CrossMark

Stimulating Scientific Reasoning with Drawing-Based Modeling

Dewi Heijnes[1] · Wouter van Joolingen[1] · Frank Leenaars[1]

Abstract We investigate the way students' reasoning about evolution can be supported by drawing-based modeling. We modified the drawing-based modeling tool SimSketch to allow for modeling evolutionary processes. In three iterations of development and testing, students in lower secondary education worked on creating an evolutionary model. After each iteration, the user interface and instructions were adjusted based on students' remarks and the teacher's observations. Students' conversations were analyzed on reasoning complexity as a measurement of efficacy of the modeling tool and the instructions. These findings were also used to compose a set of recommendations for teachers and curriculum designers for using and constructing models in the classroom. Our findings suggest that to stimulate scientific reasoning in students working with a drawing-based modeling, tool instruction about the tool and the domain should be integrated. In creating models, a sufficient level of scaffolding is necessary. Without appropriate scaffolds, students are not able to create the model. With scaffolding that is too high, students may show reasoning that incorrectly assigns external causes to behavior in the model.

Due to the advance of science and the integration of science in daily life, scientific literacy is becoming more and more important to be able to make informed decisions in today's society (DeBoer 2000; Laugksch 2000). Part of scientific literacy

✉ Wouter van Joolingen
  w.r.vanjoolingen@uu.nl

[1] Freudenthal Institute, Utrecht University, P.O. box 85.170, 3508, AD Utrecht, The Netherlands

is the ability to reason scientifically, including the construction, application, and evaluation of scientific models (Nersessian 1999; van Borkulo et al., 2008; Windschitl et al. 2008). Models and modeling are considered fundamental elements of scientific literacy (Löhner et al. 2005; Louca and Zacharia 2012; Penner 2001). While students construct models, they learn to make concrete representations of abstract ideas and their underlying causes (Windschitl et al. 2008) which can motivate them to build and use scientific reasoning skills. This is acknowledged in the central role that models and modeling play in the next-generation science standards, published in 2013 (National Research Council 2013). These standards mention the importance of understanding the roles of models and modeling in scientific reasoning.

Computer models and simulations allow students to manipulate experimental variables with ease and produce results that are more predictable and easier to visualize (de Jong and van Joolingen 1998; Rutten et al. 2012). Constructing models in addition to using them is a way for students to experience research in the same or similar way as it is performed by many scientists (e.g., Barowy and Roberts 1999; Frigg and Hartmann 2012; Zhang et al. 2006). Louca and Zacharia (2012) showed in a review that model-based learning in science education affected cognitive, meta-cognitive, social, material, and epistemological skills positively, which in turn contributed to students' learning. Modeling, the practice of creating models instead of only using them, as a learning method is more effective at attaining conceptual and operational understanding of the nature of science and developing reasoning skills than current other learning tools (Harrison and Treagust 2000). Also, modeling has been shown to increase understanding of the structural aspects of a knowledge domain (van Borkulo et al. 2011).

Modeling naturally relies on the availability of modeling tools, offering students a representation and computational framework to build, execute, and evaluate models. Such tools

should be adapted to both the domain to be modeled, including associated modeling formalisms, and to the capabilities of the students who use the tool. In this study, we investigate the relation between domain characteristics, characteristics of the modeling tool and model-based reasoning in the domain of evolution, in lower secondary biology education. In the following section, we describe the characteristics of the tool and the domain.

Modeling Tools: Drawing-Based Modeling

In order to allow the construction of models, students must have tools that allow them to define objects and variables in the model as well as means to specify their behavior (Louca and Zacharia 2012; Tisue and Wilensky 2004; Weintrop et al. 2015). Such means often utilize equations or programming code. In the lower grades of secondary school education, however, creating computer models cannot involve the same amount of mathematics and abstractness as modeling in scientific research. Therefore, modeling for students in these grades should employ the more qualitative and visual aspects of models, as indicated by Nersessian and colleagues (Magnani et al. 2012; Nersessian 1995). Models using graphical representations hide the underlying mathematics and allow students to define the model in conceptual terms. In the current article, we focus on the use of *drawings* as a means for expressing the objects and relations in models, leading to the concept of *drawing-based modeling*. Drawing has been proposed as an engaging way to learn in science, because it makes students reason in various ways and allows them to compare their drawing with observations, measurements, and emerging ideas (Ainsworth et al. 2011). This can help students to consciously relate different concepts and behaviors to each other. Ainsworth et al. suggest that drawing should be a key element in science education. They state that drawings can be used to enhance engagement, to learn to represent scientific ideas, to serve as a learning strategy, and to support reasoning in science. Students choose a specific feature to focus on when drawing. The selection of these features is a representation of their reasoning about scientific concepts.

Wilkerson-Jerde et al. (2015) state that the combination of drawing, animation, and simulation can emphasize complementary elements of the domain and of scientific reasoning. In the domain of diffusion (how does smell spread?), they show that the use of these three media combined helps organize students' ideas about the domain. These aspects of drawing make it suitable as a basis for creating computational models, thereby integrating drawing and simulation. (Van Joolingen et al. 2010).

Based on this principle, Bollen and van Joolingen (2013) created SimSketch, a modeling tool based on drawings. In SimSketch, students can create their own drawings and assign

behaviors to the objects in the drawings. For instance, by drawing an object representing the sun and another representing the earth, SimSketch allows assigning the CIRCLE behavior to the earth, instructing it to move in circular orbits around the sun. The model can then be simulated and the (drawing of the) earth will move around the sun as specified. This means that drawing-based modeling goes further than pen-and-paper drawings: it enables students to transform their drawings into simulations, allowing them to convey more complex ideas in an understandable way and observe changes in outcome when changing variables. This example application of SimSketch was used in a study by van Joolingen and colleagues (van Joolingen et al. 2015). They found that students aged 8 and above were able to create solar system models and reason about the cause of solar and lunar eclipses, supported by a drawing-based model. Drawing-based modeling allows students to not just work in a set modeling environment but to design their own experiments and draw the objects they need. Based on these results, we expect that this idea may transfer to other domains, and that students in lower secondary education will be able to create models in the domain of evolution and be able to reason about the observed results.

Model-Based Reasoning about Evolutionary Processes

In this study, we focus on the application of drawing-based modeling to the domain of *evolution*. The theory of evolution is a foundational theory for most areas within biology. This means that understanding evolutionary processes is an important goal of biology education. However, in this domain, teaching may be hampered by wrong notions about the nature of science. Examples of such notion can be on the status and nature of scientific theories ("evolution is 'just a theory'") or students' lack of explanatory coherence (Smith 2009). In the case of evolution, the acceptance as a theory is also influenced by religious stances (Nadelson and Hardy 2015). Therefore, teaching evolution should go hand in hand with teaching about the nature of science and about the causal mechanisms in scientific theories (Lombrozo et al. 2008; Perkins and Grotzer 2005). The approach chosen in the current study involves the construction of evolutionary models by students. In constructing these models, applying them and reasoning about their models, we expect students to gain insight in the model-based nature of knowledge about evolution and to be able to build chains of reasoning to explain simple evolutionary processes. In doing so, we have adapted SimSketch to become a modeling tool that can support model-based reasoning in this specific domain. In the current article, we describe the design considerations involved in creating the modeling elements needed to enable students to model evolutionary processes

and report on the reasoning they display while creating and studying such models.

Research Question

In this study, we are interested in how engaging in a drawing-based modeling task influences students' scientific, model-based reasoning, within the domain of evolution. We want to see how drawing objects and building a model affects the amount and complexity of reasoning, and how details in the design of the modeling and task environment can stimulate or inhibit scientific reasoning. This can be formulated as the following research question: *What are factors in the design of a drawing-based modeling tool that stimulate or inhibit scientific, model-based reasoning within the domain of evolution?*

Method

In this design-based research study, we adapted the drawing-based modeling tool SimSketch to create a modeling environment in which students in lower secondary education can create explanatory models of evolutionary process. We developed and evaluated this environment and the task given to the students in four iterations. In each iteration, students' utterances related to scientific, model-based reasoning were recorded and analyzed and changes for the next iteration were made based on this analysis.

Participants

Participants were selected from lower secondary classes in a school for general secondary education. These classes were all taught by the same biology teacher. For each iteration, the teacher selected students at random from one or more of her classes. In the first iteration, four students participated from grade 7, followed by six grade 7 students, 12 grade 9 students and six grade 8 students for iterations 2, 3, and 4, respectively. Each student participated only once in this study. In each iteration, a new version of the SimSketch application was used.

The Modeling Task and Environment

The task for the learner is loosely based on the way natural selection works on the snail species *Cepaea nemoralis* (Cain and Sheppard 1952). These snails are polymorphic in shell color and banding, with certain morphs appearing more in certain areas. This is explained by thrushes, who hunt the snails by sight. The thrushes are more likely to spot and eat a snail which has colors that do not match the background area in which it is present. This results in most snails having a shell color that matches the background color of the area in which they live.

The task for the students consists of creating a model of snails in two adjacent areas differing in color: a green area to represent forest and a red area to represent clay. Students are then asked to explain the results of executing the model. Birds can be set to hunt by sight, which means they will pick those snails that have the highest contrast shell color to the background color. Snails reproduce, and on each reproduction, the color of the new snail differs slightly from the parent. The magnitude of the mutation as well as the reproduction rate can be set by the student. Over time, this results in a population of green snails in the green area and red snails in the red area (see Fig. 1). Ideally, students working on this assignment would observe the changes in color and attribute them to the right causes.

The modeling task evolved over the four iterations. In the first iteration, birds always hunted on sight—students could not modify this—and a fleeing behavior was present. In later iterations, students could select the hunting behavior, but fleeing was removed. Also, the accompanying instruction varied over the iteration. Details are given below where each iteration will be discussed separately.

Procedure

In each iteration, students worked on the assignment in SimSketch during one class period of 50 min. Of this 50 min, approximately 15 (with slight variations over iterations) was spent on introducing the context and explaining the workings of SimSketch, 30 on modeling, and 5 on discussing the results of modeling.

Students worked in pairs but on separate computers (with the exception of the third iteration). They first made their own drawing and then explained their drawings to each other to come to a joint drawing. This follows from the research of Van Joolingen et al. (2012) that showed that students who combine their drawings produce higher-quality drawings than students who work individually. During the third iteration, students worked in pairs on one laptop due to the lack of available laptops. This did not seem to influence the amount of discussion about the model. An additional benefit of working in pairs is that students naturally share their thought process, allowing us to follow their reasoning without having to force them to think aloud.

During the task, a teacher was present and answered students' questions on request. Also, the teacher asked questions to the students in order to help them in case they got stuck.

Data Collection

Students' conversations were recorded using voice recorders. Also, the log files of the interaction with SimSketch were
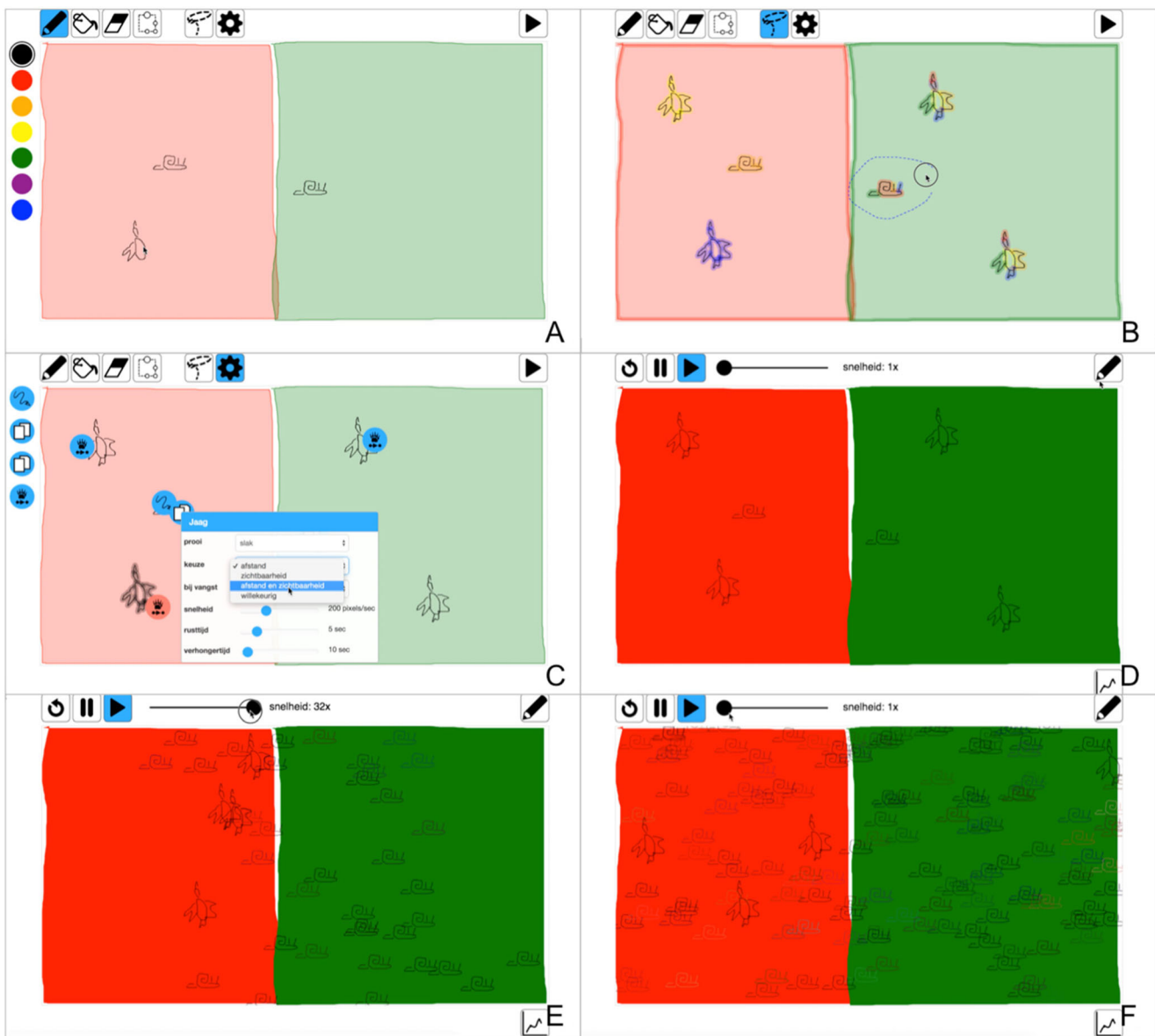
Fig. 1 Example of successful modeling of snail populations. **a** Drawing two areas (*red* for clay and *green* for forest), snails and birds. **b** Selecting those drawing strokes that should be considered by the program as one object (lines associated with a single object share the same color). **c** Assigning behaviors by dragging the icons from the left onto the objects. Adjustments can be made by clicking on the icon. **d** Starting the simulation. **e** Shortly after starting the simulation. The snails have multiplied and are spreading. **f** After the simulation has run for a few minutes, there is a population of red snails in the red area and a population of green snails in the green area

recorded in the final iteration. Unfortunately, technical difficulties prevented the recording of log files in the first three iterations.

Analysis

Students' conversations were transcribed and analyzed on two levels. First, utterances were classified in terms of their *reasoning complexity* using a method by Hogan et al. (1999). Second, utterances were qualitatively assessed on the causal reasoning about the model and the domain.

Reasoning complexity is used to gauge the quality of students' reasoning, which focuses on their ability to explain and elaborate on their understanding, rather than on comparing their knowledge to that of experts. This method therefore discerns different levels of reasoning without judgment on right or wrong answers. Though this does not show students' understanding of the subject of modeling, it does show how the process of modeling affects their ability to build coherent explanations of the processes studied.

The parts of the conversation where students discussed the assignment were selected and scored using the reasoning complexity rubric. This rubric gives a score on six complementing

components, each having five levels (zero to four) for increasing complexity of reasoning. These are as follows:

– Generativity: the number of subtopics within the discussion. Low scores are given for observations and generalizations; high scores are given for own ideas and assertions.
– Elaboration: the amount of detail given to the subtopics. Higher scores are given for more detail.
– Justification: the amount of evidence-based or inference-based support for an idea or assertion. Higher scores are given for more justifications per idea.
– Explanation: the mechanisms that are put forward to explain a phenomenon. Lower scores are awarded for single mechanisms; higher scores are given for multiple or chained mechanisms.
– Logical coherence: judged by the soundness of explanations or justifications of a phenomenon. Low scores are given for nonsensical or vague connections; high scores are given for solid and coherent connections.
– Synthesis: judged on the way different views are handled. Low scores a given to unresolved conflicting ideas, higher scores for supported rejection of one of the ideas, or combining ideas in to one, more complex idea.

Scoring scientific reasoning utterances over the subsequent versions of SimSketch and the modeling task allowed us to identify what elements can stimulate or inhibit students expressing higher-level reasoning. Utterances were scored by the first author and discussed among all authors. After scoring the final iteration, scores for the first three iterations were revisited and checked to ensure consistent scoring over all iterations.

Furthermore, student utterances were scanned for their relation with the domain in order to see if modeling in SimSketch can result in students' discovering of and reasoning with concepts of evolution.

Results

This section will discuss the findings in chronological order. For each iteration, the state of SimSketch is described, along with the results of the analysis of students' conversations and how these results are incorporated in a new version of SimSketch, instructions and assignment.

First Iteration

The design of the modeling environment started from the basic version of SimSketch 2.0, a remake of SimSketch 1.0 that was reported earlier (Bollen and van Joolingen 2013). The basic principles of SimSketch, drawing model elements and

assigning behavior to these elements, were maintained in this version, but the user interface was redesigned and the system now runs in a web browser instead of java. In the drawing process, a clear distinction was introduced between drawing, assigning behavior, and running by the separating different modes. In drawing mode, students can create and modify the drawing. In select mode, the students can combine various strokes into objects. In behavior mode, each behavior is represented by an icon that can be dragged onto an object. When selecting an assigned behavior, students can set parameters relevant to the behavior, such as the speed in the case the behavior implies a movement. In simulation mode, the simulation runs according to the model specified.

For the domain of evolution, three specific behaviors were created. The first is the "split-with-mutation" behavior resulting in an object to clone itself at a specified rate, in which the color of the new object differs slightly in a random way from the parent. The maximum size of the change can be set by the user. The second is the "hunting" behavior that makes objects hunt for objects of a specified category. In this first iteration, the user could not specify the way a hunting object picks its victims. Hunting objects hunted on sight always, picking targets with the largest contrast to their background.

In this iteration, sessions started with an introduction to SimSketch using the example of predator and prey—but without mutation. Also, a brief explanation of heredity and mutation was given. The assignment was explained orally and given in print. This first version of the assignment asked students to draw three areas: forest, sand, and swamp.

The students in this first iteration spend most of their time trying to figure out how SimSketch worked, which left them with less time to spend on modeling and reasoning. In total, the two dyads generated 22 utterances that could be classified as scientific reasoning. Table 1 shows the number of reasoning utterances for all dyads in the four iterations.

Further analysis on the students' reasoning shows that they do reason about the model, albeit not in the way that might be expected. The first segment shows hypothesis building and explanation of the hypothesis. (T is teacher, S is Student):

T: What is, you think, the influence of the colors of the areas on the snail shells—on the colors of the snails?
S: They become… darker.
T: How so?
S: Because the green color influences it or something, or it makes them have darker children.
T: Why would that make them have darker children?
S: Because they take up a bit of the pigments or something?
T: From the leaves?
S: Could be.
And, later in the conversation,

Table 1 Student's utterances that could be scored as reasoning steps, classified according to their reasoning complexity

Type	Generativity			Elaboration			Justification			Explanation			Coherence			Synthesis		
Level	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Iteration 1 (2 dyads)	7	1	2	4			4			2			1	1				
Iteration 2 (3 dyads)	13	3	6	1			7			2			3	2	2			
Iteration 3a (3 dyads)	15	2	4	3			5			1			1	2		2	1	
Iteration 3b (3 dyads)	20	1	13	8			9			5		5	1	3	1			
Iteration 4 (3 dyads)	19	2	3	2			1		1	2			2				1	1

Only utterance categories that scored more than 0 are shown. No utterances reached the highest level (4). Hence columns for this level are omitted

T: What do you think will happen with the snails in that area? Those that live on that bit?
S: Well, they get eaten a lot.
T: Why?
S: Because there are quite a lot, on a small area.
T: What is the difference with the snails that live over there?
S: Those get eaten even more because that is open field. It is a desert. Or sand.

Students needed the teacher to engage in higher-level reasoning. Interestingly, they did not see model and modeled system as separate. Although they had constructed their model themselves, they used explanations that had no source in the model. Areas were only modeled by their color, but they assigned the chance of being eaten to other properties of the field. A second example is the idea that snails pick up color from the environments, which is also not part of the model.

Because the hunting behavior was hidden for the students, it appeared to be hard for them to explain the observed behavior in terms of the model.

Second Iteration

In the second iteration, several changes were implemented. Some usability aspects were improved: the dialogs for specifying properties and behavior were moved close to the object selected, with as many preset choices as possible. For instance, when attributing the hunting behavior to an object, the prey can now be selected from a list of all available object names, instead of having to type it.

On the content level, the main change was that the birds' hunting behavior was made variable instead of being preset. Students can now choose between birds selecting prey by distance, by sight, by distance and sight, or randomly. By having to make this choice, students are stimulated to think about the implications of the different hunting behaviors. This change also allows teachers to reference the birds' hunting behavior during discussions.

To limit distractions and uncertainty for the students, the number of available behaviors was reduced: the flight behavior was not essential for building a working model and was therefore removed.

It was also apparent that the three areas left too much room for error: the colors of the areas were too close together, making it difficult to see different colored populations arising. To make the different population more visible for the students, from here on out, the assignment stated two areas of contrasting color: red to represent clay and green for forest.

During this iteration, six students from grade 7 worked with SimSketch. A brief explanation of heredity and mutation was given, and SimSketch was explained by drawing part of the snail and birds model and giving the objects different behaviors. The teacher showed the students the different hunting behaviors without explaining the implications they could have on the snail population. The written assignment had an extra sentence explaining the effect of mutation on the snails and stated the colors of the two areas to prevent confusion.

During this iteration, the average number of instances of scientific reasoning per couple was higher than the number in the first iteration: here, students showed an average of 13 instances of scientific reasoning (see Table 1). The level of reasoning was also higher: more students showed generative and coherent reasoning at higher levels.

Making the hunting behavior of the birds explicit helped during discussion:

T: How does it choose which snail it hunts?
S: On the one that is closest—random.
T: Why at random?
S: They can also go at the prettiest. But now—If they just pick the one that is closest. That is the most logical, really. That they hunt the one that, that they are like 'oh see there is the snail I'll take that one'. You know, they

are not going to fly ten kilometers to get him if they have already seen fifty.

In the conversation of another couple we see that de appearance of the two populations is noted, though the reason why the two populations come to be is not entirely clear to the students.

T: But do you also see a difference between this area and that area?

S1: They have like a bit more greener, green blue kind of. And here red, orange, yellow, pink, purple.

T: How do you think that came to be?

S1: Err… That is because—Those go to this area and the others go to that area, because they feel more at home or something?

T: Why do they feel more at home there?

S2: Because otherwise they are bullied.

Though none of the students came to the conclusion that two populations of different-colored snails emerge because of the hunting behavior of the birds, (scientific) reasoning about the snails' colors and the possible influences on them was taking place.

Third Iteration

For this version of SimSketch, some further changes were made to improve the ease of use. Most notable of these was removing the scaling option. Students were able to zoom out in simulation mode, but this interfered with modeling because this meant there was a white area around the two colored areas. This gave the snails room to spread there, meaning that birds hunting by sight would pick off only the snails in the white area because they had the highest contrast. Removing this function had no negative effects on modeling.

Also, some basic statistics tools were added to help students make sense of their observations. This included the number of objects with a certain name and the average contrast between objects and background.

The day the third iteration was tested, we had access to two ninth grade classes. To make the most use of this opportunity, we decided to try out two versions of the assignment to test the influence of the way SimSketch and the assignment was introduced on the way that students interacted with the program.

In both groups, six students worked in pairs. Suspecting that the school's computers prevented the logs from being recorded, we took three of our own laptops to the school. Because each couple therefore had access to only one computer, the assignment was adapted to have students work in pairs from the start.

In the first group, the emphasis was placed on the research question. Students were asked early on what they thought was going to happen with the colors of the snails while explaining the workings of SimSketch, and the model created was put on

second place. In the second group, emphasis was placed on explaining the SimSketch model. The research question was only written down in the assignment and not repeated during the initial instructions. This resulted in different levels of complexity of reasoning. The impact this had on the students' reasoning is shown in Table 1.

It is clearly visible that students in the second group spoke more about the modeling process in reasoning terms than did the first group (a total of 87 versus a total of 36). This places the average number of reasoning instances per couple in the first group around the same level as the first two iterations, while the average of the second group is more than twice as high. Almost half of these utterances fall in the level 1 generativity category, meaning they are simple observations. There are also, however, noticeably more higher-level reasoning patterns visible. For the first time observed during this study, students also showed higher-level reasoning while discussing among themselves, without direct stimulation from the teacher.

A comparison between the results of the two groups suggests that the explanation of the modeling tool is most important for students working with the program for the first time. Though students in the second group did spend some time wondering what the goal of the assignment was, they also spend a lot of time talking about what happened in their simulations. Even more, one of the couples came to the conclusion that birds eating more snails of a contrasting color results in two snail groups that have a color close to the background color.

One example of students showing higher-level reasoning without prompting by the teacher, though not directly related to the snails' colors, is shown below. The students observe in the simulation that all their snails (which they called Esmeralda) and birds (which they called George) are dead and explain why that would have happened. They also elaborate on what would have happened if the birds had died out first.

S1: What happened?All of a sudden they were dead.Oh whatever, we did well.

S2: We kept them alive as long as possible.Probably Esmeralda died first and then George didn't have any food left.

S1: Yes, because when George dies first then-

S2: That doesn't matter-

S1: That doesn't matter-

S2: Because Esmeralda stays alive.

One couple came to the right conclusion when prompted by the teacher and then went on to elaborate and explain their findings.

T: And, if it would hunt by 'this is the snail I can see the best'? What do you think would happen then?

S1: Then it will…

S2: Then it will—

S1: Then it will hunt by color.

S2: Yes, I think so too. Then, it will hunt… On red, it will hunt the green things and on green the red things. I think.

S1: Exactly.

T: So, what would you see happen then? What kind of snails would you get here and-

S2: Then, you would get a lot of green snails here and here a lot of red snails.

After the teacher left, the student explained further:

S2: I think that if the birds… Imagine there is a snail that has the same color red as this, that the bird can't hunt it because it can't see the snail. Hardly.

Also notable is that multiple students in the second group spoke of 'finishing the game' (Dutch: 'uitgespeeld'), while none of the students in the first group did. This could mean that students in the second group viewed SimSketch more as an environment to play in than as a school assignment, which could have influenced the way they interacted with the tool. This, in turn, could have contributed to the higher number of reasoning instances in the second group.

Fourth Iteration

Again, small changes were made to SimSketch. The "factory function," which seemed to raise confusion was replaced by the duplication function, which was placed in object selection mode. This way, the number of birds could be controlled in a more natural way.

The statistics tools were expanded to enable students to select variables to show a graph of. This allowed them to see the number of objects in a category, the contrast between a category of objects and their background, and the number of objects in each area.

Six eighth grade students worked in dyads in this iteration. This time, the teacher first introduced the assignment and refreshed or explained relevant knowledge, emphasizing that students should follow the written instructions that introduced SimSketch. The teacher then explained the different modes and behaviors in SimSketch by drawing part of the assignment. An overview of the version of SimSketch used during this iteration can be seen in Fig. 2 (also compare to Fig. 1).

Relatively few instances of scientific reasoning were recorded (see Table 1). In comparison with the seventh and ninth grade students in the previous iterations, these students seemed most hesitant to start working on the assignment. This could be connected to the way students accessed SimSketch: in previous iterations, SimSketch was available online; in this iteration, students accessed the program via a USB drive. This was done to safeguard log files. This has an

effect on the students' first impression of SimSketch. Instead of a well-known way to play games (in a browser), students had to click through multiple folders to access this unknown program.

Log files were saved locally, on the USB drive, circumventing the schools' network settings. While the analysis of students' conversations show the level of their reasoning, the log files give insight into the structure of their reasoning while working with SimSketch. The log files show the steps students take when creating their model: most students first draw the two areas, the snails and the birds, assign behaviors, and then go back and forth between playing the simulation and adjusting the behaviors. Some students initially forgot to select the lines constituting one object but amended this after playing the simulation for the first time. Notable was that some students used all behaviors on all objects. Apparently, the instructions did not make clear that not all behaviors should be used on all objects. Most students drew the snails in different colors. Though this should have no effect on the outcome of the simulation, it is notable that these students apparently misunderstood the difference between snails *becoming* different colors and snails *being* different colors. There also seemed to be some confusion between the "division" and "division with mutation" icon. Some students used both on the same object.

The higher-level reasoning that was recorded can mostly be attributed to students discussing what they should write down on the assignment sheet. Only one couple came close to the right conclusion, discussing at the start of the lesson,

S1: What do you think happens to the color of the snails' shells? That it—if it is hunted, then I think that they get a bit darker.

S2: Yes, like the shrubs. As camouflage.

Observing halfway through (though mixing up names of colors):

S2: If they are in the blue area they become blue.

But concluding:

S1: First it was blue, but then it became more green.

S2: And then, it became totally green and then-

S1: Now, it becomes dark black.

S2: Then, red.

S1: Where did you see red?

S2: Here.

S1: And then, see, yes red.

S2: And then, pink. It becomes lighter, this color.

S1: Conclusion?

S2: The colors are getting lighter.

This shows that a promising hypothesis does not always lead to the expected conclusions. The log files show that the

Fig. 2 The final version of SimSketch after four iterations. **a** Drawing mode, with selectable colors on the left. **b** Filling tool. **c** Eraser tool. **d** Modification tool, with copy tool on the left. **e** Object selection mode, with different colors indicating different objects. **f** Behavior mode, with icons of the four behaviors on the left. **g** Simulation mode, with replay, pause, and play buttons; speed adjustment slider; pencil button to go back to behavior mode; and graph button in the lower right. **h** Clicking the graph button brings up a tool in which can be selected of what category of objects a graph is shown of the number, contrast with background or location of this category
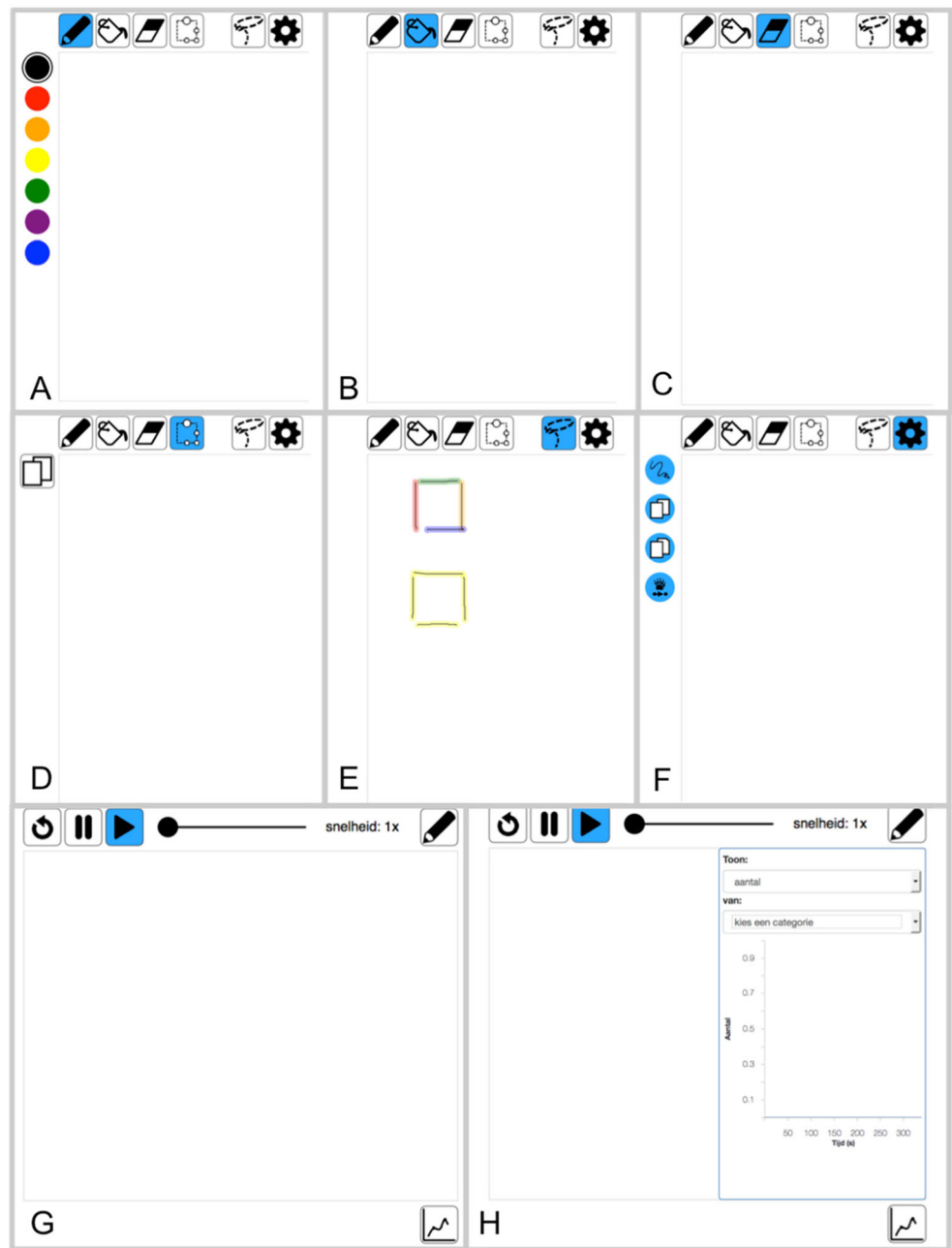
hunting behavior of the birds was not set on hunting by sight (see Fig. 3), resulting in a different change in the snails' colors. The teacher's help or personal feedback from the program would have been needed to get the students back on the right track.

This was the only iteration in which all students filled out the complete written assignment. The assignment did help to structure some students' thinking, allowing them to make a reviewable hypothesis early on. The amount of detail, both in the hypothesis and in the final conclusions, is poor. This could, however, be enhanced by stimulation by the teacher

or by spending some dedicated time discussing hypothesis building and conclusion writing.

Conclusions

Our research question addressed the factors in the design of the drawing-based modeling tool and associated tasks that influenced how students' scientific reasoning can be supported or inhibited within the domain of evolution. Whereas this exploratory study can provide only tentative answers to this
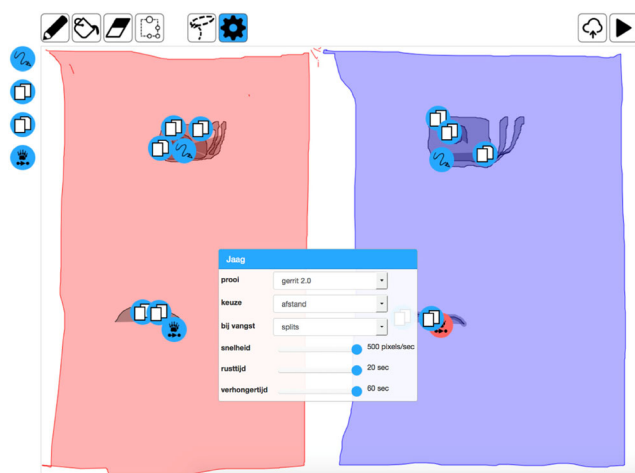
Fig. 3 Example of a students' model. The two snails have multiple division behaviors. The two birds hunt by distance

question, we observed a number of interesting interactions between the design of the environment and assignment and the resulting behavior of students.

The first obvious influence is the complexity of the model that has to be created. Simplifying the assignment and modeling tool yielded an increase in reasoning complexity, which could be seen in the move from the first to the second iteration. For the assignment, this meant going from three areas (forest, sand and swamp) to two with contrasting colors (forest and clay), which made the difference between the two populations easier to observe. Simplification was also applied to the modeling environment: those functions that were not strictly necessary for building a successful model were removed. For students who work with SimSketch multiple times, it could be beneficial to have some redundant functions, to make them reason about the things they really need and see how multiple functions influence the outcome of their model. The students in this study, however, all experienced working with a drawing-based modeling tool for the first time. To make the attainability of the assignment for these students higher, the decision was made to remove the "flee" functions and later integrate the "factory" function into the object selection tool. This made all behaviors in behavior mode essential for building a working model, thus removing distraction and uncertainties for the students. This allowed them to spend more time building models, with more certainty of going in the right direction. Similar issues were faced by VanLehn et al. (2016) who guided students in the direction of executable models by constraining options based on a known target model.

However, the opposite of simplifying can also be necessary. Students need to be able to manipulate key variables of the model to fully comprehend the processes in the simulation. In the first iteration, the hunting behavior was set to automatic and birds implicitly hunted by sight. This was done to reduce the difficulty of the assignment for the students. However, the

conversations between students and teacher showed that students did not have a clue of the cause of the changing color and they used causes which were external to the model such as "picking up color from the leaves." Having students set the hunting behavior themselves allowed for discussion of the influence of these behaviors in relation to the snails' color. This in turn allows for students to have a deeper understanding of the processes happening in their model, as brought forward in much of the modeling literature (e.g., Clement 2000, Louca and Zacharia 2012).

The comparison in the third iteration appears to show that mastering the modeling tool is the first concern before students start reasoning about the domain. The instruction that focuses on the tool seemed to engage students in a game-like mode, whereas the instruction focusing on the inquiry-based aspects resulted in a stance aimed at completing the given instruction. In the fourth iteration, the fact that the tool was started from a USB stick instead from the web also seemed to have negative impact on the students' playful behavior. This shows that the relation between learning environment design and students' playful behavior is a subtle one, as is also found in game research (Liu et al. 2014). Such playful attitude and not being scared to make mistakes may have a positive impact on students' inquiry behavior and hence on their level of scientific reasoning (Brady et al. 2015).

The teacher's support appeared necessary during the whole of each lesson. This is apparent in the hesitance many students showed and the amount of questions they asked the teacher, both about the working of the program and to give feedback on their progress. A written assignment can be used to structure reasoning processes, but feedback and support from the teacher using scaffolding practices is essential. Students often observe things happening in their model without giving reasons for why something would happen. Teachers can stimulate students' scientific reasoning by asking these scaffolding "why" questions.

Recommendations for Teachers and Curriculum Designers

Some students successfully set up models within 30 min of first encountering a drawing-based modeling tool. This suggests that with sufficient support and time, most—if not all—students will be able to successfully set up models using a drawing-based modeling tool.

For the first encounter with this new medium, it is most important to explain the basics of the tool. This can take away some of the insecurities of the students and enables them to play around with the tool without having to consult a teacher. This playing around, without directly worrying about right or wrong actions, is important for students to form their own understanding of the tool and the interactions between objects. Explaining the different functions (drawing, selecting objects,

assigning behaviors, running simulation) of the tool in the same order you would use them in the assignment can help students place the functions in a broader frame, instead of seeing them as separate units.

Scaffolding should keep students on track, without giving away answers. Students will often talk about their observations, but reasoning about the reasons why they happen can be lacking. Though assignments can have guiding questions, these can be misinterpreted (e.g., "Why do you get two different colors of snails?"—"Because the snails mutate"), while teachers' questions are more interactive and can go more in depth and support active inquiry (e.g., Rutten et al. 2015). Teachers can, for example, ask questions using the same terms and names the student uses and ask for more support behind an assumption or idea. Interaction with the teacher can also help to timely redirect students' reasoning to a promising path by using personalized questions and hints.

It is also recommended to pay attention to the limits of the modeling tool and the models that can be made with it. By discussing with students what parts of the model are realistic and which are not, what could be added to make it more realistic and which simplifications are necessary, students can gain a deeper understanding of models and how they are used. Such understanding of models is seen as central to science education (Nersessian 1995, Grünkorn and Upmeier zu Belzen 2014). By experiencing different subjects through models and discussing the link between model and reality, this could eventually lead to understanding how models are used in science.

Understanding of the subject of modeling is part of the reasoning process. During this study, it was unclear if students understood that the assignment was about evolution. For implementation in the classroom, the outcome of the students' models and their meaning could be used within the context of instruction. This could, for instance, include classroom discussion and comparing students' models with each other and with expert models of evolution.

Limitations and Further Research

Because of limited time and resources, a maximum of six students was used in each iteration. This did allow for a relatively quick turnover time between iterations but limits the scope of the conclusions. The evidence given in this paper can be seen as a proof of existence of certain ways of student reasoning. No conclusions can be drawn about the frequency of their occurrence.

For the first three iterations, SimSketch was run from a website, while in the fourth, it was run from a USB drive to allow local saving of log files. This could influence the way students viewed the modeling tool: during the first iterations, some students spoke of SimSketch as a game, while in the last

iteration, none did. This could have contributed to students' unwillingness to "play" with the program.

For further research with a larger sample, students' utterances could be coded and counted, in order to get insight in the frequency of occurrence of reasoning patterns. Also, extended use of SimSketch would give insight in how students' interaction with the modeling tool evolves over time. More attention could be paid to understanding SimSketch before working on the main assignment, using real-life contexts and to the limitations of the modeling tool in comparison with reality.

By changing the behaviors inside the tool, SimSketch can be used for a variety of subjects. Each subject could match with a different toolset, in which behaviors are placed that can be used to fulfill assignments within this subject area. By choosing, for example, an "infections" toolset, the behaviors "movement," "infection," and "resistance" would be added to the program. This would allow teachers to conveniently select those behaviors needed for an assignment and allow SimSketch to be used in a variety of educational settings. From the current study, it has become clear that for each new domain, studies are necessary to engineer the right combination of pre-set behavior, assignments, and teacher support.

References

Ainsworth S, Prain V, Tytler R (2011) Drawing to learn in science. Science 333(6046):1096–1097

Barowy W, Roberts N (1999) Modelling as inquiry activity in school science: what's the point? BT—modelling and simulation in science and mathematics education. In: Feurzeig W, Roberts N, Ruth M, Hannon B (eds) Modelling and simulation in science and mathematics education. Springer, New York

Bollen L, van Joolingen W (2013) SimSketch: multi-agent simulations based on learner-created sketches for early science education. IEEE Trans Learn Technol 6(3):208–216

Brady C, Holbert N, Soylu F, Novak M, Wilensky U (2015) Sandboxes for model-based inquiry. J Sci Educ Technol 24(2–3):265–286. doi: 10.1007/s10956-014-9506-8

Cain AJ, Sheppard PM (1952) The effects of natural selection on body colour in the land snail Cepaea nemoralis. Heredity 6(2):217–231. doi:10.1038/hdy.1952.22

Clement J (2000) Model based learning as a key research area for science education. Int J Sci Educ 22(9):1041–1053. doi:10.1080/095006900416901

de Jong T, van Joolingen W (1998) Scientific discovery learning with computer simulations of conceptual domains. Rev Educ Res 68:179–201

DeBoer GE (2000) Scientific Literacy: another look at its historical and contemporary meanings and its relationship to science education reform. J Res Sci Teach 37:582–601

Frigg R, Hartmann S (2012) Models in science. Stanf Encycl Philos:1–25. doi:10.1073/pnas.94.3.826

Grünkorn J, Upmeier zu Belzen A (2014) Assessing students' understandings of biological models and their use in science to evaluate a theoretical framework. International Journal of … 36(10):1651–1684 Retrieved from http://www.tandfonline.com/doi/abs/10.1080/09500693.2013.873155

Harrison AG, Treagust DF (2000) A typology of school science models. Int J Sci Educ 22:1011–1026

Hogan K, Nastasi BK, Pressley M (1999) Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. Cogn Instr 17(4):379–432

Laugksch R (2000) Scientific literacy: a conceptual overview. Sci Educ 84:71–94

Liu M, Rosenblum JA, Horton L, Kang J (2014) Designing science learning with game-based approaches. Comput Sch 31(1–2):84–102. doi:10.1080/07380569.2014.879776

Löhner S, van Joolingen W, Savelsbergh ER, van Hout-Wolters B (2005) Student's reasoning during modeling in an inquiry learning environment. Comput Hum Behav 21:441–461

Lombrozo T, Thanukos A, Weisberg M (2008) The importance of understanding the nature of science for accepting evolution. Evol: Educ Outreach 1(3):290–298. doi:10.1007/s12052-008-0061-8

Louca LT, Zacharia ZC (2012) Modeling-based learning in science education: cognitive, metacognitive, social, material and epistemological contributions. Educ Rev 64(4):471–492. doi:10.1080/00131911.2011.628748

Magnani L, Nersessian N, Thagard P (2012) Model-based reasoning in scientific discovery. Springer Science & Business Media, Berlin, p 2012

Nadelson LS, Hardy KK (2015) Trust in science and scientists and the acceptance of evolution. Evolution: Education and Outreach 8(1):9. doi:10.1186/s12052-015-0037-4

National Research Council (2013) Next generation science standards: for states, by states. T: National Academies Press, Washington DC

Nersessian NJ (1995) Should physicists preach what they practice? Constructive modeling in doing and learning physics. Sci & Educ 4:203–226

Nersessian NJ (1999) Model-based reasoning in conceptual change. In: Model-based reasoning in scientific discovery. Springer, US, pp 5-22

Penner DE (2001) Cognition, computers, and synthetic science: building knowledge and meaning through modelling. Rev Res Educ 25:1–37

Perkins DN, Grotzer TA (2005) Dimensions of causal understanding: the role of complex causal models in students' understanding of science. Stud Sci Educ 41(1):117–165. doi:10.1080/03057260508560216

Rutten N, van Joolingen W, van der Veen JT (2012) The learning effects of computer simulations in science education. Comput Educ 58(1):136–153

Rutten, N., van der Veen, J. T., & van Joolingen, W. R. (2015). Inquiry-based whole-class teaching with computer simulations in physics. Int J Sci Educ 37:1225–1245

Smith MU (2009) Current status of research in teaching and learning evolution: I. Philosophical/epistemological issues. Sci Educ 19(6–8):523–538

Tisue, S., & Wilensky, U. (2004). NetLogo: a simple environment for modeling complexity. In International Conference on Complex Systems. Boston

van Borkulo SP, van Joolingen W, Savelsbergh ER, de Jong T (2008) A framework for the assessment of learning by modeling BT—a framework for the assessment of learning by modeling. In: Blumschein P, Stroebel J, Hung W, Jonassen D (eds) A framework for the assessment of learning by modeling. Sense Publishers, Rotterdam, pp 179–195

van Borkulo SP, van Joolingen W, Savelsbergh ER, de Jong T (2011) What can be learned from computer modeling? Comparing expository and modeling approaches to teaching dynamic systems behavior. J Sci Educ Technol 21(267):267–275

Van Joolingen WR, Bollen L, Leenaars FAJ (2010) Using drawings in knowledge modeling and simulation for science teaching. In: Nkambou R et al (eds) Advances in Intelligent Tutoring Systems. Springer, Berlin, pp 249–264

Van Joolingen WR, Bollen L, Leenaars F, Gijlers H (2012) Drawing-based modeling for early science education. In: Cerri SA et al (eds) ITS 2012. Springer, Berlin, pp 689–690

van Joolingen W, Aukes AVA, Gijlers H, Bollen L (2015) Understanding elementary astronomy by making drawing-based models. J Sci Educ Technol 24:256–264

VanLehn K, Wetzel J, Grover S, van de Sande B (2016) Learning how to construct models of dynamic systems: an initial evaluation of the dragoon intelligent tutoring system. IEEE Trans Learn Technol 1382(c):1–1. doi:10.1109/TLT.2016.2514422

Weintrop D, Beheshti E, Horn M, Orton K, Jona K, Trouille L, Wilensky U (2015) Defining computational thinking for mathematics and science classrooms. J Sci Educ Technol 25(1):127–147

Wilkerson-Jerde MH, Gravel BE, Macrander CA (2015) Exploring shifts in middle school learners??? Modeling activity while generating drawings, animations, and computational simulations of molecular diffusion. J Sci Educ Technol 24(2–3). doi:10.1007/s10956-014-9497-5

Windschitl M, Thompson J, Braaten M (2008) Beyond the scientific method: model-based inquiry as a new paradigm of preference for school science investigations. Sci Educ 92(5):941–967 http://doi.wiley.com/10.1002/sce.20259

Zhang B, Liu X, Krajcik JS (2006) Expert models and modeling processes associated with a computer-modeling tool. Sci Educ 90(4):579–604 http://doi.wiley.com/10.1002/sce.20129