International Journal of Science Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tsed20

Moving beyond the model as a copy problem: investigating the utility of teaching about structure-preserving transformations in the model-referent relationship

Daniel K. Capps & Jonathan T. Shemwell

To cite this article: Daniel K. Capps & Jonathan T. Shemwell (2020) Moving beyond the model as a copy problem: investigating the utility of teaching about structure-preserving transformations in the model-referent relationship, International Journal of Science Education, 42:12, 2008-2031, DOI: 10.1080/09500693.2020.1810354

To link to this article: https://doi.org/10.1080/09500693.2020.1810354

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION
2020, VOL. 42, NO. 12, 2008–2031
https://doi.org/10.1080/09500693.2020.1810354

Routledge
Taylor & Francis Group

Check for updates

Moving beyond the model as a copy problem: investigating the utility of teaching about structure-preserving transformations in the model-referent relationship

Daniel K. Capps [a] and Jonathan T. Shemwell [b]

[a]Department of Mathematics & Science Education, University of Georgia, Athens, GA, USA; [b]Department of Curriculum and Instruction, The University of Alabama, Tuscaloosa, AL, USA

ABSTRACT

An important research objective in modelling instruction is defining what students should learn about the model-referent relationship, as when unguided, they tend to errantly think of models as literal interpretations of their referents. Restating this problem we say students should learn about structure-preserving transformations between models and referents; including: what gets transformed, how it is transformed, and directions of information flow in transformation. To define what students should learn of this triumvirate, we introduce two explanatory statements: models are abstractions and models have transferability. We then investigated how readily students could learn about abstraction and transferability, and how this learning related to literal interpretation, by comparing pre- and posttest scores of modelling knowledge for students (n = 175) who participated in a modelling activity that provided experiences with abstraction and transferability. A control group (n=49) took identical tests but received no relevant instruction. Results showed that modelling students: (1) improved their capability to generate abstract models; (2) gained in their understanding that models represented their referents under transformation; and (3) consequently, did not think of models as literal interpretations. Lack of gains in the control ruled out a testing effect. We concluded that knowledge of structure-preserving transformations is a viable and valuable object of instruction.

ARTICLE HISTORY
Received 12 July 2019
Accepted 11 August 2020

KEYWORDS
Modelling; nature of models; metamodelling knowledge

Authentic modelling practice in the classroom depends crucially on whether students understand how models relate to the phenomena they represent. Indeed, when this understanding is absent, students are prone to think of models as 'copies,' or literal interpretations of phenomena (Grosslight et al., 1991; Tasquier et al., 2016). Unfortunately, and despite many useful ideas for combatting the literal interpretation view (e.g. Harrison & Treagust, 2000; Lin & Chen, 2002; Saari & Viiri, 2003), instructional definition of truer conceptions of the model-referent relationship – what students should learn that models are, if not literal interpretations – has been slow to develop in the educational literature. Such definition is needed more than ever, we argue, given expectations for

CONTACT Daniel K. Capps dacapps@uga.edu
Supplemental data for this article can be accessed at https://doi.org/10.1080/09500693.2020.1810354.

modelling instruction to be a mainstream feature of science classrooms (Australian Curriculum and Assessment Reporting Authority [ACARA], 2014; NGSS Lead States, 2013; NRC, 2012; SLO, 1996).

In this article, we propose that students should build an understanding of what we call the structure-preserving transformations that take place when a model is constructed from, or instantiated on, a referent. What students should understand of these transformations, we argue, is only loosely described in the existing treatment of models and modelling in the educational literature. In an effort toward greater precision in describing these transformations, we introduce a pair of concepts, summarised here as *models are abstractions of their referents*, and *models have transferability to their referents*, that when unpacked, should help students know what is transformed in the relationship between a model and referent, how it is transformed, and the directions of information flow in transformation. Further, we introduce an approach to modelling, called synthesis, that is geared towards helping students experience, and thus learn about, abstraction and transferability as embodiments of these three aspects of structure-preserving transformation. Through an empirical study in a high school Earth science context, we present evidence of (1) how readily students can learn about abstraction and transferability through the experience of synthesis, relating this learning to the what, how, and directionality of transformation; and (2) the extent to which learning about abstraction and transferability counteracts thinking of models as literal interpretations, thus providing a useful conception of the model-referent relationship that will support authentic modelling practice.

Theoretical framework

Problems with understanding the model-referent relationship, signified by literal interpretation, are a common phenomenon among novice modellers of any age (Dogan & Abd-El-Khalick, 2008; Grosslight et al., 1991; Harrison & Treagust, 1996; Ingham & Gilbert, 1991; Lin & Chen, 2002; Ryan & Aikenhead, 1992; Tasquier et al., 2016). Indeed, the strength of literal interpretation as a tendency may be why early researchers framed it as a persistent false belief, akin to a misconception (Grosslight et al., 1991; Ryan & Aikenhead, 1992). In the wake of this framing, intervention research has tended to favour combating literal interpretation over introducing truer conceptions of the model-referent relationship that should stand in its place. Inaugurating this trend, Grosslight et al. (1991) recommended teaching students that models were not literal interpretations, but they did not specify what to teach that models were. Continuing the trend, later researchers introduced various strategies for teaching against literal interpretation, but did not define or teach qualities of models that made them not literal interpretations (Harrison & Treagust, 2000; Lin & Chen, 2002; Saari & Viiri, 2003). Extending the trend to practitioner guidance, Krajcik and Merritt (2012) warned that students easily fall prey to literal interpretation, but did not specify non-literal thinking that instructors should promote in its place.

One good reason why true (i.e. non-literal) conceptions of the model-referent relationship have not been at the centre of researchers' efforts is that such conceptions have been slow to develop in the knowledge base. This situation is evident in Table 1, which summarises existing ideas about what students should learn that models are. These ideas are drawn from two distinct areas in the literature, definitions of models for educational audiences, and definitions of what students should learn about models and modelling, alternatively

Table 1. Presence of structure-preserving transformations between the model and referent in the modelling literature.

Source	Models are …	Transformation occurs	What gets transformed	How it is transformed	Directions of information flow
Definitions for educational audiences	deliberate simplifications (NRC, 2007, p. 152)	X			
	entities that represent some aspects of a phenomenon to some degree (Passmore et al. 2014, p. 1176)	X			
	abstract, simplified, representation of a system or phenomena that makes its central features explicit and visible (Schwarz et al., 2009, p. 663)	X	X		
Specifications for metamodelling and modelling competence	'idealized representations of the original' and 'theoretical reconstructions of the original' (Krell et al., 2014, p. 114)	X		X	
	… abstract entities that (1) represent salient features (not just what is unseen) (Fortus et al., 2016, p. 790); (2) explicitly … applicable to a range of phenomena (Fortus et al., 2016, p. 790; see also Schwarz et al., 2012)	X			X
Present study	abstractions of the structure of their referents that have transferability to their referents	X	X	X	X

conceived of as metamodelling knowledge (Schwarz et al., 2009) and knowledge for modelling competence (Gilbert & Justi, 2016). The table is structured to shed light on a needed area of the specification that existing ideas do not cover, namely the transformations that occur between a model and its referent. We propose that three main characteristics of these transformations need to be described in order to define the model-referent relationship: (1) what gets transformed (or preserved under transformation); (2) how it is transformed; and (3) the directions of information flow in the transformation. Combining these characteristics as a summary statement, there is opportunity to better define what students should learn of the 'structure-preserving' transformations in the model-referent relationship.

As the third column of Table 1 shows, all of the existing ideas about the model-referent relationship indicate that there is transformation (i.e. a model is not a literal interpretation). Less consistent, as indicated by the columns flowing rightward is information about what, how, and directionality. Accordingly, we provide our own working terminology for describing structure-preserving transformations at the bottom of the table. Summarised in prose, it is that models *are abstractions of the structure* of their referents, and consequently, models *have transferability* to their referents. These concepts are depicted in Figure 1. Unpacking them, an abstraction is an idea that has been pulled away from its

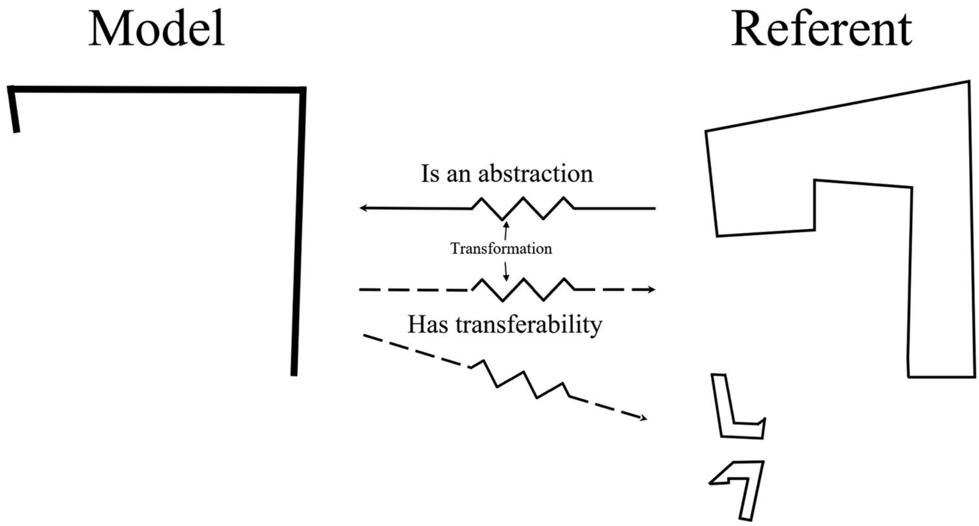


Figure 1. Abstraction and transferability, our working terminology for the structure-preserving transformations in the model-referent relationship.

source (Merriam-Webster, 2019). Models, meanwhile, are fundamentally analogues, meaning they are structures that mimic the structure of their referents (Nersessian, 2008). Combining these two ideas, mimicking structure of a referent corresponds to being an abstraction of structure, which is to say that models represent select information that has been pulled away from their referents, but with the essential structure preserved. Further, abstraction, distinct from extraction, denotes a transformation on the structure such that models differ in appearance from their referents (i.e. mimicking structure; not reflecting it). Thus, what gets transformed is the select structure of the referent, and how it is transformed is that it is an abstraction. Finally, abstraction shows one direction of information flow in the transformation. For the opposing direction, we used the term transferability to indicate how models can be instantiated on their referents. Transferring, as distinct from translating, includes the crucial idea that any instantiation of a model occurs under transformation, not directly. Also, transferability connotes the important idea that a model's structure may be instantiated on novel referents, meaning those not directly foreseen during model construction.

Returning to Table 1, the three columns at the right are intended to show precisely how abstraction and transferability carry more information about structure-preserving transformations between models and their referents than are available in any one of the existing formulations. As an example, Schwarz et al. (2009) description of an 'abstract, simplified, representation of a system or phenomena that makes its central features explicit and visible' is informative about what gets transformed (i.e. central features). However, as these authors do not unpack or emphasise the meaning of 'abstract,' they have a fairly lean description of how these central features are transformed (for us, these features are pulled away from the referent, with structure preserved). Opposite this, Krell and colleagues' description of 'theoretical reconstruction of the original' gives a strong indication of how models are transformed (i.e. rebuilt, with constraints provided by theory), but does not explicitly state what gets transformed (i.e. select structure within the original).

Meanwhile, only Fortus et al. (2016), following Schwarz et al. (2012) touch upon the two opposing directions of information flow, through their proposal that models are abstract entities that should be applicable to a range of phenomena.

One drawback of our terminology is that the word abstraction puts pressure on the usage of parts of speech, namely abstract (the adjective), abstraction (the noun), and abstracting (a verb). In the present article, we use the phrasing 'models are abstractions,' and avoid the phrasing 'models are abstract' because the adjective, abstract, often means fuzzy, or opaque, which is far from the concept of structure pulled away denoted by the noun form. Similarly, we are careful to use abstraction as a noun, as in 'models are abstractions' to denote the quality of being pulled away and avoid using it as a verb, or the act of pulling away. For the latter, we adhere strictly to 'abstracting.' This point is crucial because our intent here is to describe what models are, not how they are constructed. Indeed, modelling by overtly abstracting, which is featured in the present study and will be described in the next section, is rare in the modelling literature in science education. Instead, most authors describe abductive processes, wherein model structure is selected from a repertoire (e.g. Brewe & Sawtelle, 2018; Halloun, 1996, 1997; Hestenes, 1987, 1997) or suggested by classroom experiences (Clement & Núñez-Oviedo, 2003; Clement & Steinberg, 2002; Krajcik & Merritt, 2012). Many of these abductive processes, for example, those described by Halloun (1996, 1997) and Hestenes (1987, 1997), result in models that are highly abstract (i.e. mathematical expressions) despite the fact that overt abstracting processes are not involved. Thus, speaking carefully, although all models are abstractions, only some model construction uses abstracting processes, for example, those featured in the present study.

Purpose

Granting, for the moment, that the concepts of abstraction and transferability convey much of what students need to know about structure-preserving transformations by which models relate to their referents, it becomes important to ask how useful these twin concepts may be for supporting authentic modelling practice. Consequently, the empirical study described next had two purposes. The first was to find out whether school-age students (for us, secondary students) could make substantial progress in learning about abstraction and transferability in a short period of instruction (a 50-minute period of modelling). A positive result would indicate that these concepts of the model-referent relationship were learnable and thus viable as targets of instruction. The second purpose was to determine the extent to which increasing understanding of abstraction and transferability would run counter to literal interpretation. Such a result would show that learning these transformations between model and referent would be not only achievable but also valuable to achieve.

An abstracting process to model construction: synthesis

Asking whether it would be viable and valuable to learn about abstraction and transferability raises the practical question of how to support this learning. Our answer was to have students experience abstraction and transferability as part of model construction. Accordingly, the study presented next employed an inductive, overtly abstracting approach to model construction, shown in Figure 2, that differs markedly from typical, abductive approaches for generating abstract structure within conceptual models (e.g. Clement &

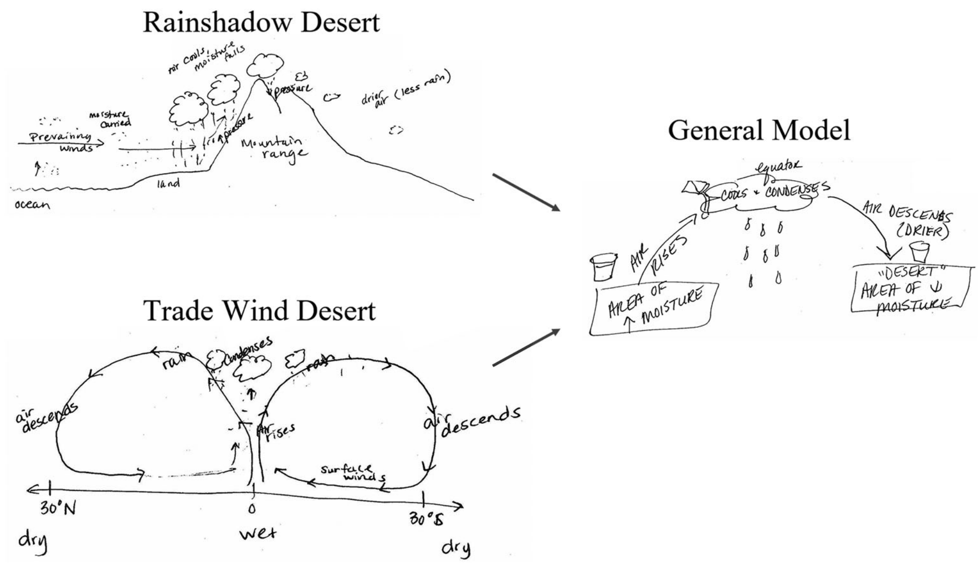


Figure 2. Students synthesized a model from two explanations of desert formation by abstracting their shared structure. Reprinted from Shemwell and Capps (2019). Reprinted with permission.

Núñez-Oviedo, 2003; Clement & Steinberg, 2002; Krajcik & Merritt, 2012). We call the approach 'synthesis modelling' because it supported learners in synthesising a unified model structure from two or more outwardly different source explanations that shared that structure under transformation. Synthesis follows what analogical learning theorists called schema induction (Catrambone & Holyoak, 1989; Gentner et al., 2003; Gick & Holyoak, 1983) or structure mapping (Gentner, 1983; Loewenstein, 2017). However, it extends these paradigms by having learners use modelling practices to represent the structure they abstract.

Figure 2 shows the basic process of synthesis, using an example that is similar to the models students constructed in the present study. At the left of the figure are two explanations for how different kinds of deserts form. Briefly, they are a rain shadow desert, in which air loses moisture as it rises over a mountain range before descending dry over the desert region, and a trade wind desert, in which, air loses moisture as it ascends in the wet equatorial region and moves poleward, eventually descending dry near the horse latitudes (30°N and 30°S). The model at the right of Figure 2 has been synthesised by abstracting and depicting the shared underlying structure of these two explanations, namely that air rises wet, moisture precipitates out, and the resulting dry air moves and descends in a new location. Complementing this process would be ensuring that the depicted structure (i.e. the model) had fidelity to each of the referents under transformation, or checking the model's transferability.

Method

To meet the study's two purposes, we engaged 175 secondary Earth science students in a two-week instructional unit on desert formation followed by a single period of synthesis modelling instruction. It was during this latter period that students were meant to experience – and hopefully learn about – abstraction and transferability. A pair of items on identical pre- and posttests measured student learning. Given the limitations of this pretest-instruction-posttest design (Shadish et al., 2002), the study included five augmenting design features, described next, to maximise the utility and trustworthiness of the evidence collected.

Design

Figure 3 shows the overall study design, five augmenting design features, and the threats or problems each of these features was meant to counter. Starting at the left of the figure, design feature 1 was the inclusion of a control group of 49 students who took the pre- and posttests surrounding a period of unrelated instruction (i.e. neither desert formation nor modelling). Scores from this group enabled us to show whether gains might have occurred merely from taking the same test twice (i.e. a testing effect). The second and third design features addressed the limitation that the tests had only two items of interest for the present study (i.e. one each for abstraction and transferability), which would otherwise severely undermine the validity of inferences about student learning. Accordingly, design feature 2 was to provide detailed information about each items' task demands and responses, so as to maximise the information upon which learning inferences were based. In this same vein, design feature 3 provided additional evidence in the form of think-aloud interviews with a sample of students as they responded to the two items of interest. Continuing on, the fourth design feature addressed the problem that students took the pretest at the start of the two-week desert modelling unit, and not immediately before the synthesis modelling period, thus raising the question of when learning about abstraction and transferability may have actually occurred. This question was important because we wanted to see if

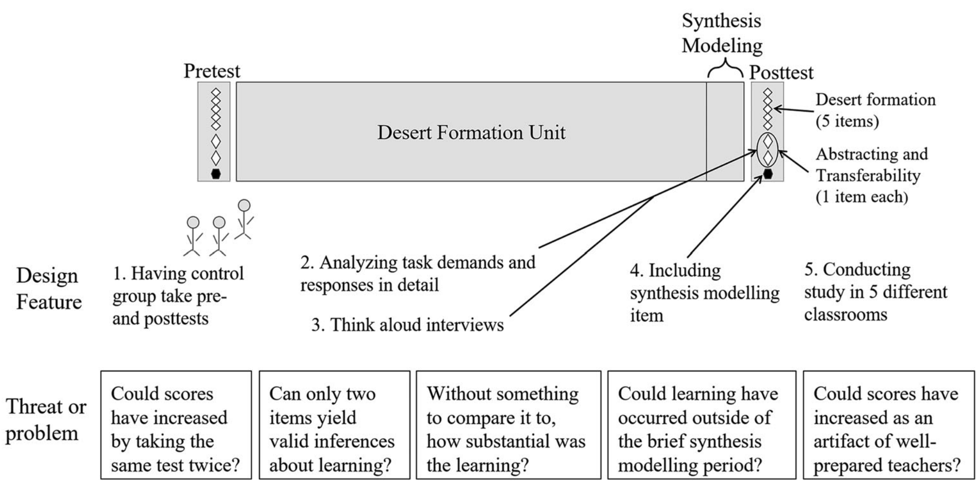


Figure 3. Study design including the five augmenting design features and the threat/problem they countered.

learning could occur during a brief, as opposed to protracted, intervention. Therefore, feature 4 was the inclusion of an item on the tests that measured how well students learned the synthesis model. A positive association between scores on this item (a proxy for success in synthesis modelling) and those for abstraction and transferability, would suggest that learning about these two concepts occurred during synthesis modelling and not before. The fifth design feature addressed the question of whether score increases from pre- to posttest could have been driven by an especially well-prepared teacher as opposed to synthesis modelling. Thus, feature 5 was to have modelling instruction span five teachers who had differing levels of preparation with synthesis modelling instruction. A comparative analysis would reveal whether the test results were unduly driven by well-prepared teachers.

Context, teachers, and student participants

Context. We conducted the study in six public, comprehensive high schools in a rural area of the northeastern United States. Each school served a small town and its outlying villages. The schools were located within about a 60-mile diameter of one another. The communities they served were predominantly white, non-Hispanic and non-affluent. Their median incomes ranged between \$33,000 and \$43,000 per year, significantly lower than the US median of \$53,657 at the time the study was conducted (United States Census Bureau, 2014).

Teachers. The modelling group had five teachers and the control group had three. Three of the modelling teachers helped develop the desert formation unit and the synthesis modelling activity and were thus especially well prepared to teach both of these. The remaining two modelling teachers learned about the instruction at a one-day professional development institute. The three control teachers were also recruited from among teachers who attended the one-day institute. All teachers received a stipend for administering the tests. The three teachers who helped develop the instructional materials received an additional stipend for this work.

All the teachers were white, non-Hispanic. They had between 2 and 22 years of teaching experience, with a median of 13 years. All were certified to teach high school science and regularly taught Earth science. The five modelling teachers taught in three different schools, with two schools each having a pair of teachers, and the third having a single teacher. The three control teachers taught in separate schools. Three of the modelling teachers were male and two were female. One control teacher was male and two were female.

Student participants. The modelling group comprised 175 students from the five modelling teachers' classrooms spread across two or three sections per teacher. The control group consisted of 49 students from one section of each of three control teachers' classrooms. The vast majority of the students in both groups were ninth graders. All were enrolled in mixed-ability Earth science classes. As Table 2 shows, the modelling and control schools were roughly comparable in terms of free and reduced lunch rates (State Department of Education, 2014) and performance on the mathematics and reading sections of the Standardised Achievement Test ((SAT) College Board, 2013), with a slight advantage for the control group. These data also show that neither student group comprised elite populations.

Table 2. Comparing student free and reduced lunch rates and Test Scores across school sites.

		% students at or above the 50th percentile[a]	
School	% students free and reduced lunch	SAT Mathematics	SAT Reading
Modelling 1	50	36.2	35.8
Modelling 2	48	47.2	45.4
Modelling 3	28	54.4	48.1
Mean modelling	42	45.9	43.1
Control 1	51	35.4	42.5
Control 2	28	49.5	45.2
Control 3	25	55.4	51.2
Mean control	35%	46.8	46.3

[a]The 50th percentile score was 460 for both mathematics and reading for the state where the study took place. Nationally, the 50th percentile was slightly higher, 490 for reading and 510 mathematics.

Instruction

In the unit leading up to the synthesis modelling period, students spent six to nine class periods learning about the two types of desert formation depicted at the left of Figure 2, rain shadow deserts and trade wind deserts (see the Supplementary Materials for details). This unit included no instruction on model abstraction or transferability. Upon completion of the unit, students engaged in the 50-minute synthesis modelling period. Their task during this period was to construct a 'general' model for desert formation (i.e. Figure 2, right) that could explain both rain shadow and trade wind deserts. The period included a pair of 10-minute model construction stages, each followed by a whole-class sharing and discussion which also took about 10 min. During the model construction stages, students worked in small groups on table-top white boards to generate their models as an abstracting process. During the sharing intervals, they presented their models to their classmates. Questions from the teacher were meant to help students work out whether and how the abstracted model could explain both types of desert formation, (i.e. rehearsing transferability). Importantly, at no point in modelling instruction did teachers or students use the terms abstraction and transferability. Instead, language centred on how the model was meant to be 'general' or 'powerful' enough to explain the formation of both types of desert in a single representation.

Measures and coding

Identical pre- and posttests were administered by modelling and control group teachers in their classrooms and took about 20 min to complete. As explained in the Design, there were three key items on the tests, the were the: abstracting, transferability, and synthesis model items. Each of these items is described in detail below. The tests also included five items to evaluate learning from the desert formation unit that was unrelated to synthesis modelling, abstraction, or transferability. Thes unrelated items, along with student results on them, are included in the Supplementary Materials.

Abstracting item. Figure 4 shows the abstracting item, so named because it required students to abstract structure from a referent to generate a model. The item showed an image of an igloo and referred to it as a specific model for a house. It then asked students to construct a general model[1] for a house using words and/or pictures. An example of abstracting preceded this prompt. It depicted a general model for a sports ball that had been abstracted

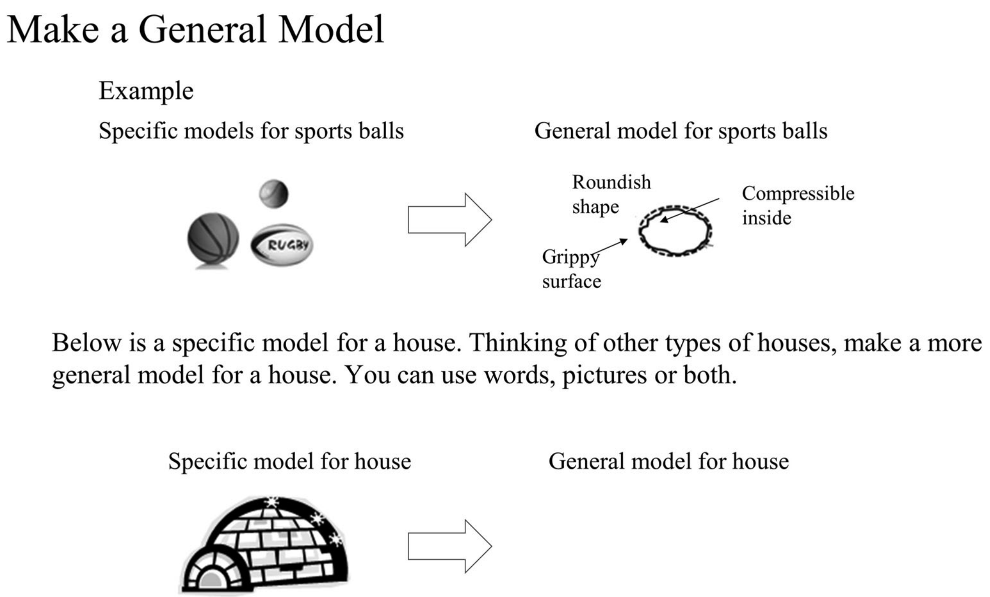


Figure 4. Abstracting item from the pre- and posttest with the example used to help students understand the task directive. Reprinted from Shemwell and Capps (2019). Reprinted with permission.

from several specific sports balls. We included the example to ensure the directive to abstract a model would be understandable to students on the pretest. Taking the prompt and example together, the item was meant to measure students' capability in abstracting, composed of recognising what was required to abstract, via the example, and then effectively meeting this requirement. We coded the item using the values 2, 1, or 0 according to whether all, some, or none of the structural elements in students' responses were abstracted. We labelled these levels full, partial, and no capability in abstracting (see Table 3). Since students could optionally respond in words or drawings, we coded these aspects separately and took the higher of the two scores. Interpreting the scores, we reasoned that good performance on the abstracting item would depend at least in part on knowledge of abstraction, especially knowledge of the goal or endpoint of the abstracting process. We used the think-aloud interviews as a check on whether this was so, examining students' verbal outputs for evidence of their thinking as they completed the item.

Transferability item. The transferability item measured whether students understood that models often represent scenarios that differ outwardly from each other (i.e. models have diverse referents). The item asked students to state the potential advantages of a general model. It read:

> Scientists use models to describe and explain things. One scientist made a specific model of a situation. Another scientist made a more general model of the same situation. Explain the potential advantages of the general model over the specific model.

As shown in Table 4, we coded the responses 2, 1, or 0 based on whether students stated that general models had diverse referents (level 2), stated that they had multiple referents without actually stating that they were diverse (level 1), or gave an irrelevant response

Table 3. Coding for the abstracting item.

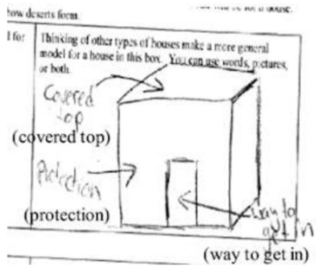
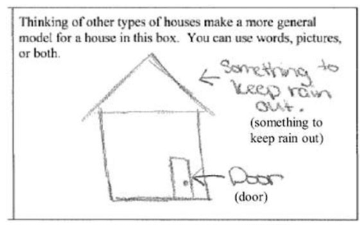
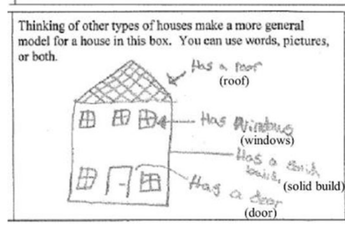
Code	Criterion	Example Response	Inferred Capability
2	All elements abstracted		Full
1	At least one element abstracted		Partial
0	No elements abstracted		None

Table 4. Coding for the transferability item.

Score	Criterion	Example response	Inferred Knowledge
2	Applicable to diverse referents	The general model of the situation is a more practical model to use because it could represent the same type of situation with a number of different circumstances	Knew that models had diverse referents
1	Applicable to multiple referents	A general model would be able to explain more situations than a specific one	Possibly knew that models had diverse referents
0	Irrelevant	The general model can be easier to explain to people and to understand. It can also be easier to make than the specific model	No knowledge of diverse referents

(level 0). We believed that a student responding at level 1 might have diverse referents in mind without expressing this idea explicitly. Equally possible, such a student had in mind multiple referents that were, in fact, outwardly similar (i.e. non-diverse). Given this uncertainty, we labelled level 1 responses as 'possibly knowing' that models had diverse referents. By contrast, a level 2 response required a definitive indication that a model's referents would be outwardly different from each other. Thus, we labelled this type of response as 'knowing' that models had diverse referents. Presumably, this knowing entailed an understanding that transformation was required to instantiate the model on any one of the referents. As a check on this presumption, we analysed the think-aloud interviews for evidence of thinking about transformation.

Synthesis model item. The purpose of the synthesis model item was to identify students in the modelling group who were successful in the synthesis process. The proxy for success was whether or not students used the synthesis model to explain desert formation, as opposed to reverting to pre-synthesis rain shadow or trade wind desert explanations. Accordingly, the item prompt read, 'In general, why do deserts form where they do?' Student responses were coded into three categories: (a) synthesis model; (b) one or both pre-synthesis explanations; or (c) no relevant explanation.

Item coding procedure. Two researchers independently coded a random sample of 20% of the data for each item and then checked for interrater agreement. Interrater agreement exceeded 85% for all items. Once agreement was established, a single researcher coded the remaining data for a given item.

Think-aloud procedure. The think-aloud interviews were conducted with 12 students from one modelling classroom about two weeks after they took the posttest. Working alone in the presence of a research assistant who audio-recorded the session, each student talked out loud continuously (see Ericsson & Simon, 1998) while they responded in writing to the abstracting and transferability items. In a procedural error, the researcher failed to collect written responses from two students working the transferability item. Thus, we had only 10 students with complete data for this item. Students' written responses during the think-alouds were coded according to the item scoring criteria of Tables 3 and 4. Their corresponding verbal outputs were transcribed and open coded by both authors working together to generate categories of thinking relative to the ideas outlined under each item description. Finally, a frequency table was generated by tallying whether or not each category of thinking was present in a given response.

Results

Abstracting

Table 5 shows the test results for the abstracting item. In the modelling group, there was a measurable shift in the distribution of scores from before to after instruction. At pretest, nearly 60% of students scored at the lowest level, no capability in abstracting, compared to about 40% of students who scored at either partial or full capability. After instruction, the distribution flipped, with about 40% having no capability in abstracting, and roughly 60% with partial or full capability. The change was most pronounced at the full capability level. At pretest, only 6% of students were in this category, compared to 19% at posttest. The overall shift in scores was statistically significant in a chi-square test, χ^2 (2, $N = 175$) = 17.43, $p < .001$, with alpha set at .025 to account for a simultaneous test on the transferability item. For the control group, by contrast, gains were negligible. Thus, there was

Table 5. Results for the abstracting item.

	Capability in abstracting		
	0 – None	1 – Partial	2 – Full
Group	*N* (%)	*N* (%)	*N* (%)
Modelling			
Pretest	101 (57.7)	64 (36.6)	10 (5.7)
Posttest	72 (41.1)	70 (40.0)	33 (18.9)
Control			
Pretest	34 (69.4)	13 (26.5)	2 (4.1)
Posttest	31 (63.3)	18 (36.7)	0 (0.0)

no discernable benefit from taking the test twice, making it unlikely that the gains in the modelling group occurred due to a testing effect.

Results from the think-aloud interviews for the abstracting item indicated modelling students' scores for 'capability in abstracting' (i.e. full, partial, none) were rooted in genuine, albeit rudimentary, understanding of abstracting. As shown in Table 6, three forms of thinking emerged in relation to this understanding as students worked on the item: anticipating transferability, engaging in the abstracting process itself, and reflecting on the nature of an abstraction.

Thoughts anticipating transferability reflected two complementary requirements that students set for their models which we describe as escaping the constraints of surface details, and accommodating diverse referents. Five of eight correctly-scoring students (i.e. those with scores of partial or full capability on written responses) verbalised at least one of these ideas about transferability. In an example of escaping constraints, a student verbalised that he was drawing 'a little pointy roof' and continued on to say, 'it doesn't have to be pointy, though.' Thus, he knew that he should escape the constraint of 'being pointy' to represent the essential characteristics of a roof, though he did not say how to do this (As the drawing was not an abstraction of this feature, his written response score was accordingly lower.). For the second type of requirement, accommodating diverse referents, a student mused before drawing that 'an ordinary house can have anything from four walls, for a really boring one, to many more.' Thus, and again while not saying how to do it (i.e. by abstracting), she recognised that she needed to accommodate houses of differing complexities in her model. Parenthetically, it is telling that thinking about transferability surfaced on the abstracting item. Specifically, it shows the

Table 6. Results from the think-aloud interviews for the abstracting item.

		Verbal output (think-aloud)				
		Anticipating transferability		Engaging in the abstracting process		Reflecting on abstraction
Student	Written response score (capability in abstracting)	Escaping surface details	Accommodating diverse referents	Broadening depictively	Broadening conceptually	Various
1	Full	X	X			X
2	Full	X			X	
3	Full			X		
4	Full		X		X	
5	Partial	X		X		X
6	Partial			X		
7	Partial				X	X
8	Partial		X		X	
9	None					
10	None					
11	None	X				
12	None					
Example Responses		Got a little pointy roof, it doesn't have to be pointy though	Like an igloo has one wall, and an ordinary house can have anything from four walls for a really boring one to many more	Like if I can do the general shape of the house … and hmm … that's kind of a general	there is a solid shell on the outside of the house	If you really really strip it all down, what you have got

reciprocal relationship between the two concepts in student thinking. To us, it also underscores the distinct value of transferability as a separate idea from abstraction, a point we argue in the Discussion.

Engaging in the abstracting process consisted of expanding each aspect of the referent to make it more general, what we call broadening. Students broadened both conceptually, by bringing in generalised ideas, and depictively, by drawing generalised shapes during brief pauses in their verbal output. Seven of eight correctly-scoring students exhibited at least one of these two processes. In an example of conceptual broadening, a student decided that the exterior walls in their model would be 'a solid shell on the outside.' Depictive broadening was exemplified by a student who said 'like if I can do the general shape of the house … and [drawing two circular shapes] hmm … that's kind of a general.' The fact that students conceptualised abstracting as broadening was not surprising, given that both their experience with synthesis and the item prompt focused on constructing a 'general' model. Therefore, we interpreted it as valid thinking about abstraction within the context of the study. On the other hand, having increased breadth, or generality, is a narrow conception of abstraction that would not apply to many types of models. We explore this issue further in the Discussion.

Finally, three correctly-scoring students reflected on the nature of abstraction by verbalising what makes a model 'general.' One talked about removing extraneous features to uncover the essence of the phenomenon, 'if you really, really strip it down, what you have got.' A second discussed the importance of not complicating the model with specific features, 'I guess we should just keep it basic.' The third expressed the need for broad applicability, as in 'a good model for that would be something pretty generic.' All of these show evidence that the students' ideas about abstracting extended beyond mere procedure to encompass knowledge of the endpoint of the process, an abstraction.

Finally, and to summarise the abstracting think-aloud results, the information in Table 6 makes it clear that students' 'full' or 'partial' capability in abstracting as measured by the item reflected rudimentary knowledge of abstracting, not sophisticated knowledge. Their notion of how to abstract was to broaden particular aspects, a crude conception, which though effective in the context of the item, would not work in many modelling scenarios (see Discussion for more on this point). Similarly, their ideas about abstractions as being generic or having removed surface features from the original neglect, or at least under-conceptualize, the idea of pulling essential structure away from the original that lies at the heart of abstraction. Nevertheless, and especially when combined with the tendency to anticipate transferability, the envelope of students' ideas shows them forming the goal to abstract, and having functional, albeit rough, ideas of how to achieve this goal. Arguably, this result represents a significant step forward in capability, which is what we wanted to uncover.

Transferability

Table 7 shows the results for the transferability item. Looking at the modelling group, scores again shifted toward higher levels of understanding after instruction. On the pretest, only about 20% of modelling students either knew or possibly knew that models could accommodate diverse referents. At posttest, over 40% of students showed clear evidence of understanding this point (level 2), and another 20% possibly understood it (level 1). A chi-square test for the modelling group, with the alpha value again set at .025,

Table 7. Results for the transferability item.

	Knowledge that models had diverse referents		
	0 – Did not know	1 – Possibly knew	2 – Knew
Group	*N* (%)	*N* (%)	*N* (%)
Modelling			
Pretest	140 (81.4)	20 (11.6)	12 (7.0)
Posttest	67 (38.3)	35 (20.0)	73 (41.7)
Control			
Pretest	40 (85.1)	5 (10.6)	2 (4.3)
Posttest	38 (77.6)	7 (14.3)	4 (8.2)

indicated the shift from pre- to posttest was statistically significant, χ^2 (2, $N = 175$) = 75.37, $p < .001$. As with abstraction, we observed no corresponding shift in the control group for transferability, making it unlikely that gains in the modelling group reflected a testing effect.

Similar to the abstracting item, the think-aloud interviews for the transferability item showed that modelling students' knowledge of transferability (i.e. knowing, possibly knowing, and not knowing models had diverse referents) was rudimentary, but genuine. As shown in Table 8, there were two aspects of this understanding: range of applicability and breadth of explanation. Both aligned well with knowing that general models could accommodate diverse referents.

Range of applicability encompassed three types of utterance. The first two expressed either covering multiple scenarios or covering a general class of scenarios, in line with the demarcation in coding written responses as 'possibly knew' and 'knew,' that models must accommodate diverse referents. As an example of the former, a student said 'the general model could explain other situations.' An example of the latter was, 'it [a general model] can explain more about a genre of situations instead of just one.' A comparison of the presence or absence of these utterances to the written response codes at the left of the table shows that some students whose written responses were confined to

Table 8. Results from the think-aloud interviews for the transferability item.

		Verbal output (think-aloud)				
		Range of applicability			Breadth of explanation	
Student	Written response score (knowledge of diverse referents)	Covers multiple scenarios	Covers a class of scenarios	Not restricted to one scenario	Gives broad explanation	Foregoes precise details
1	Knew		X	X	X	
9	Knew		X	X	X	
2	Possibly knew	X		X		
4	Possibly knew				X	X
11	Possibly knew	X		X		
5	Possibly knew		X	X		
3	Possibly knew	X	X	X		
12	Possibly knew		X	X		
10	Did not know		X		X	X
6	Did not know	X				
Example Responses		the general model could explain other situations	it can explain more about a genre of situations instead of just one	not [about] just that one specific situation	gives the general idea of the entire thing	[does not give] a clear idea … of the specific event

multiple scenarios actually verbalised thoughts about diverse scenarios while forming these responses. Thus, our interpretation of level 1 responses as 'possibly knew' about diverse referents seemed fitting. Shifting to the third type of utterance for range of applicability, which was that a model was not restricted to a single scenario, these utterances were often ambiguous as to multiple scenarios vs. class scenarios. Nevertheless they consistently indicated that models stood at a distance from their referents, suggestive of transformation. As an example, a student said, 'it's not about just that one specific situation.' Seven of the eight students who had correct written responses (i.e. levels 1 or 2) had verbal outputs of either the 'class' or 'not restricted' types.

Turning to the second aspect of understanding transferability, breadth of explanation, two complementary types of utterances expressed this idea about general models: gives a broad explanation and foregoes precise details. These were verbalised by four students overall. As an example of the broad explanation type, a student said that the general model 'gives the general idea of the entire thing.' Exemplifying the forgoing of precise details, a student said 'it does not give a clear idea of a specific event.' As evidenced by these examples, both types of utterance were more about abstraction than transferability. Thus, just as the abstracting item generated complementary thoughts about transferability, the transferability item generated thoughts about abstraction. To us, this complementarity was fitting, given the reciprocal nature of the two ideas.

Predictive validity of learning the synthesis model

Table 9 is a contingency analysis comparing the distribution of student responses on the synthesis model item to the distributions for the abstracting and transferability items. It shows that modelling students who spontaneously explained desert formation in terms of the synthesis model were much more likely to score well on the abstracting and transferability items than students who relied on pre-synthesis ideas to form their explanations. For abstracting, about a third of students who relied on the synthesis model to explain desert formation, 32%, had full capability in abstracting, while only 9% who relied on pre-synthesis ideas had this capability. Similarly, 64% of students who based their explanations on the synthesis model also knew that models should have transferability to diverse referents, compared to 29% who based their explanations on pre-synthesis ideas. These results, though perhaps easily explained by the principle that students who do well on one test item are apt to do well on others, are at least consistent with the

Table 9. Strength of association between scores for abstraction and transferability and knowing the synthesis model (modelling group only).

	Capability in abstracting			Transferability – Having diverse referents		
How students explained desert formation	None *N* (%)	Partial *N* (%)	Full *N* (%)	Did not know *N* (%)	Possibly knew *N* (%)	Knew *N* (%)
No relevant explanation *N* = 43 (25)	19 (44)	17 (40)	7 (16)	21 (49)	9 (21)	13 (30)
Using pre-synthesis source *N* = 70 (40)	36 (51)	28 (40)	6 (9)	35 (50)	15 (21)	20 (29)
Using synthesis model *N* = 62 (35)	17 (27)	25 (40)	20 (32)	11 (18)	11 (18)	40 (64)

claim that students learned about abstraction and transferability during the brief synthesis modelling period and not the longer desert formation unit that preceded it.

Score differences across teachers

The design of the study precipitated the question of whether learning about abstraction and transferability could have been driven more by especially well-prepared teachers than by synthesis modelling (see Figure 3, Design Feature 5). Such an outcome would obviously undermine claims about learnability. Accordingly, we compared the five teachers' test scores, taking the mean increase from pre- to posttest (i.e. difference scores) for the sum of the abstraction and transferability items for each teacher. The results were that two of the modelling teachers did have mean difference scores that were much higher than the other three, $M_{\text{Higher2}} = 1.2$, SD = 1.0 compared to $M_{\text{Lower3}} = 0.64$, SD = 0.86. Further, the two higher-scoring teachers were among the three who had helped develop the desert formation unit and synthesis modelling instruction. On the other hand, the results were not driven solely by these teachers: the mean for the lower-scoring teachers was still sizable and much higher than that of the control, $M_{\text{control}} = 0.13$, SD = 0.62. Thus, while two well-prepared teachers did make a large contribution to the overall result, all five made contributions.

Discussion

One of the two purposes of the study was to find out whether students could learn about abstraction and transferability in a brief intervention, indicating the viability of these concepts as targets for instruction in modelling. The results show that students in the modelling group did, in fact, make substantial progress in learning about both concepts during a 50-minute class period, with no explicit teaching of either concept. For abstraction, they increased their ability to generate the abstract model structure. Think-aloud interviews showed students to be conscious of the need for abstraction, and that they were also armed with useful, if rudimentary, ideas about how to achieve it (i.e. through broadening). For transferability, students learned that 'general' models could accommodate referents that were outwardly different from each other, reflecting tacit awareness that a transformation is involved when such a model represents any particular referent. This awareness was backed by ideas, visible in the think-aloud interviews, to the effect that models must escape the constraints of specific scenarios and capture the essence of a class of them. Therefore, and taking the two main results together, it seems that student understanding of abstraction and transferability were viable – perhaps even promising – targets for instruction.

Importantly, students' knowledge of abstraction and transferability represented a limited understanding of what these terms were meant to convey about structure-preserving transformations in models. Their understanding can be evaluated in terms of the three characteristics of the transformations in the model-referent relationship described earlier (see Table 1). Considering first what gets transformed, students' notion of this would have been information that all members of a class of phenomena had in common, as in the components of a dwelling or pathways for moisture transport; but would not have gone as far as structure that was purposefully selected. It is unlikely they would have thought of the components they transformed as 'structure', as this was not something they learned explicility. Furthermore, as students transformed all of the shared structures they encountered, they might not have conceptualised the components as being 'selected', as one might, for instance, in the case of a solar system model that emphasised orbital eccentricity, but did not include other information.

Second, students' notion of how structure is transformed (i.e. their notion of abstraction) was that structural elements should be generalised, by broadening them conceptually or depictively. Accordingly, students did not develop a conception of how and why a model might be an abstraction apart from being a generalisation. Thus, they might not recognise a schematic diagram like a subway map as being an abstraction, as the abstraction would have been achieved through simplification, not generalisation. Similarly, they might not see how exaggerating the eccentricity of orbits in a solar system model would be an abstraction, since emphasising this element of structure would appear to make the model less general, not more. These limitations make sense given that the abstracting processes they used emphasised generalisation. However, even if students had been exposed to a broader experience of abstracting, keeping to this one model construction process would likely have constrained their ideas about model abstraction. For example, they might struggle to recognise an equation abducted from a pattern of data as an abstraction, and hence, a model, because it would not come from the expected abstracting process.

Lastly, for the directions of information flow in transformation, students knew that a model which had been abstracted from a set of sources, would in turn, accommodate a range of referents. This was a tacit, as opposed to explicit, understanding of the opposing directions of information flow. Arguably, however, it was a well-grounded understanding, as evidenced by modelling students' responses to the abstracting and transferability items as a pair: on the first item, they knew to abstract the structure of a dwelling that would accommodate diverse referents; on the second, they stated that models in general should represent diverse referents. Moreover, think-aloud responses showed students thinking about the bi-directionality of transformation as they responded to each of the items. For the abstracting item, they thought of the need for the model to have transferability; for the transferability item, they thought of how a model was an abstraction. Thus, while students did not explicitly know that transformation was bi-directional in model-referent relationships, they seemed very well prepared to learn this important idea.

Returning to our overall claim that the data show the viability of teaching about abstraction and transferability as ways of conceptualising structure-preserving transformations in models, it is evident that the observed learning represented steps, not strides, toward the sophisticated understanding of these transformations. However, as students' ideas about abstraction and transferability were in a good general direction, and because students generated these organically from their modelling experience (i.e. they did not memorise information we gave them), we do claim these steps represent a positive start on this understanding, and thus a positive indicator of their viability as targets of instruction.

Our second purpose was to determine if there was a relationship between learning about abstraction and transferability and thinking of models as literal interpretations, which would serve as an indication of the value of these concepts as targets for instruction. According to the results, learning about abstraction and transferability directly opposed literal interpretation. Students were keenly aware that the 'general' model they constructed could not, and indeed, should not, retain the surface features of its referents. Instead, they understood the model to be a transformation on its referents. Similarly, they knew that general models, having broad explanatory power, could not (and should not) reflect their referents directly. Moreover, learning on these two fronts was not visibly inhibited by stable beliefs to the effect that models should be literal interpretations. This outcome suggests that literal interpretation is not a misconception as indicated by early researchers on this phenomenon (e.g. Grosslight et al., 1991; Ryan & Aikenhead, 1992). More likely (and framed here as a hypothesis), it is a natural response to a vacuum of ideas about structure-preserving transformations that occur between a model and its referent. Reasoning from this hypothesis, given that the obvious criterion of quality for a model is fidelity to its referent, then students who do not anticipate transformation between the model and its referent will naturally swing toward capturing all available information, and hence, literal interpretation. On the other hand, if students have some knowledge about structure-preserving transformations and why they are needed – even burgeoning knowledge as in the present study – there is no reason that literal interpretation would intrude on their thinking.

In light of the preceding hypothesis, two distinct recommendations for instruction can be advanced for combatting literal interpretation. Both are focused on teaching a true conception of the model-referent relationship as opposed to simply teaching against literal interpretation. The first would be to employ model construction processes that include well-defined cognitive supports for generating abstract model structure. In the present study, we employed synthesis, which supports structure-abstraction processes. However, abductive approaches could also provide cognitive support by having defined processes for arriving at an abstract model structure. The distinct selection practices found in physics education modelling exemplify an abductive approach (Brewe & Sawtelle, 2018; Halloun, 1996, 1997; Hestenes, 1987, 1997).

The second recommendation would be to teach directly about structure-preserving transformations. As just discussed, students should experience phenomena directly, in the model construction process, but they should also reflect on these experiences to develop their ideas. This experiential learning could be supplemented with auxiliary activities like going over examples of models that highlight various transformations. Useful activities abound in the modelling literature, for instance showing several models of a phenomenon that transform structure in different ways (Harrison & Treagust, 2000; Lin & Chen, 2002). There could also be a procedural emphasis within model construction activities, for instance emphasising these transformations by challenging students to construct a model that captures structure best while actually resembling the referent the least. Gallery walks and other formative assessment tasks could then query for this quality.

Both of the above recommendations would depend on providing language for describing structure-preserving transformations. In the present study, we have proposed using language around abstraction and transferability to do this. Reviewing each of these terms, we think that 'abstraction of structure' has value because it carries fairly specific information about both what gets transformed (i.e. essential structure) and how it is transformed (i.e. it is an abstraction). Further, it is a simple expression and thus easy to remember. On the other hand, as we pointed out in the front of the article, the term abstraction puts pressure on usage, especially the need to differentiate abstraction, the noun describing what all models are, from abstracting, the verb describing a unique way of constructing models. Nonetheless, we think the advantages of parsimony go far toward outweighing the need for careful usage.

In a similar vein, we argue that the phrase 'models have transferability' is useful as a compact way of describing opposing directions of information flow in transformation. One advantage of the root word, transfer, is that it has established meaning in educational circles, namely the application of an idea under transformation instead of directly. Thus, having transferability is arguably more useful than similar terms like having instantiability, which might otherwise be preferred for their precision. Transferability also raises the question of superfluity, on the argument that 'abstraction of structure' could stand alone to cover both directions of information flow, not just model to referent, but also referent to model. That is, to say a model's structure has been pulled away from a referent is sufficient to establish that transformation is required to map that structure onto the referent. Thus, perhaps transferability should not rise to the level of a defined characteristic of models. We argue that it should rise to this level, if only for pedagogical reasons. Abstraction is largely a descriptive concept, indicating what a model is as a representation (i.e. an idea that is pulled away from its referent(s)). By contrast, transferability has a more functional orientation, indicating what a model does as a representation (i.e. corresponds to its referent(s) under transformation). Arguably, this functional quality makes transferability useful as a back-door way of thinking about what it means for a model to be an abstraction. To use an example from earlier in this discussion, an equation abducted from data might fall outside a student's conception of an abstraction, if, as in the present study, they had only learned about abstraction through abstracting processes. But on the subject of transferability, the equation would fit the student's conception very well, through its applicability to diverse referents. Thus, recognising transferability would offer an entry point for understanding why the equation is an abstraction. Putting this idea into operation, students might be given the task of evaluating a mixed set of models and non-models. They could use transferability as an initial operational criterion for discriminating the two phenomena. Presumably, a similar task that asked students to judge whether or not a set of representations were abstractions would be more challenging because it would be harder to operationalise abstractness.

One area of immediate utility for the concepts of abstraction and transferability (or similar terminology for structure-preserving transformations) is assessment frameworks for models and modelling (e.g. Fortus et al., 2016; Krell et al., 2014; Schwarz et al., 2009; Schwarz et al., 2012; zu Belzen et al., 2019). As an example, Upmeir zu Belzen and colleagues (2019) advanced a framework with three levels of thinking about models in relation to their referents: copies; idealised representations; and theoretical constructions. With the present study in mind, these levels might be reframed as models are not transformations; models are unspecified transformations; and models are abstractions. To us, this revised version would provide a clearer, more complete, and more unidimensional definition of what to assess of students' thinking about models than the existing proposal (see Shemwell and Capps, 2019 for detailed arguments). As a more dramatic example, Fortus and colleagues' (2016) framework had four levels of sophistication in thinking about the model-referent relationship. At the lowest level, models were seen as literal interpretations of a single phenomenon. The three higher levels, however, struggled to convey what made models not literal. They were: showing what was unavailable to the senses; being combinable to represent more than one phenomenon; and having the capacity to represent unknown phenomena. Reminiscent of our study's transferability item, the two highest levels indirectly indicated transferability by pointing to models' capacity to represent diverse referents. Thus, they might be rendered more directly and unidimensionally as models are translatable (i.e. to multiple referents) and models are transferable (i.e. to diverse referents).

The present study was not designed to test the efficacy of synthesis modelling compared to other approaches to model construction for learning about abstraction and transferability. Nevertheless, the evidence is suggestive on this point, prompting us to make two observations. First, synthesis modelling seemed to help students learn these two concepts, presumably because it gave them a generative way to experience the concepts. Specifically, students were compelled to abstract the underlying structure in order to accommodate multiple, outwardly different sources. And, as structure emerged, they tested its transferability to each of the referents. Although there was no direct evidence for these processes, they seem probable for reasons articulated in studies of analogical transfer (Catrambone & Holyoak, 1989; Gentner et al., 2003; Gick & Holyoak, 1983). Secondly, and as a caveat to the first observation, our version of synthesis left students with a narrow conception of abstraction, namely generalisation, and a limited concept of transferability, that is, applicability to diverse referents. Thus, although synthesis may be useful for teaching about abstraction and transferability, it cannot stand on its own. Nor is synthesis the only available option. As pointed out previously, an abstracting process like synthesis is just one way to help students generate models that are obvious abstractions and have salient transferability (e.g. consider abduction). Similarly, as indicated by earlier portions of this discussion, there are many ways to experience, and thus learn about abstraction and transferability outside of the model construction process. Thus, while it is our proposal to teach that models are abstractions that have transferability, we are not proposing that this should be done through synthesis.

Conclusion

At the outset of this article, we explained how literal interpretation, or the tendency for students to think that a model should reflect every available feature of its referent, continues to cast a shadow on the landscape of modelling instruction. To be sure, in the era since Grosslight and colleague's (1991) initial discovery of literal interpretation, modelling researchers have avoided suggesting that this phenomenon stems from persistent false beliefs (e.g. Crawford & Cullin, 2004; Harrison & Treagust, 2000; Krajcik & Merritt, 2012; Lin & Chen, 2002; Saari & Viiri, 2003). On the other hand, authors have not directly challenged the misconception view. Perhaps it is time to do so. In the present study, we saw no evidence of stable beliefs about veridicality getting in the way of a truer understanding of the model-referent relationship. On the contrary, once students were introduced to structure-preserving transformations in this relationship via synthesis, they readily generated them in a new context, and they explained their advantages in terms of explanatory power. If we have read this evidence correctly, then much more is to be gained by teaching what models are than teaching what they are not. To make this happen, instructors will need defined ways for students to think about structure-preserving transformations in the model-referent relationship. As a start, we have proposed a working framework to the effect that models are abstractions, and they have transferability. Doubtless, these constructions can be improved upon. Indeed, improving on them would seem a worthwhile effort given the present study's finding that, under the right conditions, these kinds of ideas are accessible to students after a brief experience of model construction.

Note

1. This term was meant to cue thinking about the "general model" students had constructed during synthesis.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Directorate for Education and Human Resources [grant number 1720996].

ORCID

Daniel K. Capps http://orcid.org/0000-0002-2017-9218
Jonathan T. Shemwell http://orcid.org/0000-0001-8681-0378

References

Australian Curriculum and Assessment Reporting Authority. (2014). Foundation to Curriculum: Science. http://www.australiancurriculum.edu.au/science/curriculum

Brewe, E., & Sawtelle, V. (2018). Modelling instruction for university physics: Examining the theory in practice. *European Journal of Physics*, *39*(5), 054001. https://doi.org/10.1088/1361-6404/aac236

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1147. https://doi.org/10.1037/0278-7393.15.6.1147

Clement, J. J., & Núñez-Oviedo, M. (2003). *Abduction and analogy in scientific model construction*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Philadelphia, PA.

Clement, J. J., & Steinberg, M. S. (2002). Step-wise evolution of mental models of electric circuits: A "learning-aloud" case study. *The Journal of the Learning Sciences*, *11*(4), 389–452. https://doi.org/10.1207/S15327809JLS1104_1

College Board. (2013). Research. Retrieved from: https://research.collegeboard.org/

Crawford, B. A., & Cullin, M. J. (2004). Supporting prospective teachers' conceptions of modelling in science. *International Journal of Science Education*, *26*(11), 1379–1401. https://doi.org/10.1080/09500690410001673775

Dogan, N., & Abd-El-Khalick, F. (2008). Turkish grade 10 students' and science teachers' conceptions of nature of science: A national study. *Journal of Research in Science Teaching*, *45*(10), 1083. https://doi.org/10.1002/tea.20243

Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity*, *5*(3), 178–186. https://doi.org/10.1207/s15327884mca0503_3

Fortus, D., Shwartz, Y., & Rosenfeld, S. (2016). High school students' meta-modeling knowledge. *Research in Science Education*, *46*(6), 787–810. https://doi.org/10.1007/s11165-015-9480-z

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7* (2), 155–170. https://doi.org/10.1207/s15516709cog0702_3

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, *95*(2), 393–408. https://doi.org/10.1037/0022-0663.95.2.393

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38. https://doi.org/10.1016/0010-0285(83)90002-6

Gilbert, J. K., & Justi, R. (2016). *Modelling-based teaching in science education* (Vol. 9). Springer International Publishing.

Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, *28*(9), 799–822. https://doi.org/10.1002/tea.3660280907

Halloun, I. (1996). Schematic modeling for meaningful learning of physics. *Journal of Research in Science Teaching*, *339*(9), 1019–1041. https://doi.org/10.1002/(SICI)1098-2736(199611)33:9<1019::AID-TEA4>3.0.CO;2-I

Halloun, I. (1997). Schematic concepts for schematic models of the real world: The Newtonian concept of force. *Science Education*, *82*(2), 239–263. https://doi.org/10.1002/(SICI)1098-237X(199804)82:2<239::AID-SCE7>3.0.CO;2-F

Harrison, A. G., & Treagust, D. F. (1996). Secondary students' mental models of atoms and molecules: Implications for teaching chemistry. *Science Education*, *80*(5), 509–534. https://doi.org/10.1002/(SICI)1098-237X(199609)80:5<509::AID-SCE2>3.0.CO;2-F

Harrison, A. G., & Treagust, D. F. (2000). A typology of school science models. *International Journal of Science Education*, *22*(9), 1011–1026. https://doi.org/10.1080/095006900416884

Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics*, *55*(5), 440–454. https://doi.org/10.1119/1.15129

Hestenes, D. (1997, March). Modeling methodology for physics teachers. In *AIP conference proceedings* (Vol. 399, No. 1, pp. 935–958). AIP.

Ingham, A. M., & Gilbert, J. K. (1991). The use of analogue models by students of chemistry at higher education level. *International Journal of Science Education*, *13*(2), 193–202. https://doi.org/10.1080/0950069910130206

Krajcik, J., & Merritt, J. (2012). Engaging students in scientific practices: What does constructing and revising models look like in the science classroom? *Science & Children*, *49*(7), 10–13.

Krell, M., zu Belzen, A. U., & Krüger, D. (2014). Students' levels of understanding models and modelling in biology: Global or aspect-dependent? *Research in Science Education*, *44*(1), 109–132. https://doi.org/10.1007/s11165-013-9365-y

Lin, H. S., & Chen, C. C. (2002). Promoting preservice chemistry teachers' understanding about the nature of science through history. *Journal of Research in Science Teaching*, *39*(9), 773–792. https://doi.org/10.1002/tea.10045

Loewenstein, J. (2017). Structure mapping and vocabularies for thinking. *Topics in Cognitive Science*, *9*(3), 842–858. https://doi.org/10.1111/tops.12280

Merrian-Webster. (2019). *Abstract*. Retrieved from: https://www.merriam-webster.com/dictionary/abstract

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K–8*. R.A. Duschl, H.A. Schweingruber, & A.W. Shouse, Eds. Washington: The National Academies Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.

Nersessian, N. J. (2008). *Creating scientific concepts*. MIT Press.

NGSS Lead States. (2013). *Next Generation science Standards: For States, By States*. The National Academies Press.

Passmore, C., Gouvea, J. S., & Giere, R. (2014). Models in science and in learning science: Focusing scientific practice on sense-making. In *International handbook of research in history, philosophy and science teaching* (pp. 1171–1202). Springer. https://doi.org10.1007/978-94-007-7654-8_36

Ryan, A. G., & Aikenhead, G. S. (1992). Students' preconceptions about the epistemology of science. *Science Education*, *76*(6), 559–580. https://doi.org/10.1002/sce.3730760602

Saari, H., & Viiri, J. (2003). A research-based teaching sequence for teaching the concept of modelling to seventh-grade students. *International Journal of Science Education*, *25*(11), 1333–1352. https://doi.org/10.1080/0950069032000052081

Schwarz, C., Reiser, B. J., Acher, A., Kenyon, L., & Fortus, D. (2012). MoDeLS: Challenges in defining a learning progression for scientific modeling. In A. Alonzo & A. Gotwals (Eds.), *Learning progressions in science* (pp. 101–137). Sense Publishers.

Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., Shwartz, Y., Hug, B., & Krajcik, J. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, *46*(6), 632–654. https://doi.org/10.1002/tea.20311

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

Shemwell, J. T., & Capps, D. K. (2019). Learning abstraction as a modeling competence. In A. Upmeier zu Belzen, D. Krüger, & J. van Driel (Eds.), *Towards a competence-based view on models and modeling in science education* (Vol. 12, pp. 291–307). Springer. https://doi.org/10.1007/978-3-030-30255-9_17

SLO. (1996). *Voorlichtingsbrochure havo/vwo Algemene natuurwetenschappen* [Information brochure on public understanding of science].

State Department of Education. (2014). *School nutrition*. Retrieved from: https://www.state.gov/doe/home

Tasquier, G., Levrini, O., & Dillon, J. (2016). Exploring students' epistemological knowledge of models and modelling in science: Results from a teaching/learning experience on climate change. *International Journal of Science Education*, *38*(4), 539–563. https://doi.org/10.1080/09500693.2016.1148828

United States Census Bureau. (2014). Quick facts: Median household income 2014. Retrieved from: https://www.census.gov/quickfacts/table/PST045214

zu Belzen, A. U., van Drie, J., & Krüger, D. (2019). Introducing a framework for modeling competence. In *Towards a competence-based view on models and modeling in science education* (pp. 3–19). Springer.