

Chapter 2

Model-Based Reasoning

Abstract Model-based reasoning consists of cycles of proposing, instantiating, checking, revising to find an apt model for a given purpose in a given situation, and reasoning about the situation through the model. Results from cognitive research can help us understand and assess both the experiential and reflective aspects of model-based reasoning. This chapter reviews research on model-based reasoning and the inquiry cycle to define aspects of model-based reasoning that can be used to guide assessment design.

Broadly speaking, inquiry is the process by which scientists and engineers formulate and investigate questions about the natural world in order to formulate answers, explanations, predictions, designs, or theories (NSES). Developing inquiry skills means being able to reason through fundamental concepts and relationships to understand and interact with particular real-world situations—in short, reasoning through models. Because scientific models embody hard-won, powerful, knowledge about how the world works, students do need to learn about important models in disciplinary areas of science. Many science-education researchers regard model-based reasoning as a pivotal way to unify content, the activities of inquiry, and teaching and learning (Stewart and Gobert, *op cit.*; Buckley, 2012); Clement 2000); Gilbert & Justi, 2016; Hestenes, 1987).

2.1 Scientific Models

A model is a simplified representation that focuses on certain aspects of a system (Ingham & Gilbert, 1991). Its entities, relationships, and processes constitute its fundamental structure. They provide a framework for reasoning across any number of unique real-world situations. The model abstracts salient aspects of the situations and goes further by viewing them as instances of recurring mechanisms, causal relationships, or connections across scales or time points that are not apparent on the surface. It formalizes experience usually by many people, tested, argued,

extended, and accumulated sometimes over centuries. Frigg and Hartmann (2006) provide an overview of models in science, and Harrison and Treagust (2000) give a typology of models as they are used in STEM education and in practice.

This brief concerns the explicit models that scientists create and use, and are targets of learning in science and engineering. The focus is not simply models as representations, but models as epistemic tools: Ways to understand the world, to interact with it, and to change it (Gilbert & Justi, 2016). A scientific model is a community resource—a particularly technical special case of what cognitive anthropologists call cultural models (Strauss & Quinn, 1998). The system of concepts, relationships, and processes that constitute a model extends beyond the mind of any individual. It is manifest in books and tools, in activities both formal and informal, in ways of seeing the world, and in patterns one can interact with the world and others. A web of interrelated ideas and activities spans individuals, is contributed to by many, is used by many more, and is enriched with every use (Latour, 1987). Science education aims to bring students into the community—to acquaint them with key concepts and relationships of important models, to be sure, but further to empower them to interact with the ideas and with people in the practically useful ways that are mediated through scientific models.

A broad conception of models highlights similarities in the kinds of thinking and activities that occur across a range of models. We want to ground *design patterns* on broad similarities in order to support task design across a broad range of content and contexts. More focused *design patterns* can be constructed for particular classes of models and representations. They would provide more focused support for particular areas of science or kinds of tasks.

For our purposes, models can be as simple as the change, combine, and compare schemas in elementary arithmetic (Riley, Greeno, & Heller, 1983), or as complex as quantum mechanics, with multiple forms of representation, advanced mathematical formulations, and interconnections with other physical models. Models can contain or overlap with other models. Relationships among the entities can be qualitative, hierarchical, dynamic, and spatial. Some models concern processes, such as the stages of cell division in meiosis. Some relationships can be qualitative (if Gear A rotates clockwise, Gear B must rotate counterclockwise), and some support quantitative or symbol-system representations and operations (if Gear A has 75 teeth and Gear B has 25, Gear B will rotate three times as fast as A). There can be different models for the same phenomena. The wave and particle models for light connect in some important aspects (amount of energy) but differ in ways that are useful for different problems (diffraction patterns versus the photoelectric effect).

Figure 2.1 suggests some central properties of model-based reasoning (a simplification of Greeno, 1983). The lower left plane (A) shows phenomena in a particular real-world situation. A mapping is established between this situation and, in the center plane (B), the patterns expressed in terms of the entities, relationships, and properties of the model. This is the “semantic” layer of the model. Reasoning is carried out in these terms. This process constitutes the reconception of the situation shown at the lower right (E). It synthesizes particulars of the situation with the abstracted structure of the model—a “blended space” for reasoning, as Fauconnier

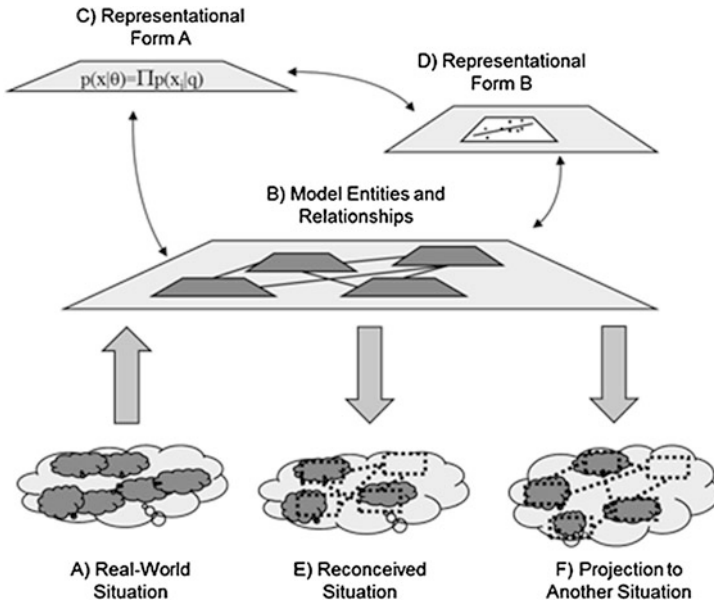


Fig. 2.1 Reconceiving a real-world situation through a model

and Turner (2002) call it. The processes and relationships of the model are used to make inferences such as explanations about the current real-world situation, and inferences about other situations (F) such as predictions or designs for artifacts (Swoyer, 1991). Above the layer of entities and relationships are symbol systems (C and D are two of possibly several) that further support reasoning in the semantic layer of this model, such as diagrams, matrix algebra, and computer programs.

Figure 2.1 also suggests properties that are important for understanding how models are used. The real-world situation is depicted as nebulous, whereas the entities and relationships in the model are crisp and well defined. Not all aspects of the real-world situation have corresponding representations in the model. On the other hand, the model conveys ideas and relationships that the real-world situation does not. The situation as reconceived through the model shows a less-than-perfect match to the model, but it provides a framework for reasoning that the situation itself does not.

The validity of a model does not address a two-way relationship between a model and reality, but a four-way relationship among a model, reality, a user, and a purpose (Suárez, 2004). As the statistician George Box said, “all models are wrong, but some are useful.” Being able to construct models that suit both the situation and the purpose at hand is central to model-based reasoning. Reasoning within models’ narrative spaces and manipulating their symbol systems are important, but they are not enough.

The strategies, the procedures, and the rules of thumb that enable one to put a model to practical use are the kinds of “epistemic games” (Collins & Ferguson, 1993) students must learn if they are to develop their capabilities for reasoning with

models. Students learn to reason in these ways *by* reasoning in these ways—in specific and real problems, in classrooms, in projects, in games, in hobbies. Ideally, support and feedback sharpens their reasoning and makes its generalizable structure explicit. Through these experiences, students begin to build increasingly broad and more generally applicable resources for both working with particular models and for the processes for reasoning with models (Schunn & Anderson, 1999).

Assessment of students' thinking and activities helps instructors guide their learning, and helps curriculum developers generate activities that fully reflect the targeted learning. The model-based reasoning *design patterns* bring out essential, recurring aspects of the processes of model-based reasoning, in ways that connect them to assessment arguments and help educators develop tasks to draw them out, whether focusing on particular aspects or on their interplay in investigations.

2.2 The Inquiry Cycle

In traditional science education, students are presented with models and asked to apply them to problems (Stewart & Hafner, 1991). But model-based reasoning in practice is characterized by the processes of proposing, instantiating, checking, and revising to find an apt model in a given situation. A model-based reconception is often provisional. Hypothesized missing elements can be used to evaluate the quality of the representation, and prompt a user to revise or to abandon a particular model. The hypothesized relationships then guide actions that change real-world situations and lead to further cycles of inquiry, understanding, and action. The depiction of the inquiry cycle in Fig. 2.2 (from White, Shimoda, & Frederiksen, 1999) is useful for highlighting aspects of model-based reasoning as they are used in inquiry and as they can be addressed in assessment.

Students can be presented, or propose on their own, a question that can be addressed by the concepts and principles in a scientific domain, then determine

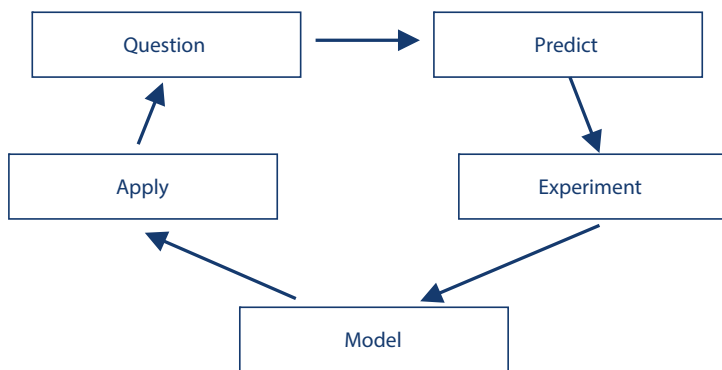


Fig. 2.2 The inquiry cycle

what observations might bear on its solution. They may be presented with, or gather themselves, data about the natural world, then build a model to account for patterns in the data. Once they have formulated a model, they may be asked to test the model by making predictions about further observations and determining whether it holds up in light of new information or requires modifications. If so, the cycle of model-building, model-checking, and model-revision continues, each stage requiring its own particular kind of reasoning.

Typically, students are introduced first to simpler forms of models and inquiry (e.g., provided substantial scaffolding to guide their investigations) and are then gradually exposed to more complex models (as described in the example in Box Genetics-1) and more independent situations for using them (Gotwals & Songer, 2010; Hammer, Elby, Sherr, & Redish, 2005; Redish, 2004; Songer, Kelsey, & Gotwals, 2009).

The multifaceted nature of model-based reasoning holds implications for both instruction and assessment. An instructor's decision to highlight to certain aspects will require assessment attuned to those aspects. The focus of instruction, and thus of assessment, for a new model may initially be reasoning through that model with data that are known to be appropriate. Alternatively, an instructor may want to see students work through cycles of inquiry with a model that is already familiar to the students. These latter tasks allow a focus on the self-monitoring and organizational capabilities required to coordinate the aspects of reasoning that interact when fitting models iteratively.

Students do not develop competence across all aspects of model-based reasoning at the same rate and depth. A student may be more facile with some aspects of inquiry in some content domains than others—and even for different investigations in the same domain (Mislevy, 2017; Ruiz-Primo & Shavelson, 1996). Instructors and assessment designers must consider the interplay between models and model-based reasoning, and where they want to focus attention. For example, an exercise meant to highlight model-checking could use a model familiar to students. An exercise to expand students' capabilities with a new model could employ a model-checking technique that students are familiar with from a previous lesson. The task designer must consider the extent to which declarative knowledge of a model's structure and components—as opposed to reasoning with and through the model—are to be stressed. Making this determination depends not simply on what is in the task but on the relation of that task to the experience of the examinee. This relationship may be known (e.g., as in local assessments embedded in instruction) and leveraged to sharpen the evidentiary focus of a task. Conversely, the relationships may be unknown (e.g., as in large-scale accountability tests), so that information about examinees' substantive knowledge about a model and their capability to use it are confounded. Sect. 2.3.4 says more about how these choices affect the evidentiary value of tasks in different assessment uses.

2.3 Some Relevant Results from Psychology

There are two basic modes of human cognition. Kahnemann (2011) called them “fast thinking” and “slow thinking;” Norman (1993) described them as experiential and reflective: “The experiential mode leads to a state in which we perceive and react to the events around us, efficiently and effortlessly. The reflective mode is that of comparison and contrast, of thought, of decision making. ... Both modes are essential to human performance (p. 15, 20).

Model-based reasoning involves both. As Giere (1987) put it,

My general view is that scientific theories should be regarded as continuous with the representations studied in the cognitive sciences. There are differences to be sure. Scientific theories are more often described using written words or mathematical symbols than are the mental models of the lay person. But fundamentally they are the same sort of thing (p. 143).

This section notes some results from research in cognitive psychology and learning science that are useful for understanding model-based reasoning, how people become proficient, and then how they might be assessed.

2.3.1 *Experiential Aspects of Model-Based Reasoning*

A person forming a mental model to understand a situation activates, assembles, and particularizes elements from long-term memory to create an instance of a model that is tailored to the task at hand. Walter Kintsch’s “construction-integration” (CI) model of text comprehension (Kintsch, 1998) provides insights into the process. Kintsch and Greeno (1985) apply the CI perspective to understanding reasoning with models. In one of their examples, the models of interest are Change, Combine, and Compare arithmetic schemas, and the problem is figuring out how a problem situation correspond to these models.

For a simple word problem, model formation takes place in working memory, incorporating features of the situation from sensory memory and information from long-term memory. Features of the situation activate elements of long-term memory, which can in turn activate other elements of memory or guide a search for new features in the situation. A person’s goals and affective state also influence what models are activated. This construction phase (the C in CI theory) is initiated by features of stimuli in the environment and activates associations from long-term memory-whether or not they are relevant to the current circumstances.

A “situation model” emerges from the integration (the I in CI theory) of mutually reinforcing elements among the immediate stimuli and the retrieved patterns. The situation model constitutes the person’s comprehension of the situation. Particular elements of the real-world situation are synthesized with more generalized patterns from that individual’s previous experience. Ideally, in the case of scientific models, the person activates appropriate chunks of formal models, and its elements correspond to elements in the real-world situation. Now the situation is comprehended in

terms of the salient elements and relationships in the scientific model (Larkin, 1983). This model formation sets the stage for further reasoning by activating, to the extent the person has developed them, associations of many sorts—narratives, representations, procedures, strategies, examples, and personal experiences.

The same cognitive processes also take place when students reason with partial, incomplete, fragmentary, and intuitive building blocks rather than with correct scientific models (diSessa, 1993, calls them phenomenological primitives, or “p-prims”). The resulting situation model again draws on patterns from the student’s past experience, which together provide an understanding of the situation upon which to base further reasoning and action. Unlike the situation model of an expert, however, this understanding may be based on superficial features of the situation or misconceptions; for example, the “continuous push” p-prim that an object will keep moving only if some force is continuously applied to it. Such understandings often suffice for everyday life. But they are not cast in terms of coherent conceptions that connect diverse situations and link them to effective procedures and strategies. People reasoning in this way are employing model-based reasoning, but not through the models that are the targets of science instruction.

Successfully forming a cognitive situation model around a scientific model requires not only the availability of the formal elements of the scientific model from long-term memory, but cues to activate them and to then relate them to the real-world situation (Redish, 2004). Experts have more information in long-term memory about models than do novices, but more importantly, they have more effective connections among them—including the conditions of when they are useful (Glaser, Chi, & Farr, 1988). Experts’ model formation is streamlined by extensive use, to accommodate more rapid access, larger chunks, and routinized.

For example, Chi, Feltovich, and Glaser (1981) asked novices and experts in physics to sort cards depicting mechanics problems into stacks of similar tasks. Novices grouped problems in terms of surface features such as pulleys and springs. Experts organized their groups around more fundamental principles such as equilibrium and Newton’s Third Law, each group containing a variety of spring, pulley, and inclined plane tasks. The experts’ categories reflect a well-practiced model formation process for understanding real-world situations in terms of principles that are not apparent on the surface. Their situation models are linked, in turn, to mathematical representations for solving problems (Model Use), for criteria to evaluating its suitability (Model Evaluation), and to strategies and procedures for carrying out these activities.

2.3.2 Reflective Aspects of Model-Based Reasoning

While scientific models can ground an individual’s understanding about a situation, they also are cultural tools that people can use to think and act together—a special case of what Wertsch (1998) calls mediated action. Seeing model-based reasoning as action underscores how science is not merely a matter of models, formulas, and

procedures, but ways of thinking, talking, and acting in the world, through patterns of knowledge and understanding that have built up within a community of practice.

Processes analogous to the CI model take place in conscious, explicit, model-based reasoning; that is, reasoning among people, using tools and external representations, occurring over minutes, hours, or years rather than milliseconds. Tools and external representations embody key relationships to enable computation and capture intermediate results to help overcome the limitations of working memory (Markman, 1999). The cognitive activation of relevant information in an individual's long-term memory is echoed externally in literature searches and conversations with colleagues. The external counterparts of refocusing a gaze are now generating scatterplots, looking for trends and outliers, and re-expressing residuals in a different format. The elements of a tailored, synthesized, and integrated model can be drawn from different domains, and reconfigured through multiple drafts of an article. The correspondence between the elements of real-world situations and the entities in an instantiated scientific model may require repeated attempts to determine just what to address, at what level of detail, and in what representational form to achieve the goals at hand. These are cycles of Model Formation, Model Evaluation, Model Elaboration, and Model Revision.

Managing one's own activities in their full complexity over time requires being able to reflectively monitor one's progress, evaluate the effectiveness of work, keep track of where one is, and determine next steps. These are metacognitive skills associated with model-based reasoning. White, Shimoda, and Frederiksen (1999) cited Piaget (1976)'s argument that reflecting on one's cognition reflects an advanced stage of development, and Vygotsky's (1978) claim that children progress from relying on others to help regulate their cognition to being able to regulate it themselves. Chapter 11 draws on this work for the *design pattern* for creating tasks to assess how students coordinate aspects of model-based reasoning within more encompassing activities. Self-regulation can be scaffolded as an option to design instruction to help students develop these skill, and to design assessments that either support them or put greater demands on them to assess them at higher levels.

2.3.3 *Higher-Level Skills*

Educators agree on the importance of higher-order skills such as critical thinking, problem-solving, systems thinking, and, to the present concern, model-based reasoning. There is less agreement on just what these terms mean. What is the nature of such skills, and how are they acquired? How they might be assessed? Research sheds light on the issue, and highlights design decisions that must be made in different ways to make the terms meaningful for particular purposes in particular contexts (the "use cases" described in Sect. 2.3.4).

These results follow from a view of learning as developing resources through experiences in specific contexts (Bransford & Schwartz, 1999; Hammer et al.,

2005). Building resources for, say, model-based reasoning starts in work with particular models, simple ones at first. The work is entwined with knowledge and skills connected with those models, and the particular problems and contexts in the situation at hand. Further experience begins to encompass more complex models, more complicated situations, and more sophisticated reasoning, always in the context of particular models and purposes. To the degree that the more general concepts and representations of working with systems are brought to the surface, learners begin to organize resources that can be adapted more readily to new models and more advanced practices (Schwartz et al., 2009). Students shift from seeing models as correct or incorrect to models as encompassing explanations for multiple aspects of a phenomenon. They develop more nuanced reasons to revise models. More advanced activities present challenges such as constructing a model to aid their own sense-making, and seeing model building as a way to generate new knowledge.

Still, engaging in what would be called “model-based reasoning” in any particular situation will jointly require resources for the substance, the context, and the practices that are involved. It is only through experience with multiple models in multiple contexts that students begin to develop more general capabilities they can bring to bear in new situations (National Research Council, 2000; Perkins & Salomon, 1989).

Constructs like “model formation” and “model revision” thus call out similarities as they appear to an outside observer, across what people do in situations that vary considerably in context and substance. Any assessment of model-based reasoning must therefore always face design decisions about the models, the content, and the context that are at issue. Critical questions for an assessment designer include what students know about the content and context, and what the designer knows about what the students know. *Assessment use cases* are helpful for thinking about these design issues.

2.3.4 Implications for Assessment Use Cases

The term “assessment” spans a broad array of ways and purposes for gathering information about what students know, can do, or might work on next. An assessment use case is a recurring configuration of people, information, contexts, and purposes. Model-based reasoning tasks have an inherent complexity because they necessarily involve some content, some context, and some practices. The interplay among these factors and the relationship to students’ backgrounds holds different implications for assessing model-based reasoning in the four use cases described below. Keeping the use case in mind while referring to the design patterns for support helps a designer make appropriate choices. It is not the features of a task alone that determine its evidentiary value, but the match of the task to the purpose and the students who will be assessed (Gorin & Mislevy, 2013).

Use Case 1: Formative assessment during learning activities

In this use case, inferences about students are used for feedback to further learning. It could be to a teacher, a learning system, or the students themselves. A significant factor of a task's value is how it matches up with what is known about students' previous experiences: A task may be quite complex, but for students working with this model at this time, some aspects will be known to be familiar and thus minor sources of challenge. Much of the knowledge that is necessary but irrelevant to the learning target is known to be familiar, and certain aspects of knowledge or modeling processes are targeted as the primary challenge. The evidentiary value of a task under these conditions can be quite high *for the targeted inferences*, because it is matched to local purposes about these students and takes advantage of local knowledge about their current and past experiences.

Use Case 2: Large-scale student-level accountability assessment

Consider a state accountability test where every student in Grade 6 is administered at a randomly-selected set of tasks, to estimate scores for individuals. The tasks are assigned without consideration of the matchups of the previous use case. Research on large-scale performance assessments shows that a student's performance on complex tasks assigned without knowing how the facets of the task match up with the students' previous experiences often does not convey very much information about how she would fare with a different, equally acceptable, task (Linn, 2000). The more diverse the test-takers, the greater the effect. There is low generalizability from how a student performs from one context to another or with one model to another, with respect to what is nominally "the same scientific process skill."

Use Case 3: Summative assessment in a course of instruction

This use case blends features of the two discussed above: assessments are integrated with a course of learning, but are used with higher stakes for individuals, such as a course grade or a certification. The College Board's Advanced Placement (AP) examinations are an example. Like both the accountability tests of Case 2 and the educational surveys of Case 4 below, AP examinations are large-scale assessments, developed and evaluated outside the local learning context. But because the College Board provides syllabi, sample tasks, evaluation rubrics, and instructional support materials for AP courses, many aspects of the critical student/task matchup are in place before the examination.

Use Case 4: Large-scale educational surveys

In large-scale educational surveys such as the National Assessment for Educational Progress (NAEP), samples of students are administered assessments to survey achievement across jurisdictions and to support research on its correlates. It is similar to Use Case 2 in that tasks are administered to students about whom relatively little is known. But it differs as to the intended claims: Not inferences about individuals, but about distributions of performance, relationships with demographic and educational background variables, and patterns of performance on some more complex tasks. In the last of these, rich work products such as log files of students' actions are obtained, providing evidence about the processes by which students perform tasks: their choices, the way they use tools, the steps they take, where they run into problems, and so on (for examples, see the 2014 NAEP Technology and Engineering Literacy (TEL) assessment: http://www.nationsreportcard.gov/tel_2014/).