden gerealiseerd als er geschikte ondersteuning wordt geboden. Wij stellen dat, om de behoeften en moeilijkheden van leerlingen te kunnen identificeren, het noodzakelijk is om de aard van de redeneerprocessen die zij gebruiken tijdens modelleren te onderzoeken. Daarnaast verwachten wij dat de mate waarin computermodelleren leidt tot diepe redeneerprocessen en tot een betere prestatie afhangt van kenmerken van de leerling en van eigenschappen van de leeromgeving. Onze keuze om het onderzoek te richten op de invloed van het epistemologisch begrip van leerlingen, het communicatiemedium en de motivatie van leerlingen op de redeneerprocessen van leerlingen wordt toegelicht. Tenslotte presenteren we de vier onderzoeksvragen die in de studies van dit proefschrift worden behandeld.

HOOFDSTUK 2

Het doel van hoofdstuk 2 is de aard van de redeneerprocessen te bepalen die leerlingen tijdens het uitvoeren van een computerondersteunde modelleertaak gebruiken. In dit hoofdstuk richten we ons op de volgende hoofdonderzoeksvraag:

Welke redeneerprocessen, die door leerlingen tijdens een computergestuurde modelleertaak gebruiken, vereisen ondersteuning?

Deze vraag omvat de volgende subvragen:

1) Wat zijn de relevante eigenschappen van redeneerprocessen van leerlingen tijdens het computermodelleren?
2) Wat onderscheidt succesvolle van minder succesvolle leerlingen?
3) Welke redeneerprocessen zijn voor leerlingen moeilijk uit te voeren?

Om de eerste subvraag te beantwoorden werd een analyseschema ontwikkeld op basis van beschreven processen in de literatuur. Door dit schema toe te passen op de verbale interacties van de samenwerkende tweetallen werd het schema in enkele iteratieve stappen verfijnd. Het resulterende analyseschema heeft de volgende dimensies: a) type redeneerproces, b) onderwerp en c) type argumentatie. Redeneerprocessen als analyseren of verklaren omvatten veelal een aantal achtereenvolgende uitingen door beide leerlingen in een tweetal. Om deze reden kozen we niet voor losse uitingen als eenheid van analyse maar voor het episodeniveau, waarbij een episode een korte periode behelst waarin een coherent en continu gesprek plaatsvindt over één enkel onderwerp. Verbale transcripten werden gesegmenteerd in episodes en gescoord met behulp van het analyseschema. Het analyseschema bleek een bruikbaar instrument te zijn om zowel het vóórkomen als de kwaliteit van de redeneerprocessen die leerlingen gebruiken, te meten.

De tweede onderzoeksvraag, over het verschil tussen meer en minder succesvolle modelleerders, werd eerst kwalitatief beantwoord op grond van drie casebeschrijvingen. Daarna werden de bevindingen getoetst in een kwantitatieve analyse van de gehele dataset. In de kwalitatieve analyse vonden we dat leerlingen die modellen van een hoge kwaliteit hadden geconstrueerd (de meer succesvolle leerlingen) hun redeneerprocessen rechtvaardigden door te verwijzen naar hun ervaringskennis of naar hun natuurkundekennis. De minder succesvolle leerlingen besteedden meer tijd aan zogenaamd ‘modelfit’ gedrag, dat wil zeggen dat deze leerling voornamelijk waardes van de individuele grootheden in hun model manipuleerden om de output van hun model zo goed mogelijk overeen te laten komen met de empirische data. De kwantitatieve analyse bevestigde een significante positieve correlatie tussen inductief redeneren met een verwijzing naar voorkennis en kwaliteit van het model; en een significante negatieve correlatie tussen kwantificeren van grootheden zonder argumentatie (indicatief voor modelfit gedrag) en kwaliteit van het model.

Met betrekking tot de derde subvraag, over welke redeneerprocessen complex zijn voor leerlingen, vonden we dat leerlingen op drie niveaus moeilijkheden ondervonden. Voor wat betreft het niveau van *taakperceptie*, vonden we bewijs voor modelfit gedrag. Dit suggereert dat leerlingen niet op zoek waren naar een conceptueel aanvaardbare verklaring voor het gedrag van het gemodelleerde verschijnsel. Op het niveau van de *inhoud*, vonden we dat leerlingen problemen hadden met het begrijpen van interacties tussen grootheden in hun model en met het relateren van hun relevante voorkennis aan het gemodelleerde verschijnsel. Tenslotte, op het niveau van de *modelleertool*, vonden we dat leerlingen, ook na het doorwerken van een tutorial over systeemdynamisch modelleren, het modelleerformalisme nog niet goed begrepen.

Uit deze bevindingen concluderen we dat geschikte ondersteuning ontworpen zou moeten worden om het redeneren van leerlingen tijdens computermodelleren te ondersteunen op elk van de drie niveaus waarop leerlingen moeilijkheden ondervinden. Ondersteuningsmaatregelen die wij voorstellen zijn: a) stimuleer leerlingen om hun voorkennis te gebruiken tijdens het modelleren, b) benoem specifieke modelleer (sub-) doelen, c) geef leerlingen de mogelijkheid hun model te toetsen aan meerdere datasets, d) vraag leerlingen modellen te maken van verschijnselen waarvan ze al wat kennis hebben, e) laat leerlingen eerst een kwalitatief model maken en f) biedt een topdown instructie aan waarin leerlingen het dynamische gedrag van hun model beter leren begrijpen.

HOOFDSTUK 3

De studie die beschreven wordt in hoofdstuk 3 dient twee doelen. Ten eerste onderzoeken we het niveau van het epistemologisch begrip van leerlingen over modellen en modelleren. Het epistemologisch begrip van leerlingen omvat hun beeld van: a) de aard van modellen, b) het doel van modellen, c) de evaluatie van modellen en d) het construeren en reviseren van modellen. Om het epistemologisch begrip van leerlingen te kunnen bepalen, gebruikten we de drie algemene niveaus die in de studies van Carey & Smith (1993) en Grosslight et al. (1991) zijn gedefinieerd. Leerlingen die op niveau 1 zitten, zien een model als een eenvoudige replicatie van de werkelijkheid. Niveau 2 omvat het beeld dat modellen gemaakt worden met een specifiek doel en dat het model niet exact overeen hoeft te komen met het verschijnsel dat gemodelleerd wordt. Niveau 3 wordt gekenmerkt door de perceptie dat modellen geconstrueerd worden om ideeën te ontwikkelen en te testen. We bepaalden het niveau van het epistemologisch begrip van leerlingen voor elk van de vier bovenstaande categorieën. Omdat epistemologisch begrip vaak contextgebonden is, werd, in tegenstelling tot andere studies, het epistemologisch begrip van leerlingen binnen de specifieke taakcontext (de computerondersteunde modelleertaak) gemeten.

Het tweede doel van deze studie is om de relatie tussen het niveau van redeneren van leerlingen tijdens het computermodelleren en hun epistemologisch begrip te onderzoeken. Onze verwachting was dat een verfijnde epistemologie (een niveau 3 begrip) gerelateerd is met het gebruik van meer diepe redeneerprocessen en minder oppervlakkige processen tijdens het modelleren. Op grond van het analyseschema dat in hoofdstuk 2 wordt beschreven, conceptualiseerden we diepe processen als episodes waarin leerlingen elaboreren op de modelleertaak en waarin ze hun eigen kennis gebruiken. Episodes waarin leerlingen niet op de materie elaboreren en geen eigen kennis gebruiken werden als oppervlakkige processen gelabeld.

Onze resultaten laten zien dat meer dan de helft van de leerlingen een gemiddeld (niveau 2) epistemologisch begrip hadden van modellen en modelleren. Leerlingen rapporteerden dat modellen vereenvoudigde representaties van de werkelijkheid zijn en dat het construeren en reviseren van modellen het toevoegen, verwijderen en kwantificeren van grootheden omvat. Ook meldden ze dat modellen worden geëvalueerd door de model output met de empirische data te vergelijken of door te beslissen of alle relevante variabelen in het model zijn opgenomen. Voor wat betreft het epistemologisch begrip van leerlingen betreffende het doel van modellen blijkt dat de meeste leerlingen gelijkmatig verdeeld waren over niveaus 2 en 3. Leerlingen rapporteerden dat modellen gebruikt worden om een overzicht te geven van de variabelen en de relaties daartussen (niveau 2) of dat modellen bruikbaar zijn voor het maken van voorspellingen (niveau 3). In het algemeen vonden we dat slechts weinig leerlingen een niveau 1 epistemologisch begrip hadden van modellen en modelleren.

Bij het beantwoorden van de tweede onderzoeksvraag vonden we een significante positieve correlatie tussen het gebruik van diepe redeneerprocessen tijdens het modelleren en het epistemologisch begrip van leerlingen. Daarnaast vonden we een significante negatieve correlatie tussen het gebruik van oppervlakkige redeneerprocessen en het epistemologisch begrip.

Op grond van deze bevindingen, stellen wij de hypothese dat het epistemologisch begrip van leerlingen invloed zal hebben op de aard van de redeneerprocessen die ze gebruiken tijdens het computermodelleren. Er zou daarom ondersteuning moeten worden ontwikkeld die gericht is op het bevorderen van het epistemologisch begrip van leerlingen.

HOOFDSTUK 4

In de studie die wordt beschreven in hoofdstuk 4 onderzoeken wij de verschillende invloed van face-to-face-communicatie en chatcommunicatie op: a) de modelleeractiviteiten die leerlingen uitvoeren, b) de redeneerprocessen (diepe en oppervlakkige processen) die leerlingen gebruiken en c) de kwaliteit van het model dat leerlingen construeren.

Chat maakt samenwerking tussen groepen leerlingen die fysiek van elkaar gescheiden zijn mogelijk. Een nadeel van chat in vergelijking met face-to-face-communicatie is dat de set modaliteiten waarmee leerlingen kunnen communiceren beperkt is. Dit roept de vraag op wat de invloed is van chatcommunicatie op de modelleerprestaties van leerlingen in vergelijking met face-to-face-communicatie. In deze studie staat de volgende onderzoeksvraag centraal:

Wat is het effect van chatcommunicatie versus face-to-face-communicatie op de modelleeractiviteiten van leerlingen, op de redeneerprocessen die zij gebruiken tijdens een computergestuurde modelleertaak en op de kwaliteit van hun model?

In andere studies werd gevonden dat leerlingen die chat gebruiken, beperkt worden in hun interactie door een gebrek aan communicatie-cues. We verwachtten dat deze beperking van chatcommunicatie ertoe kan leiden dat leerlingen worden gehinderd tijdens het modelleren óf juist dat leerlingen hierdoor gedwongen worden om efficienter te modelleren. Deze overwegingen leiden tot twee alternatieve hypothesen over het effect van face-to-face communicatie versus chatcommunicatie: 1) Leerlingen in de face-to-face-conditie zullen hoger scoren op modelkwaliteit, zullen relatief minder tijd besteden aan oppervlakkige redeneerprocessen en relatief meer tijd besteden aan diepe redeneerprocessen in vergelijking met de leerlingen in de chatconditie. De twee condities zullen niet significant van elkaar verschillen voor wat betreft het aantal modelleeractiviteiten die leerlingen uitvoeren. 2) Leerlingen in de face-to-face-conditie zullen lager scoren op modelkwaliteit en zullen relatief meer tijd besteden aan oppervlakkige redeneerprocessen en relatief minder tijd besteden aan diepe redeneerprocessen in vergelijking met de leerlingen in de chatconditie. Daarnaast, zullen de leerlingen in de face-to-face-conditie meer modelleeractiviteiten uitvoeren dan leerlingen in de chatconditie.

We vonden dat leerlingen in de face-to-face-conditie significant meer van hun tijd besteedden aan oppervlakkige redeneerprocessen en relatief minder tijd besteedden aan één van de zes diepe redeneerprocessen (inductief redeneren en verwijzen naar modelleercomponenten) in vergelijking met leerlingen in de chatconditie. Tenslotte vonden we dat leerlingen in de face-to-face-conditie meer modelleeractiviteiten (zoals het saven en het simuleren van hun model) uitvoerden dan leerlingen in de chatconditie. Hoewel de kwaliteit van de modellen in beide condities niet significant verschillend was, zijn de resultaten meer in overeenstemming met hypothese 2. De verwachting van hypothese 2, dat leerlingen in de chatconditie voor het gebrek aan communicatie cues compenseren door hun communicatie ‘samen te persen’ wat uiteindelijke leidt tot efficiënter modelleergedrag, werd ook ondersteund in de additionele analyses.

Daarnaast vonden we dat de communicatie in de chatconditie niet alleen verliep via de chatbox maar ook via de gedeelde modelrepresentatie.

HOOFDSTUK 5

In hoofdstuk 5 wordt een conceptueel model getest, voor de relaties tussen motivatie, redeneerprocessen en modelkwaliteit van leerlingen binnen de context van een computergestuurde modelleertaak. Meer specifiek onderzoeken wij of de doeloriëntatie en self-efficacy van leerlingen invloed hebben op de redeneerprocessen die zij gebruiken tijdens het computermodelleren en of de relatie tussen motivatie en modelkwaliteit gemedieerd wordt door redeneerprocessen. De doeloriëntatie van leerlingen behelst hun perceptie van leersucces en hun motieven om zich in te spannen voor een gegeven leertaak. In deze studie richtten wij ons op twee divergente doeloriëntaties: Mastery-approach goal orientation en performance-avoidance goal orientation. Een mastery-approach goal orientation omvat de motivatie om (leer-)vaardigheden en kennis te verwerven en te ontwikkelen en om het leermateriaal zo goed mogelijk te begrijpen en te verwerken. Performance-avoidance goal orientation is gericht op het vermijden van falen en van het laten zien van de eigen onbekwaamheid aan anderen. Self-efficacy wordt gedefinieerd als de mate waarin een leerling zichzelf binnen een bepaald kennisdomein bekwaam acht.

We verwachtten dat mastery-approach goal orientation en self-efficacy positief correleren met het gebruik van diepe redeneerprocessen tijdens het computermodelleren, terwijl we verwachtten dat performance-avoidance goal orientation positief correleert met het gebruik van oppervlakkige redeneerprocessen. Daarnaast was de verwachting dat de relatie tussen motivatie en modelkwaliteit gemedieerd wordt door het niveau waarop leerlingen redeneren.

In tegenstelling tot de meeste studies, waarin gekeken wordt naar de relatie tussen motivatie en verwerkingsprocessen op basis van zelfrapportage-vragenlijsten, maten wij het redeneren van leerlingen op basis van procesobservaties, zoals beschreven in de vorige hoofdstukken.

Pad-analyses lieten zien dat self-efficacy en mastery-approach goal orientation significant positief gerelateerd zijn met elkaar en met het gebruik van diepe redeneerprocessen tijdens het modelleren. Verder bleken zowel self-efficacy als mastery-approach goal orientation een positief verband te hebben met modelkwaliteit, maar de relatie tussen deze variabelen werd gemedieerd door het gebruik van diepe redeneerprocessen. Er werd geen significante relatie gevonden tussen performance-avoidance goal orientation en oppervlakkig redeneren. Ook was de relatie tussen oppervlakkig redeneren en modelkwaliteit niet significant.

Een implicatie voor de onderwijspraktijk zou kunnen zijn dat onderwijsstrategieen, die een mastery-approach goal orientation zouden stimuleren en die de self-efficacy van leerlingen zouden bevorderen, kunnen leiden tot het gebruik van meer diepe redeneerprocessen tijdens het modelleren en uiteindelijk tot een hogere prestatie.

HOOFDSTUK 6

In hoofdstuk 6 presenteren en bediscussiëren we de resultaten van de studies in de voorafgaande hoofdstukken uit deze dissertatie. Daarnaast bespreken we enkele algemene kwesties met betrekking tot de mate van generaliseerbaarheid van ons onderzoek en met betrekking tot de onderzoeksmethode die we voor onze studies hebben gekozen. Tenslotte worden de implicaties voor de onderwijspraktijk besproken.

Het door ons ontwikkelde meetinstrument voor de redeneerprocessen van leerlingen bleek geschikt om, op basis van wat leerlingen tijdens het modelleren zeggen, hun cognitieve processen af te kunnen leiden. Verder bespreken we het risico van het over-of onderschatten van de kwaliteit van de redeneerprocessen van leerlingen in die gevallen waarin de bijdragen van leerlingen binnen een tweetal ongelijk is verdeeld. Dit risico bleek grotendeels hypothetisch, aangezien in de meeste tweetallen leerlingen in gelijke mate bijdroegen tijdens het modelleren. Tenslotte bespreken we de aanname die aan onze keuze voor het onderzoeken van het niveau van redeneren (diep en oppervlakkig redeneren) van leerlingen in de hoofdstukken 3 tot en met 5 ten grondslag ligt. We veronderstellen dat diepe processen de optimale voorwaarden scheppen om te presteren in verschillende vakgebieden, terwijl oppervlakkige processen vaak geassocieerd worden met lage prestaties. Daarnaast was het ons doel om dieper op de bevindingen uit andere studies, die ook deze dichotomie toepassen, in te gaan. In deze studies worden de relaties tussen epistemologie of motivatie enerzijds en het niveau van redeneren tijdens het uitvoeren van een taak anderzijds vaak gesuggereerd maar slechts zelden onderbouwd.

We behandelen *de modelleertaak* die we in de hoofdstukken 2 tot en met 5 gebruiken. In tegenstelling tot de meeste studies in het veld van computergestuurd modelleren, kregen leerlingen in onze studies de mogelijkheid om hun computermodel te testen aan de hand van empirische data. Ook stellen we dat de modelleertaak zodanig was ontworpen dat leerlingen diepe redeneerprocessen moesten toepassen om de taak bevredigend af te kunnen ronden.

Met betrekking tot *de mate van generaliseerbaarheid van ons onderzoek* betogen we dat de bevindingen van onze studies gegeneraliseerd kunnen worden naar andere vormen van samenwerkend modelleren binnen een gedeelde computeromgeving. De conclusies en implicaties van ons onderzoek in deze dissertatie zijn dan ook louter van toepassing op deze context, aangezien onze metingen van de motivatie en het epistemologisch begrip van leerlingen gecontextualiseerd zijn. Ook is de invloed van het communicatiemedium op de redeneerprocessen van leerlingen sterk afhankelijk van kenmerken van de specifieke leertaak, waardoor het moeilijk is om te generaliseren naar andere taakcontexten.

We bespreken ook de validiteit van het meten van de motivatie en het epistemologisch begrip van leerlingen op één specifiek moment tijdens het onderzoek, in plaats van dat deze constructen op verschillende momenten tijdens het modelleren werden gemeten. We suggereren dat in de toekomst onderzoek gedaan zou kunnen worden naar hoe veranderingen in de motivatie en/of het epistemologisch begrip van leerlingen de redeneerprocessen die zij tijdens het computermodelleren gebruiken beïnvloeden.

We benadrukken dat het niet mogelijk was om te bepalen in hoeverre leerlingen iets geleerd hebben over het gemodelleerde verschijnsel of over systeemdynamisch modelleren in het algemeen; en in hoeverre hun modelleervaardigheden toegenomen zijn. Toekomstig onderzoek zal moeten uitwijzen hoe het redeneren van leerlingen tijdens het computermodelleren een invloed heeft op wat zij uiteindelijk leren.

Op grond van onze bevindingen presenteren we tenslotte suggesties om het redeneren van leerlingen tijdens het modelleren te verbeteren. Ten eerste zou het redeneren van leerlingen ondersteund moeten worden op alle drie de niveaus (taakperceptie, inhoud en modelleertool) waarop ze moeilijkheden ondervinden. Ten tweede zouden leerlingen moeten worden aangespoord om te reflecteren op de aard, het doel, de evaluatie en op het construeren en reviseren van modellen om een productieve epistemologie te bevorderen. Ten derde, kunnen docenten een mastery-approach goal orientation bevorderen door bijvoorbeeld persoonlijke feedback te geven aan de leerlingen of door het belang van het begrijpen van het gemodelleerde verschijnsel te benadrukken. Daarnaast kan de self-efficacy van leerlingen worden vergroot, bijvoorbeeld door ze tijdens het modelleren aan te moedigen. Tenslotte verwachten wij, op grond van de bevindingen die worden beschreven in hoofdstuk 4, dat chat door de beperkingen die deze communicatievorm oplegt kan leiden tot een meer beknopte en taakgerichte communicatie.

De typen ondersteuning die we voorstellen kunnen het redeneren van leerlingen tijdens het computermodelleren tot een zekere hoogte bevorderen en ze zouden als zodanig kunnen worden geïmplementeerd in een curriculum waarin leerlingen gaan computermodelleren. Toch stellen we dat toekomstig onderzoek zich zou moeten richten op een nadere karakterisering van de invloed van epistemologie, motivatie en communicatie medium op het redeneren van leerlingen gedurende een heel curriculum, om te kunnen beslissen welk type ondersteuning op de lange termijn nodig is.

CURRICULUM VITAE

Patrick Sins was born on January 17th 1979 in Maastricht (The Netherlands), where he completed his secondary education at the Trichter College. He studied Cognitive Psychology at the University of Maastricht and graduated cum laude in 2001, specializing in Educational Psychology. His Master's thesis concerned two studies which examined the generation of biomedical and clinical knowledge in medical students and expert physicians and was awarded with the best thesis award of the Psychology faculty of the University of Maastricht.

In August 2001, Patrick started his PhD-project at the Graduate School of Teaching and Learning of the University of Amsterdam in the field of computer-based modeling. He was involved in the EC-funded project Co-Lab which aimed at designing, developing, and evaluating new learning arrangements based on the pedagogical ideas of collaborative inquiry learning and modeling. Patrick was the treasurer of the organizing committee of the JURE (Junior Researchers of EARLI) conference in Amsterdam in 2002. In 2003, he was a member of the executive committee of the JURE conference.

Currently, Patrick works on several research and educational development projects for the SCO-Kohnstamm Instituut, the Stichting ICT op School, and for the Graduate School of Teaching and Learning. In addition, he teaches various courses at the Graduate School of Teaching and Learning.