Conceptual understanding of conceptual modeling concepts: a longitudinal study among students learning to model

Dirk van der Linden[1,2,3], Hend*erik* A. Proper[1,2,3], and
Stijn J.B.A. Hoppenbrouwers[4,2,3]

[1] Public Research Centre Henri Tudor, Luxembourg, Luxembourg
{dirk.vanderlinden,erik.proper}@tudor.lu
[2] Radboud University Nijmegen, Nijmegen, the Netherlands
[3] EE-Team, Luxembourg, Luxembourg
[4] HAN University of Applied Sciences, Arnhem, the Netherlands
stijn.hoppenbrouwers@han.nl

Abstract. In this paper we will discuss our investigation into the conceptual understanding that students have of common concepts used for conceptual modeling (e.g., actors, processes, goals). We studied if and how those understandings may change over time during a student's progress through their academic curriculum. To do so, we performed a longitudinal study with a group of students starting computing and information science studies at Radboud University Nijmegen. We followed them from the beginning of their studies as they learned new theories, techniques, and languages for modeling. We focused on investigating whether their conceptual understandings changed as they became acquainted with new languages and techniques, and whether there were correlations between the introduction of such educational stimuli and changes in their conceptual understanding. We discuss the seeming lack of connection between these stimuli and such changes, and reflect on what this means for the training of people in conceptual modeling.

Keywords: conceptual modeling, conceptual understanding, longitudinal study, learning modeling, training of students, semantic differential

1 Research Objectives

We are interested in understanding whether students develop specific conceptualizations, or perhaps stronger, conceptual prejudices when it comes to conceptual modeling concepts. As most academic programs are focused on training well-rounded people who can orient them in new conceptual environments, we could assume that the point is not to steer people into specific, narrow interpretations (i.e., conceptualizations that strongly bias people into accepting one kind of thing as correct), but instead to focus on opening their minds to many different, equally valid, viewpoints from which they can analyze multiple situations (i.e., to steer them in a direction where their conceptualizations allow for many possible correct things). Concretely, we will treat the following research questions in this paper:

1. Do the conceptualizations students have of modeling concepts become more refined or nuanced as they progress through their studies?
   – If there is such a change, is it of a discrete or continuous nature?
   – If there is such a change, is it one-directional or reversible?
2. Is there a correlation between the educational stimuli students receive and the possible change in their conceptualizations of modeling concepts?
   – Do conceptualizations take the form of the semantics of a specific language or approach?

2 Method

Materials: The concepts we look at are ACTORS, EVENTS, GOALS, PROCESSES, RESOURCES, RESTRICTIONS and RESULTS. The different semantic dimensions we investigate are whether they are believed to naturally occur, are intentional or unintentional in nature, are a logical necessity or not, physically exist or not, and if they are vague or not. They will be respectively referred to as *natural, human, composed, intentional, necessary, material* and *vague*. Combinations of these features can then be used to characterize a given concept, for example a resource typically being a non-human, non-composed, material thing. These concepts and dimensions result from previous research we reported on in [2], in which we analyzed the specifications of modeling languages and methods used for different aspects of enterprise modeling.
Participants: We report on a longitudinal study amongst computing and information science students at Radboud University Nijmegen. We initially gathered students in the very first session at the beginning of their studies, at which 46 students enrolled to participate. Of these, 19 actually participated in the first phase. Over the course of the study, several students either stopped responding (without specific reason given), stopped because they changed their study program, or because they dropped out entirely. At the final measurement, 9 people participated, however, because one of them had not participated in an earlier phase we were forced to reduce the total set down to 8 complete measurements of the total timespan. All participated voluntarily and received no compensation for their participation.
Procedure: For the procedure, we adapted the basic technique of developing a semantic differential (taking into account the quality criteria set out by [5]) which we have detailed in earlier work [3]. We assume here that the selection of study participants, concepts and semantic dimensions to investigate and materials needed for them have already been done, as detailed in the materials above. For each semantic dimension we wanted to investigate we selected a set of 5 adjectives from an earlier pilot study, which ensures a significant reaction for that dimension [5]. We then constructed a differential with a page for each concept in which we included (1) a priming task to ensure participants responded in the context of conceptual modeling, (2) a differential in which each of the adjectives

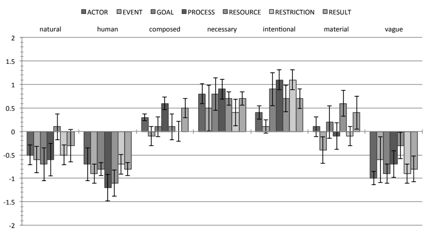
were presented to each participant in a random order, and they were asked to rate them on a 5 point Likert scale. We started the study at the beginning of the students studies so that we would have a null measurement. From then on, at the end of each semester, students received an email inviting them to a LimeSurvey implementation of the semantic differential, where they were also asked to detail what courses they had followed, and what new languages or techniques (if any) they were introduced to. In each phase we sent out 3 reminders to participants if they had not yet responded, and afterwards reduce the set of active participants down to those that participated.

Processing: The resulting data from the semantic differential was processes to calculate an average score for each concept-dimension combination based on the individual adjectives used to describe that dimensions. From this we constructed a vector for each concept, which contained scores ranging from 2.0 to -2.0, describing for each dimension how it relates to that concept. We considered scores $\geq$ 1.0 as positive judgments, and scores $\leq$ -1.0 as negative judgments. Other scores were considered as neutral. These judgments were then used to calculate a percentage wise breakdown of the amount of different polarities (i.e., negative or positive connotations) found for each concept.

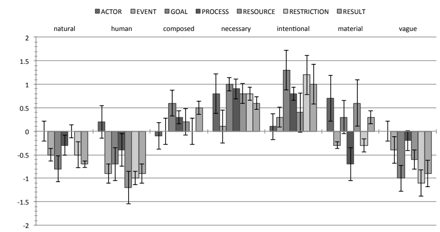
3 Results & Discussion

We present a visualization of the concept-dimension scores in Fig. 1. It shows the averaged results for each concept-dimension combination for each phase, with error bars showing the range of individual results. The polarities we calculated which show the relative amount of positive, neutral and negative responses to each concept-dimension combination are shown in Table 1, with some potentially interesting ones detailed in Table 2. Due to the amount of people that dropped out during the study, and the generally low initial response rate we cannot guarantee a strong external validity due to the lack of statistical generalizability. This could have potentially been prevented by including multiple, parallel groups of students (originating from different universities). However, this would lead to a strong heterogeneity of the results because different academic institutes and programs focus on different aspects. As such, whether those results could be first combined in order to create a larger coherent set of data is debatable as well. Nonetheless, the results here are still a thorough examination of specific individuals, and can be used to reason about the effects found in them, and to what degree measurement of their conceptual understanding is a feasible, and useful endeavor.

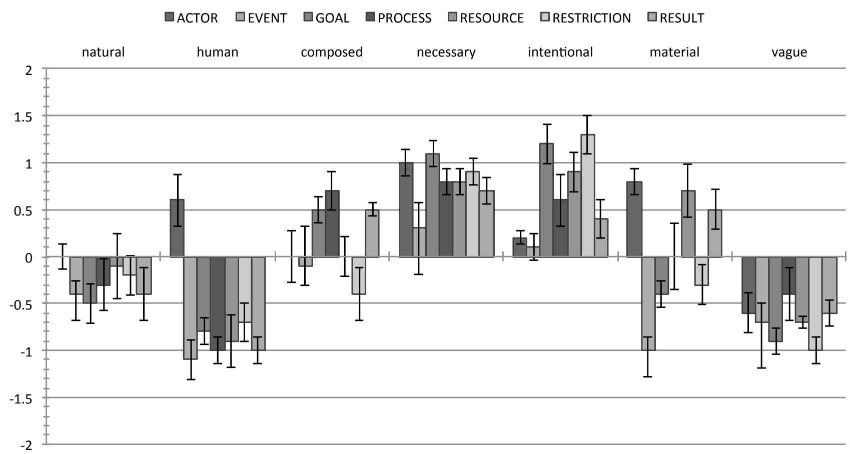
We can answer our primary research question by looking at both Fig. 1 and Table 1. In Fig. 1 we see that there is not a clearly obvious shift for any of the concepts or individual concept-dimension combinations to a particular understanding. For this to happen, the bars should either gradually or suddenly switch from ranging to one of the extremes to the other, or stay neutral in the middle. However, as we can see over time the general pattern of all the results stays similar to a sine wave, not having any of its constituents change too much. The semantic dimensions natural, human, and vague stay mostly negative for



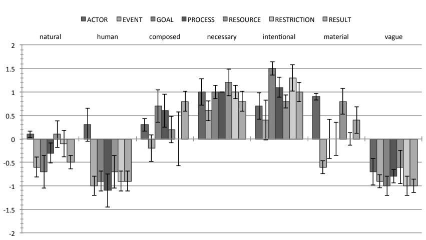
(a) Phase 1



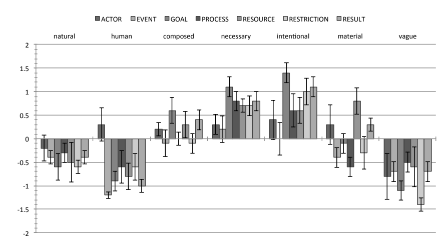
(b) Phase 2



(c) Phase 3



(d) Phase 4



(e) Phase 5

Fig. 1. Visualization of the average concept-dimension scores and individual variations for each phase of the longitudinal study.

most concepts, while the dimensions composed, necessary, intentional stay positive. The dimension material is the one dimension in which we clearly see both positive and negative polarities for different concepts, although these particular concept-dimension combinations still do not seem to change much over time. We can look in more detail at Table 1 to compare the actual distribution of the polarities for the concepts through time to verify this lack of systematic change. Here we also see that, while there are subtle variations from phase to phase in the relative amount of negative, neutral, and positive responses, there does not seem to be a significant gradual change increasing or decreasing over time for any of them. However, some specific concept-dimension combinations, do seem to have gradual shifts in one direction, which are documented in Table 2.

Table 1. Average polarities over all concept-dimension responses for each phase of the longitudinal study. Polarity scores of individual concept-dimension combinations are excluded due to space constraints, but are available upon request.

polarity	actor	event	goal	process	resource	restriction	result
				phase 1			
neg	26%	31%	31%	27%	21%	30%	20%
neu	53%	57%	46%	43%	47%	51%	54%
pos	21%	11%	23%	30%	31%	19%	26%
				phase 2			
neg	10%	30%	26%	20%	17%	26%	27%
neu	63%	57%	46%	63%	56%	51%	49%
pos	27%	13%	29%	17%	27%	23%	24%
				phase 3			
neg	7%	36%	31%	24%	24%	30%	26%
neu	57%	56%	41%	53%	46%	49%	49%
pos	36%	9%	27%	23%	30%	21%	26%
				phase 4			
neg	10%	37%	30%	30%	20%	29%	29%
neu	60%	49%	37%	37%	49%	44%	36%
pos	30%	14%	33%	33%	31%	27%	36%
				phase 5			
neg	17%	33%	31%	29%	24%	33%	24%
neu	59%	59%	39%	51%	44%	46%	47%
pos	24%	9%	30%	20%	31%	21%	29%

As a result, our first subquestion becomes irrelevant. However, the second subquestion is still interesting to look at, as the data do show that there are sometimes shifts for specific concept-dimension combinations where the polarity changes, and reverses again over the course of our study. While this might also be attributed to individual or contextual factors, it can hint at the flexibility of the students in their conceptual understandings while focusing on a specific way of thinking and working (e.g., because in one semester they work in a different paradigm than the others). Our second main question, and its related subquestion can be answered by also taking into account the educational stimuli students received (omitted due to space constraints). There do not seem to be specific systematic shifts that can be correlated with educational stimuli, nor do they seem to be systematically widening or refining to fit a specific way of thinking that could be attributed to them (e.g., the strong fact-oriented thinking approach of ORM). Given that students used several languages and techniques almost from the beginning of their studies until the final measurement, one could have expected to see some kind of development towards fitting those ways of thinking. However, given the lack of specific shifts into particular conceptual understandings discussed for question 1, this seems unlikely as well.

As touched upon earlier, there were some specific concept-dimension combinations that did show a development towards a specific conceptual understanding. Some possibly interesting ones are shown in Table 2. These patterns all show an example of a different polarity gaining or losing ground, which all translate into the willingness of a specific person accepting or rejecting a particular thing as being a good example of that modeling concepts. When we see that someone has a much stronger negative view on a particular thing (i.e., here the humanity of results), it becomes obvious that during modeling sessions those might come to the foreground when people clash on their conceptualization of the universe of discourse. Finding such specific strong polarized concept-dimension combinations might thus be a useful aid in steering such discussions to avoid communication breakdowns.

Table 2. Some interesting shifts of conceptual understanding in the results of the average (i.e., all participants) polarity scores.

polarity	p1	p2	p3	p4	p5	primary trend
		humanity of results				
neg	60%	60%	80%	70%	80%	
neu	40%	40%	20%	30%	20%	stronger negation
pos	0%	0%	0%	0%	0%	
		necessity of results				
neg	0%	10%	10%	0%	0%	
neu	50%	40%	30%	30%	30%	stronger acceptance
pos	50%	50%	60%	70%	70%	
		vagueness of actor				
neg	60%	20%	40%	50%	50%	
neu	40%	80%	60%	50%	50%	slight decrease in negation
pos	0%	0%	0%	0%	0%	
		naturality of actor				
neg	40%	0%	0%	10%	20%	
neu	50%	70%	80%	90%	70%	increase in neutrality
pos	10%	30%	20%	0%	10%	

Given that much training is done with a specific purpose (e.g., to make people acquainted with a specific subject, and steer them in a particular way of seeing things), it is disheartening to see both such a seemingly chaotic development of the conceptual understandings we measured, and a lack of correlation to the educational stimuli. However, given other recent studies into the way people learn modeling languages, this might not be entirely unexpected. In a study [4] on how well people understood different process modeling languages without formally being taught them, Recker and Dreiling showed that once someone had mastered or knew one particular language, the threshold to go to a different, similar one was very low (e.g., going from BPM to another process modeling language like EPCs, or similarly going from using i* to GRL for goal modeling). They concluded stating that it seemed not useful for an IT-oriented university curriculum to include teaching students multiple languages just for the sake of doing so, as they would likely be able to master them on their own when needed to. Given this understanding, one could perhaps infer that such continued educational stimuli (e.g., additional languages and techniques) should not necessarily be expected to have significant effect on the cognitive make-up of a student, which would include their basic conceptual understandings of the concepts used in those languages and techniques. Instead, such programs could perhaps focus more on exposing students to radically different languages and techniques, which have such different conceptual basis that they would learn a new way of looking at things. Thus, we agree with these studies that it seems less useful to train people to be modelers by teaching them every possible language for a particular focus, but that instead we should focus on opening their minds to different viewpoints that other stakeholders and modelers might have.

Acknowledgements. This work has been partially sponsored by the *Fonds National de la Recherche Luxembourg* (`www.fnr.lu`), via the PEARL programme.

References

1. van der Linden, D.J.T., Hoppenbrouwers, S.J.B.A., Lartseva, A., Proper, H.A.: Towards an investigation of the conceptual landscape of enterprise architecture. In: T. Halpin et al. (ed.) Enterprise, Business-Process and Information Systems Modeling, LNCS, vol. 81, pp. 526–535. Springer Berlin Heidelberg (2011)
2. van der Linden, D., Hoppenbrouwers, S., Lartseva, A., Molnar, W.: Beyond terminologies: Using psychometrics to validate shared ontologies. Applied Ontology 7, 471–487 (2012)
3. Recker, J.C., Dreiling, A.: Does it matter which process modelling language we teach or use? An experimental study on understanding process modelling languages without formal education. In: 18th Australasian Conference on Information Systems. pp. 356–366. Toowoomba, Australia (2007)
4. Verhagen, T., Meents, S.: A framework for developing semantic differentials in is research: Assessing the meaning of electronic marketplace quality (EMQ). Serie Research Memoranda 0016, VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics (2007)