How Do Teams of Novice Modelers Choose An Approach? An Iterated, Repeated Experiment In A First-Year Modeling Course*

Philippe J. Giabbanelli[1][0000-0001-6816-355X] and Piper J. Jackson[2][0000-0002-7025-5063]

[1] Department of Computer Science & Software Engineering, Miami University, Oxford OH 45056, USA,
https://www.dachb.com
giabbapj@miamioh.edu

[2] Department of Computing Science, Thompson Rivers University, Kamloops BC V2C 0C8, Canada
pjackson@tru.ca

Abstract. There are a variety of factors that can influence the decision of which modeling technique to select for a problem being investigated, such as a modeler's familiarity with a technique, or the characteristics of the problem. We present a study which controls for modeler familiarity by studying novice modelers choosing between the only modeling techniques they have been introduced to: in this case, cellular automata and agent-based models. Undergraduates in introductory modeling courses in 2018 and 2019 were asked to consider a set of modeling problems, first on their own, and then collaboratively with a partner. They completed a questionnaire in which they characterized their modeling method, rated the factors that influenced their decision, and characterized the problem according to contrasting adjectives. Applying a decision tree algorithm to the responses, we discovered that one question (*Is the problem complex or simple?*) explained 72.72% of their choices. When asked to resolve a conflicting choice with their partners, we observed the repeated themes of mobility and decision-making in their explanation of which problem characteristics influence their resolution. This study provides both qualitative and quantitative insights into factors driving modeling choice among novice modelers. These insights are valuable for instructors teaching computational modeling, by identifying key factors shaping how students resolve conflict with different preferences and negotiate a mutually agreeable choice in the decision process in a team project environment.

Keywords: Agent-Based Model · Cellular Automata · Education · First-Year Experience · Team-Based Modeling.

* We thank faculty at Furman University (Kevin Treu, Paula Gabbert, Bryan Catron, Chris Healy, Andrea Tartaro, Chris Alvin) for many fruitful discussions on teaching a topic-based introductory course.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

1 Introduction

For a given project, modelers can choose from a very large number of modeling techniques, both discrete (e.g., Agent-Based Modeling, Cellular Automata) or continuous (e.g., System Dynamics). *Ideally*, this choice would be driven by the availability of data, the project scope negotiated with the end-user, or the model performance given available computing resources [3]. However, all of these factors overlook an essential aspect of the model-building process: models are created *by modelers*. As most available literature provides little justification on the choice of a particular modeling approach, our previous study performed a survey of practitioners and found that the selection of methods primarily depended on factors related to the modelers, as 92% of respondents admitted that they chose the most familiar method [34]. However, 87% of respondents also declared that they chose a method based on the problem characteristics. This paradox suggests that a modeler may look at a problem and then decide on one method, convinced that it is the best fit, while a modeler with a different experience would make a case for another method in the same context. These two choices may have to be reconciled, since computational modeling is often constructed as a team-based discipline. In this paper, we examine the process through which novice modelers within the educational setting of a first year course negotiate the choice of a modeling approach.

A survey of 51 professionals in simulation education showed that 95.2% of programs on simulation education include a project, which involves a team in 92.1% of cases [20]. It is thus very common for students to be faced with the problem of reconciling multiple viewpoints when designing a model collaboratively, starting with the choice of a modeling approach. Although there has been research for several decades on how people behave when building a model, much of the focus has been on scenarios involving a single modeler [30, 37] rather than collaborative settings. When a team is studied, the focus may be on the dialogues between subject-matter experts (also called 'domain experts') and modelers rather than among the modelers themselves [8]. Similarly to Peter Rittgen, we conducted experiments involving groups of students who were provided with a textual description of four problems and asked to model them by choosing among techniques [23]. While Rittgen's experiment involved ARIS-EPC, Petri Nets, UML, DEMO, our approach focused on the choice between two closely related techniques: Cellular Automata (CA) or Agent-Based Models (ABM). Our experiments were repeated over two classes and used a scaffolding technique starting with independently choosing and justifying a model, then reviewing the choices made by a student with different arguments, and finally working in pairs to reach a consensus.

Our three main contributions are as follows:

(1) Through experiments, we identify a *core set of three problems that lead to high divergence among students*. These problems can be used by instructors to intentionally maximize divergent thinking and practice skills in collaborative problem solving for an introductory course on computational science.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

(2) By applying machine learning on experimental data, we show that the initial choice of agent-based models or cellular automata is primarily motivated by the perceived level of complexity of the problem.
(3) Through a thematic analysis of narratives by teams, we explain how students have a different perception of agent-environment interactions compared to interactions among agents. When a problem evokes interactions with the environment (e.g. through mobility), students choose Agent-Based Modeling; conversely, the absence of such interactions often justifies their use of Cellular Automata. In contrast, interactions among agents can lead to either option through the notion of networks or neighborhoods.

The remainder of this paper is organized as follows. In section 2, we introduce the setting in which we taught the course, including the teaching philosophy, core material, and institutional factors. In section 3, we explain how we designed experiments and which measurements were recorded at each step. Our results in section 4 are subdivided to focus on problems leading to different choices, explaining individual choices, and examining the process of co-creation as a pair. Finally, we contextualize our findings in the broader domain of collaborative problem solving in computational science and discuss the potential for applications to other areas such as processes in organizations.

2 The Setting: Teaching Philosophy and Implementation

The course was offered at Furman University, which is a liberal arts institution located in South Carolina, USA. The Computer Science Department used topic-based introductory courses to introduce key concepts of computer science to both majors and non-majors. The motivation is to "contextualize computing in a real-world, interdisciplinary problem upfront, and show how a variety of computer science topics apply to solving problems in that context" [32].

In the same manner as personal preferences and experiences shape a modeler's actions, simulation education depends on the teaching philosophy of the instructor [29]. Our teaching philosophy for this course, titled 'CSC 105: Virtual Worlds', rests on the four objectives presented at ICCS2016 [11]:

① We provide an *overview of the field followed by three modeling techniques*, covering both individual- and aggregate-level models. Our implementation of this objective was similar to courses at peer institutions, such as 'Modeling and Simulation for the Sciences' at Wofford College [27]. Specifically, students were exposed to system dynamics as aggregate models, then to cellular automata and agent-based models as individual-level models. In contrast to higher-level classes such as 'Introduction to Computational Modeling and Data Analysis', we do not include topics such as Markov chains or coupling models (i.e. hybrid modeling) [28].

② We cover *one programming environment*, starting from basic syntax. We complemented conceptual lectures by using the widely adopted `NetLogo` in weekly hands-on labs involving paired programming. We also emphasizing best practices from a software engineering standpoint [33].

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

③ We practice through *interdisciplinary projects*, based on years of experience in interdisciplinary curricular activities [12]. This objective is also made necessary given that the course includes the general student population, who may not major in computer science. Projects provide opportunities for creative expression, which we see as a cornerstone of a student-centered approach [9].
④ We develop *critical thinking skills*. In line with departmental practices [31], we used scholarly readings. The emphasis was on identifying limitations and on contrasting studies rather than in original writing, which is covered in a separate first-year seminar course [2].

As part of the learning objectives for the course, students should be able to choose between Cellular Automata (CA) and Agent Based Models (ABMs) in a given application context, program their model in `NetLogo`, and analyze simulation results. At a high-level, CA and ABMs are similar as they are both *discrete*, *individual-level* modeling techniques [1]. Consequently, students have to identify an initial configuration for the individual entities (e.g., initial state for the *cells* and/or baseline values for the *agents*), provide update rules that are applied to these entities at discrete 'ticks', and decide when the simulation should end. In a CA, the update rules can change the *state* of each cell based on neighboring cells, time, or probabilities. Examples include biological models such as the spread of an infection within a body [18] and the growth of tumor [19], or geographical models such as forest fires [6] and land use [36]. An ABM optionally includes a CA, which may serve to represent a physical substrate such as a soil model. An ABM necessarily includes agents, which can interact with the space (e.g., animals foraging) and/or with each other. Agents may be equipped with elaborate anthropomorphic notions, such as manipulating others, making errors, or having a range of personalities [15]. As an ABM requires more data (for calibration and validation) and a deeper theoretical understanding than a CA (to craft rules), applied computational models in fields such as obesity have gradually shifted from CA in early research to ABMs as they gained maturity [13]. Given the introductory nature of the course, we focused on homogeneous and synchronous CA [26] (e.g., all cells are updated at the same time) and we did not cover the connection between ABM and geographical information systems [16].

3 Experiments and Measurements

The course was offered on two occasions (Fall 2018 and Spring 2019), which allowed for repeated measurements. Our experiment involved three consecutive parts (Figure 1). First, students were given four problem statements and invited to create their own. In each of the five problems, a student independently chose to use either CA or ABM. Students had to argue for their choice and provide a complete design including a description of states, transitions, initial configuration, and condition to stop a simulation. In the second part, the instructor identified pairs of students who had made different modeling choices on at least some of the four shared problems. Each student received the anonymized submission from his or her pair-mate, such that they were unaware of each other's

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

identity and hence unable to communicate directly. Each student reviewed the submission and sent it to the instructor, who then passed it onto the other student. Finally, the two students met each other and had to decide on *one* technique for each of their five problems. For problems in which they argued for different solutions in part 1, students had to explicitly write how they resolved the difference.

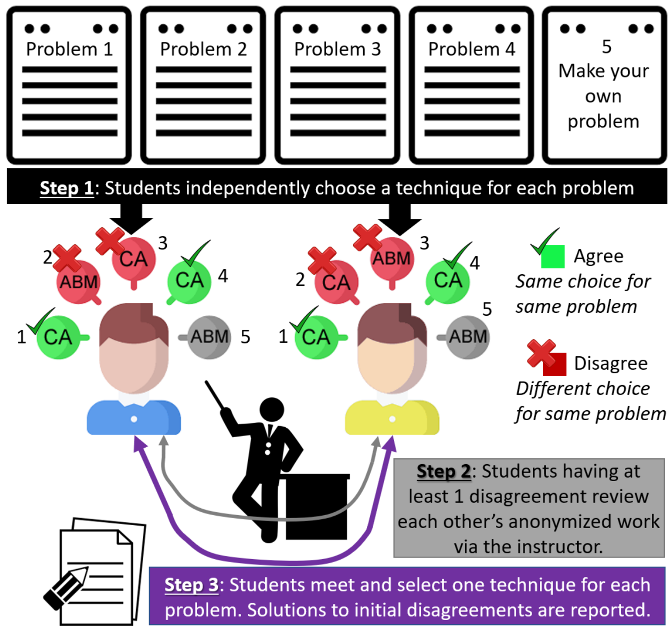

Fig. 1. Three steps process through which pairs of students resolve differences in their choice of ABM or CA.

In Fall 2018, the four problem statements started with "Spreading the flu in a classroom", described as follows:

> A handful of hard-working students may make the selfish choice of coming to class when they know they have the flu, thus infecting others and causing problems throughout the community. This model seeks to capture how flu spreads in the classroom. We consider that students are either infectious, or not (e.g., susceptible, infected but not yet infectious, recovered). A student may cough/sneeze/talk, which (to really simplify) produces a cloud of infectious droplets invisible to the human eye. If a person walks through this cloud, the droplets can land in the mouth/nose or be inhaled into the lungs, and this person can get sick. Droplets survive in the air for a few hours.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

The next problem statement was titled "Spreading HIV on campus":

> The Centers for Disease Control and Prevention (CDC) announced last August that sexually transmissible diseases (STDs) in US reached record high. The Human Immunodeficiency Virus (HIV) is an STD. We are interested in creating a model that simulates how residential Furman students may be getting HIV via unprotected sex over the course of an academic year. Once a person is infected, this person is extremely infectious (via unprotected sex) during the next few weeks. After treatment has occurred, the virus may drop to the point where it becomes undetectable in the body, and the consensus is that the risk of transmitting HIV then becomes statistically negligible.

In the third case, students investigated "Landslides":

> Because of flooding in North Carolina, the stability of some surfaces has been affected and they're now more prone to landslides. For a given land, we're interested in creating a model that can simulate how the land moves when the landslide takes place (i.e., assuming a landslide is triggered we want to know where the land goes). As vegetative structures may affect the dynamics of the landslide, note that the lands in the area of interest all have some shrubs and trees.

Finally, landslides were revisited in the last problem:

> Assume the setting and goals are identical to problem (3) above. In addition, assume that landowners with peculiar hobbies also have bongos, oryxes, kudus, and lechwe grazing on the land. As they graze, they damage shrubs (but not trees) and thus remove some of the vegetative structures that hold the land together.

Problem statements were re-written in Spring 2019 to provide additional data, thus limiting the possibility that results are an artifact of our initial problem statements. The new problem 1 was "Peer-influence on smoking":

> Smoking is partially driven by social norms, that is, whether peers smoke or endorse smoking. We want to model how students at Furman may choose non-smoking, vaping, or cigarette smoking. We are particularly interested in modelling peer influence on these choices within the academic community.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

The second problem focused on “Animal migrations”:

> Animals migration are based on the availability of food, climate, and migration of other animals. For example, birds or grazing animals may want to be close to similar animals, but too many will result in lack of sufficient nutrients. We are particularly interested in modelling animal migrations in the Carolinas, where migrating species include many types of ducks, swans, snow geese, various other birds for fall coastal migrations (e.g., warblers, grosbeaks, tanagers, orioles, vireos), fall mountain migrations (e.g., hawks, eagles, falcons) or springtime mountain migrations (e.g., thrushes, flycatchers). As a model is a simplification, you are not expected to become an expert in bird ecology to answer this question!

The third problem examined “Shopping malls”:

> Retail sales in shopping malls are important to model: where do people go if we make changes in the mall? How can we promote traffic? How can we charge rent for a specific shop based on how much traffic it can get? To answer such questions, we want to model the influence of foot traffic, customers, location, and neighboring shops, on retail sales in a shopping mall. You can pick your favorite mall (e.g., Haywood mall) if it helps you to think of something concrete.

Finally, the fourth problem was about “Laughter”:

> Some people have a contagious laughter that will make others laugh too. We want to simulate how laughter may spread in a room full of people. Laughter originates from one person, who may or may not have a ’contagious laugh’. Then, others may laugh or not.

For all students, we collected information on their gender (male, female) and whether they liked students with whom they worked during paired-programming sessions (which does *not* include their pair-mate on the experiment). In Spring 2019, we also administered two questionnaires: one upon completion of step 1 to understand how students made their selection (e.g., were they confident? did they feel it was appropriate for the problem?), and the other upon completion of step 3 to characterize how they resolved differences (e.g., was it easy to come to an agreement? were they impacted by the strength of their partner’s argument?).

4 Results

4.1 Which problems lead modelers to make different choices?

We were able to construct 8 pairs with at least one different modeling choice across the four problems (out of 9) in Fall 2018 and 4 such pairs (out of 7) in

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

	Fall 2018 (9 pairs)				Spring 2019 (7 pairs)			
	Q1	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**
Students using CA (%)	33.33	11.11	94.44	27.78	71.43	14.29	7.14	100
Students using ABM (%)	66.67	88.89	5.56	72.22	28.57	85.71	92.86	0.00
Pairs with differences	6/9	2/9	1/9	3/9	4/7	2/7	1/7	0/7
↪ *Number of differences*	4 had two diff., 4 had one				3 had two diff., 1 had one			
Pairs using CA (%)	44.4	0	100	11.1	85.7	0	0	100.
Pairs using ABM (%)	55.6	100	0	88.9	14.3	100	100	0
Differences solved using CA	3	0	1	0	3	0	0	0
Differences solved using ABM	3	2	0	3	1	2	1	0

Table 1. Modeling choices of individuals and pairs across questions and semesters.

Spring 2019. Our ability to create pairs that need to resolve differences depended on the extent to which each question led to a consensus. As shown in Table 1, the problem of "spreading the flu in a classroom" (Fall 2018 Q1) sharply divides students both as individuals and as pairs. The next questions leading to the most differences was "Peer-influence on smoking (Spring 2019 Q1) followed by our revisited landscape problem (Fall 2018 Q4). All other questions had less variations in modeling choices among individuals and no variation in pairs. This first result thus provides an experimentally established set of questions that can be used to promote differences in modeling choices.

4.2 Can we explain how modelers choose?

We analyzed the rationale for the individual modeling choices based on the questionnaire administered in Spring 2019. For each of the five problems (four created by the instructor and their own), students characterized their choice of a modeling method (confidence in the choice they made? Was it easy to decide?), rated the extent to which five factors influenced their choice (complexity of implementation, preference for the modeling technique, appropriateness for problem, explanatory power, flexibility of model), and characterized the problem by choosing between pairs of adjectives (social/physical, complex/simple, rules/ideas, individual/group, spatial/relational, decisions/behaviors).

The data for analysis thus consists of 14 questions, answered by each respondent (n=11) for five problems thus forming 55 entries. We used the supervised machine learning approach of binary classification to explain the modeling choice (CA or ABM) as a function of the 14 questions. The baseline prediction accuracy achieved by a 0R rule (i.e. simply looking at how often students tend to choose CA or ABM without considering the questions) was 54.54%, thus any classification model able to truly explain choices based on the questions would have to outperform this baseline. Using a decision tree algorithm without the restriction of depth, we found that *one question suffices to correctly explain 72.72% of choices*: seeing a problem as simple or complex. Complex problems overwhelmingly resulted in ABMs (21 out of 26 times), simple problems in CA (28 out of

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

32 times), and problems that could not be characterized by either term were still ABM (6 out of 8 times). If the notion of simple versus complex was removed, then no combination of the remaining 13 questions had explanatory power as the accuracy fell to 52.72%, which is below the baseline.

We examined whether simple dynamics could predict whether the final modeling choice in a pair would follow the initial decision of one student or the other. Since gender is occasionally used as mediating variables in studies on teamwork, we analyze the final choice in a mixed pair would espouse the initial decisions of the male student. Due to a small number of mixed teams, we only noted this situation in 3 out of 5 pairs, which is close to parity. We also tested the theory that a team may be led by a more 'knowledgeable' student. However, in pairs with different final exam grades, differences were resolved in favor of the student with *the lowest grade* GPA in 7 out of 10 pairs. The problem of identifying the nature and impact of leadership in a team would require further studies and a larger sample size, thus allowing a more fine-grained analysis of whether demographics or personality factors play a role in driving the final decision of a modeling team.

4.3 How do modelers co-create models?

In the survey administered upon resolution of modeling choices in a team, all students stated that resolving differences was very easy. To further characterize *how* differences were resolved, we examined the narratives provided by each pair on each resolution. In this section, we discuss the themes identified across narratives and briefly exemplify them through selected quotes in which students' *names were anonymized*.

Two themes were present across most narratives: mobility and decision-making processes. The need for an ABM was overwhelmingly motivated by the *perceived* need for entities to move over a space and/or engage in complex decision-making activities that require a cognitive framework. In contrast, the choice of a CA was justified by the perceived *absence* of these needs. For example, consider the flu problem. One group conceptualized students as "confined to one space within the classroom" and this absence of mobility resulted in a CA. Another group similarly debated the matter of movements:

> *"Simba claimed that [the flu] moved based on proximity because [it] cannot decide where it wants to go. Mufasa claimed that the flu was in the students who could walk around and make choices [...]. In the end, Mufasa agreed that the model is being used to show where the flu is and it is true that the flu cannot move so we agreed to explain this model as being a CA."* (Problem 1, Fall 2018)

In contrast, several other groups endorsed the hypothesis for the same problem that "agents can walk" thus leading to an ABM:

> *"Bernard believed that using Cellular Automata was the best approach and that we could have the students be stationary in an area, as if they*

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

were in sitting at desks. Bianca believed that using an Agent-Based modeling system was best because she has been in classroom environment that are hands-on more recently rather than lectures-based meaning the students would come in contact with others more often."

(Problem 1, Fall 2018)

An ABM is defined by interactions between agents and the environment as well as among agents. While the agents-environment interactions (e.g. through spatial mobility) were a recurring motivation for ABM, interactions among agents were less commonly discussed. In addition, the presence of these interactions was equally likely to motivate the use of a CA or an ABM, since both models include interactions (either through a neighborhood or via networks). For instance, consider the model of peer pressure over smoking. In one team, "the agents are interacting with other types of agents and also with the states of the environment, which are defining qualities of an ABM." However, for another team, these interactions can be handled by a CA:

"We believe that this situation is less about agent interaction and more about spatial orientation. The state of the smokers, non-smokers and vapers is more influenced by their neighbors than by individual decision making. Peer pressure often comes from people that are close to you, usually your friends. We will assume that friendship is stable and you will not randomly leave to find new friends. This implies that if you have many friends that smoke or vape it is more likely that you will smoke or vape. This is well illustrated by a Cellular Automata because while you interact with your own friends, they also interact with other friends who interact with other friends and so on and so forth."

(Problem 1, Spring 2019)

Through the narratives, we also notice that the resulting model and its justification is far from a one-sided triumph of one student's ideas over another. Indeed, students describe how features of the model are obtained *collaboratively*:

"We took things that Esmeralda's ABM could represent better and simplified them to fit the CA model. We used Quasimodo's base CA model and added those simplified elements for a more complete simulation, adding components like coughing and entering/leaving the classroom."

(Problem 1, Fall 2018)

"When we had peer reviewed each other's responses, we decided to try to combine each other's ideas (Mulan had the idea of cells being stable or unstable and Yao had thought of trying to show land movement through differences in elevation). This allowed us to be able to use both ideas to attempt a better model." (Problem 3, Fall 2018)

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

5 Discussion

In order to improve the practice of computational modeling and simulation, it is essential to understand the process by which we construct models, from the initial stages of defining the problem, designing the model, implementing it in software and testing its behavior. Certainly, the hard-earned experience and well-developed expertise that guide the decisions of experienced modelers are an invaluable component of building successful models. However, novices provide an excellent opportunity to examine which other factors may be highly influential, especially from a fresh perspective that is not invested in a certain way of doing things. Our study employs a mixed methods approach to investigate many aspects of the initial decision process as experienced by two groups of new modelers. We found that, at least in the case of considering cellular automata versus agent-based models, the apparent complexity of the problem under consideration is the key determining factor in terms of which technique was chosen. We also identified common themes found in the justification of model choice in cases where there was disagreement about the ideal technique. For problems where mobility and/or decision-making were key aspects needing to be modeled, agent-based models were preferred over cellular automata.

The exercise described here provides a sample of how modeling choices can be studied in an instructional setting. One key component is the set of sample problems which are designed to provoke consideration and discussion within the frame of the modeling techniques being considered. Another important component is the kind of data gathered by the study. Demographic factors did not appear to be highly influential among our groups, but further experimentation could help to clarify this facet of collaborative modeling. This kind of exploration also prompts us as educators to consider what we need to be teaching alongside the technical aspects of model building. If problem complexity is a driving factor in decision making (as we have seen here), should we encourage the consideration of other aspects? Should we provide more guidance on how to consider complexity?

Thus the primary contributions of our work pertain to simulation education and the practice of computational science in the classroom. Our work can also be situated within the broader theme of education and training in Collaborative Problem Solving (CPS), which is often motivated by the fact that most professional work is accomplished by teams within organizations [14]. Oppl further highlights how human work in organizations presents several salient characteristics that are also found in our study context [21]. Organizational actors can reach a shared understanding by working on *shared conceptual models*. As stated by Oppl, "existing approaches in general assume that the contributing actors have existing modeling skills [but] actors operatively involved in a work process do not necessarily have these modeling skills" and the task of model creation cannot be left to a third-party expert since the active and direct involvement of actors in the modeling process is "beneficial for the collaborative construction of a shared understanding" [21]. Our work thus also contributes to a growing

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

evidence base on the process of collaborative modeling and the negotiations that are involved [24].

A benefit of performing our experiments in one course at one institution is that we have a relatively consistent student population for analysis. However, this lack of diversity is a limitation when it comes to assessing the factors involved in shaping the co-creation of models by students of various levels, in different institutions, or involved in other curricula on computational modeling. In particular, the *experience* of students may be a mediated factor, since "conceptual modeling is often thought of as a skill that improves with experience" [35]. For example, studies on novice modelers by Powell and Willemain found several issues such as taking shortcuts [38, 22]. More recent empirical studies on the creation of models by students have confirmed that experience leads to different patterns [17]. It would thus be of particular interest to complement our study of freshman (1st year, 1st semester) with follow-up examinations in courses focused on rising sophomore (2nd year, 2nd semester) or senior (4th year).

A second limitation of our approach is the reliance on a prepared in-person session requiring a joint decision. In other words, students had to reflect in detail on each other's proposed idea then met in-person with the objective of finding one modeling technique. Results may thus be different if there is less incentive to achieve a joint decision, which may affect the willingness of participants to engage in co-creation and find a consensus. Results could also be affected by a switch to a remote scenario, which introduces an element of technology-mediated collaboration (e.g. via Zoom, WebEx, and similar platforms). An asynchronous scenario may rely even more on technology, for example through software for distributed model negotiation such as `COMA` in which modelers can propose a model, support or challenge a proposal (by tracking arguments for/against), and view the latest agreed upon version [25].

Although our study has collected detailed qualitative and quantitative data on modeling choices and reconciliation, software provide additional opportunities to track the series of steps taken by modelers, for example via a replay function [7]. The time series of modeling steps can be of particular interest if the models are *structured* rather than provided as narratives, for instance by using flow diagrams to document transitions of states for both cells in a cellular automaton [10] or agents in an agent-based model. A structured graphical notation may be better aligned with the task of model creation [4] and it supports new analytical tasks. If modelers co-construct a model as an annotated flow diagram, then automated metrics from network analysis become available both as a means of analyzing the evolution of collaboratively created artifacts and as a feedback tool for students [5]. Future research may include the development of tools that support asynchronous collaborations and mine structures as they are generated to either offer guidance to students or inform the instructor.

References

1. Badham, J.: A compendium of modelling techniques. http://i2s.anu.edu.au/sites/default/files/integration-insights/integration-

DOI: 10.1007/978-3-030-77980-1_50

insight_12.pdf (2010)
2. Barr, V., et al.: Computer science topics in first-and second-year seminar courses. In: Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education. pp. 643–644 (2017)
3. Brooks, R.J., Tobias, A.M.: Choosing the best model: Level of detail, complexity, and model performance. Mathematical and computer modelling **24**(4), 1–14 (1996)
4. Cachero, C., Meliá, S., Hermida, J.M.: Impact of model notations on the productivity of domain modelling: An empirical study. Information and Software Technology **108**, 78–87 (2019)
5. Chounta, I.A., et al.: When to say" enough is enough!" a study on the evolution of collaboratively created process models. Proceedings of the ACM on Human-Computer Interaction **1**(CSCW), 1–21 (2017)
6. Encinas, A.H., et al.: Simulation of forest fire fronts using cellular automata. Advances in Engineering Software **38**(6), 372–378 (2007)
7. Forster, S., Pinggera, J., Weber, B.: Toward an understanding of the collaborative process of process modeling. In: CAiSE Forum. pp. 98–105 (2013)
8. Frederiks, P.J., Van der Weide, T.P.: Information modeling: The process and the required competencies of its participants. Data & Knowledge Engineering **58**(1), 4–20 (2006)
9. Giabbanelli, P.J.: Ingredients for student-centered learning in undergraduate computing science courses. In: Proceedings of the Seventeenth Western Canadian Conference on Computing Education. pp. 7–11 (2012)
10. Giabbanelli, P.J., Baniukiewicz, M.: Visual analytics to identify temporal patterns and variability in simulations from cellular automata. ACM Transactions on Modeling and Computer Simulation (TOMACS) **29**(1), 1–26 (2019)
11. Giabbanelli, P.J., Mago, V.K.: Teaching computational modeling in the data science era. Procedia Computer Science **80**, 1968–1977 (2016)
12. Giabbanelli, P.J., Reid, A.A., Dabbaghian, V.: Interdisciplinary teaching and learning in computing science: three years of experience in the mocssy program. In: Proceedings of the Seventeenth Western Canadian Conference on Computing Education. pp. 47–51 (2012)
13. Giabbanelli, P.J., Tison, B., Keith, J.: The application of modeling and simulation to public health: Assessing the quality of agent-based models for obesity. Simulation Modelling Practice and Theory **108**, 102268 (2021)
14. Graesser, A.C., et al.: Advancing the science of collaborative problem solving. psychological science in the public interest **19**(2), 59–92 (2018)
15. Grantham, E.O., Giabbanelli, P.J.: Creating perceptual uncertainty in agent-based models with social interactions. In: 2020 Spring Simulation Conf. (SpringSim). IEEE/ACM (2020)
16. Heppenstall, A.J., Crooks, A.T., See, L.M., Batty, M.: Agent-based models of geographical systems. Springer Science & Business Media (2011)
17. Kavak, H., Padilla, J., Diallo, S., Barraco, A.: Modeling the modeler: an empirical study on how modelers learn to create simulations. In: 2020 Spring Simulation Conference (SpringSim). pp. 1–12. IEEE (2020)
18. Köster, T., Giabbanelli, P.J., Uhrmacher, A.M.: Performance and soundness of simulation: A case study based on a cellular automaton for in-body spread of hiv (2020)
19. Mallet, D.G., De Pillis, L.G.: A cellular automata model of tumor–immune system interactions. Journal of theoretical biology **239**(3), 334–350 (2006)

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50

20. de Mesquita, M.A., da Silva, B.C., Tomotani, J.V.: Simulation education: a survey of faculty and practitioners. In: 2019 Winter Simulation Conference (WSC). pp. 3344–3355. IEEE (2019)
21. Oppl, S.: Supporting the collaborative construction of a shared understanding about work with a guided conceptual modeling technique. Group Decision and Negotiation **26**(2), 247–283 (2017)
22. Powell, S.G., Willemain, T.R.: How novices formulate models. part i: qualitative insights and implications for teaching. Journal of the Operational Research Society **58**(8), 983–995 (2007)
23. Rittgen, P.: Negotiating models. In: International Conference on Advanced Information Systems Engineering. pp. 561–573. Springer (2007)
24. Rittgen, P.: Collaborative modeling-a design science approach. In: 2009 42nd hawaii international conference on system sciences. pp. 1–10. IEEE (2009)
25. Rittgen, P.: Self-organization of interorganizational process design. Electronic Markets **19**(4), 189 (2009)
26. Schönfisch, B., de Roos, A.: Synchronous and asynchronous updating in cellular automata. BioSystems **51**(3), 123–143 (1999)
27. Shiflet, A.B., Shiflet, G.W.: An introduction to agent-based modeling for undergraduates. Procedia Computer Science **29**, 1392–1402 (2014)
28. Silvia, D., O'Shea, B., Danielak, B.: A learner-centered approach to teaching computational modeling, data analysis, and programming. In: International Conference on Computational Science. pp. 374–388. Springer (2019)
29. Smith, J.S., Alexopoulos, C., Henderson, S.G., Schruben, L.: Teaching undergraduate simulation—4 questions for 4 experienced instructors. In: 2017 Winter Simulation Conference (WSC). pp. 4264–4275. IEEE (2017)
30. Srinivasan, A., Te'eni, D.: Modeling as constrained problem solving: an empirical study of the data modeling process. Management Science **41**(3), 419–434 (1995)
31. Tartaro, A.: Scholarly articles in the introductory computer science classroom. Journal of Computing Sciences in Colleges **34**(2), 188–198 (2018)
32. Tartaro, A., Cottingham, H.: A problem-based, survey introduction to computer science for majors and non-majors. Journal of Computing Sciences in Colleges **30**(2), 164–170 (2014)
33. Vendome, C., Rao, D.M., Giabbanelli, P.J.: How do modelers code artificial societies? investigating practices and quality of netlogo codes from large repositories. In: 2020 Spring Simulation Conference (SpringSim). pp. 1–12. IEEE (2020)
34. Voinov, A., et al.: Tools and methods in participatory modeling: Selecting the right tool for the job. Environ Modell Softw **109**, 232–255 (2018)
35. Wang, W., Brooks, R.J.: Empirical investigations of conceptual modeling and the modeling process. In: 2007 Winter Simulation Conference. pp. 762–770. IEEE (2007)
36. White, R., Engelen, G.: Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. Environment and planning A **25**(8), 1175–1199 (1993)
37. Willemain, T.R.: Insights on modeling from a dozen experts. Operations Research **42**(2), 213–222 (1994)
38. Willemain, T.R., Powell, S.G.: How novices formulate models. part ii: a quantitative description of behaviour. Journal of the Operational Research Society **58**(10), 1271–1283 (2007)

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77980-1_50