COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 36 (2012) 1–61

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/j.1551-6709.2011.01207.x

Misconceived Causal Explanations for Emergent Processes

Michelene T. H. Chi,[a] Rod D. Roscoe,[b] James D. Slotta,[c] Marguerite Roy,[d]
Catherine C. Chase[e]

[a]*Department of Psychology, Arizona State University*
[b]*Learning Sciences Institute, Arizona State University*
[c]*Ontario Institute for Studies in Education, University of Toronto*
[d]*Medical Council of Canada*
[e]*Department of Psychology and the Human Computer Interaction Institute,Carnegie Mellon University*

Received 19 January 2010; received in revised form 5 April 2011; accepted 14 April 2011

Abstract

Studies exploring how students learn and understand science processes such as *diffusion* and *natural selection* typically find that students provide misconceived explanations of how the patterns of such processes arise (such as why giraffes' necks get longer over generations, or how ink dropped into water appears to ''flow''). Instead of explaining the patterns of these processes as emerging from the collective interactions of all the agents (e.g., both the water and the ink molecules), students often explain the pattern as being caused by controlling agents with intentional goals, as well as express a variety of many other misconceived notions. In this article, we provide a hypothesis for *what* constitutes a misconceived explanation; *why* misconceived explanations are so prevalent, robust, and resistant to instruction; and offer one approach of *how* they may be overcome. In particular, we hypothesize that students misunderstand many science processes because they rely on a generalized version of narrative schemas and scripts (referred to here as a Direct-causal Schema) to interpret them. For science processes that are sequential and stage-like, such as cycles of moon, circulation of blood, stages of mitosis, and photosynthesis, a Direct-causal Schema is adequate for correct understanding. However, for science processes that are non-sequential (or emergent), such as diffusion, natural selection, osmosis, and heat flow, using a Direct Schema to understand these processes will lead to robust misconceptions. Instead, a different type of general schema may be required to interpret non-sequential processes, which we refer to as an Emergent-causal Schema. We propose that students lack this Emergent Schema and teaching it to them may help them learn and understand emergent kinds of science processes such as *diffusion*. Our study found that directly teaching students this Emergent Schema led to increased learning of the process of *diffusion*. This article presents a fine-grained characterization of each type of Schema, our instructional intervention, the successes we have achieved, and the lessons we have learned.

Correspondence should be sent to Michelene T. H. Chi, Department of Psychology, Arizona State University, PO Box 871104, Tempe, AZ 85287-1104. E-mail: Michelene.Chi@asu.edu

Keywords: Causal explanations; Diffusion; Emergence; Learning; Misconceptions; Simulations; Science processes

1. Introduction

This article addresses why students have such difficulty understanding many concepts of processes. Very little research in psychology and related disciplines has studied process concepts; instead the focus has been predominantly on object or entity concepts (Gentner & Kurtz, 2005; Medin, Lynch, & Solomon, 2000). This work focuses on process concepts because they are associated with an acute learning problem that has been documented mostly in the science education literature for several decades. The problem is referred to as having naïve conceptions or misconceptions, which are qualitative incorrect explanations that are expressed prior to and even after formal instruction. Their existence and robust persistence after formal instruction and other direct forms of interventions has been documented in over 8,338 papers, as tracked by Duit (2008). For example, one misconception that students hold about electrical current is that it contains or is comprised of substance-like entities that fill the circuit from point to point, affecting each component (such as a light bulb) in turn within the circuit as the substance-like entities arrive at the component (Closset, 1983; Reiner, Slotta, Chi, & Resnick, 2000; Shipstone, 1984). In contrast to this sequential ‘‘circuit-filling model,’’ a more accurate ‘‘bicycle-chain model’’ of electrical current is that electrons are in all parts of a simple circuit, all moving at the same time, and with the same flow rate when the circuit achieves a steady state (Dupin & Johsua, 1984; Grotzer & Sudbury, 2000; Perkins & Grotzer, 2005).

Although process misconceptions are discussed mostly in the science education literature, process concepts are prevalent in many other disciplines, and misconceptions are rampant in these other domains as well. Examples include concepts such as supply and demand in economics, function in mathematics, and historical or political events (Lohmann, 1994; Torney-Purta, 1994). Moreover, during the decades of work on misconceptions, instructional interventions focusing on remediating deep-rooted misconceptions have not met with great success. Many articles have reviewed the existence of misconceptions and cited efforts that fail to remove them (Chi & Ohlsson, 2005, p. 387; Confrey, 1990).

In this article, we address the *what*, *why*, and *how* of misconceptions. That is, we provide an account of *what* the nature of a misconceived qualitative explanation *is* in a much more detailed way than the common description included in our electrical current example: the misconceived explanation being that electrical current is more like a circuit-filling model and should instead be more like a bicycle-chain model. Moreover, our explanation of *what* misconceived explanations *are* will avoid resorting to the use of expert and technical jargons such as the ‘‘same flow rate’’ or ‘‘steady state.’’ In brief, for *what* it is, we propose that a misconceived explanation is an incorrect ‘‘inter-level’’ causal explanation for explaining the pattern of a process from the collective interactions of the agents. For *why* such explanations are misconceived, we propose that students are relying on a familiar linear, narrative-like structure or schema that they have developed for narratives, and scripts to interpret ''emergent'' processes, resulting in misconceptions. They should instead be using a complex causality or ''emergent schema'' to interpret and learn these non-sequential emergent processes in both formal and informal contexts (Chi, 2005, 2008; Jacobson & Wilensky, 2006; Perkins & Grotzer, 2005; Wilensky & Resnick, 1999). Finally, *how* might we remediate such misconceived explanations? The method of instruction we propose is derived from our theoretical analyses of *what* misconceptions are and *why* they are generated. Based on our theory, we offer a novel way of remediating misconceived explanations, consisting of teaching students a concept-general ''emergent schema'' of which they have no knowledge.

This article is divided into four sections. The first section lays out our theoretical analyses of the *what* and *why* of misconceptions. The second section points out the challenges of *how* to overcome misconceptions through instruction. The third section describes our preliminary instructional intervention attempt. The final section highlights the main ideas, assertions, and assumptions of our work and situates our work in the context of some prevailing ideas, assumptions, and assertions.

2. What are misconceptions and why are they prevalent and robust? An analysis of processes

Our account of the *what* and *why* of misconceived explanations is interwoven and presented in this first section. The argument and theoretical work consist of (a) analyzing the kind of schema that students have developed from their everyday encounters with ''sequential'' processes (to be called a Direct-causal Schema or a Direct Schema), and abstracting a generalized set of characteristics for such a Direct Schema, (b) decomposing processes into two levels, a macro or pattern level and a micro or agent level, (c) defining ''sequential'' processes and showing how they can be interpreted by a Direct Schema, (d) defining non-sequential ''emergent'' processes and showing how they can be interpreted by an Emergent Schema, and finally (e) showing that misconceptions arise from applying the Direct Schema to interpret and explain ''emergent'' processes. Early renditions of these ideas have been described in Chi (2005), but now we further expand upon and clarify these ideas.

2.1. An analysis of stories, everyday events, and sequential processes

Although little research has focused on students' understanding of processes, two lines of psychological research do shed light on how processes are understood. One line of work is related to how children comprehend stories (Bower, Black, & Turner, 1979; Mandler & Johnson, 1977; Rumelhart, 1978; Stein & Glenn, 1979). These classic papers described comprehension of simple stories as developing and instantiating a *narrative schema*, an organized knowledge structure consisting of the following components: (a) an *initiating or triggering event* or initial conditions, which may cause a *protagonist* (the central character) to elicit some internal responses and (b) which may result in *overt actions* or a *series of actions or interactions*, (c) the series of (inter)actions are *logically* related with *causal* or *enabling* relations and then (d) carried out in the interest of pursuing or attaining a *goal*, and (e) when the goal is attained, the story ends or *concludes*, and the (inter)actions *terminate*. The italicized characteristics, in essence, constitute the components of a narrative schema for understanding simple stories.

A second line of research relevant to processes is how adults and children understand everyday events for which they have personal experiences by formulating *scripts* (Schank & Abelson, 1977). In a restaurant script, for example, children form ideas such as: an *initiating event* (feeling hungry), followed by a series of *actions* (the hungry person going inside a restaurant, sitting down, ordering from a menu, etc.), in pursuit of satisfying a *goal* (of not being hungry). When the goal is met, then one is *done* and may leave the restaurant (i.e., the process of eating at a restaurant is over).

We can abstract a common set of qualitative characteristics for both narrative schemas and scripts. For example, aside from an *initiating event* that triggers the process, the process itself typically has an *agent with some controlling or special status*, in the sense that this central agent partially controls the processes and makes decisions; there is also a *series of (inter)actions* occurring *sequentially*, undertaken in attempting to achieve a *goal*. Moreover, these (inter)actions are *logically related*, with relations such as *dependency*, *enabling*, or *causation*; and *intentionally* undertaken to achieve a *goal*. Finally, these (inter)actions conclude or *terminate* when the goal is met. We can consider this set of common characteristics to represent a Direct Schema for understanding processes such as narratives and everyday events.

It may be reasonable to assume that children have developed such a Direct-causal Schema from their daily everyday experiences and their exposure to stories, and such a Direct Schema is adequate for them to interpret and understand everyday common processes. For instance, when children watch wolves hunting prey on the Discovery channel, they can understand a hunt because, like a story, there is an alpha wolf (an *agent* like a protagonist) in the hunt who is more responsible or more *controlling* in the hunt and thus may have some *special status*; there is a *series of interactions* occurring *sequentially* in attempting to achieve the *goal* (e.g., some wolves chase the prey, followed by other wolves distracting the prey, followed by other wolves blocking the prey's escape path, followed by the final kill). These series of interactions are *logically related*, with relations such as *dependency* (the final kill does not occur until the prey is tired and cannot escape), *enabling* (the wolves that blocked the escape path did so in order to enable the alpha wolf to catch the prey more readily), or *causation* (the prey is captured quickly because the alpha is strong and fast or because the ''blockers'' are blocking particularly well). These interactions are *goal directed* (the wolves act this way because they are hungry), and they conclude or *terminate* when the goal is met (after the final kill and the wolves have eaten, they are no longer hungry so the hunt is done).

We assume that children, when they enter school, also rely on this same Direct Schema that they have developed for everyday processes to understand and interpret processes that they have to learn in the context of their science classes. For example, if they have to read about the human circulatory system, they might interpret it using a Direct Schema, in which the heart is the *agent/component* with a more *special status* (in pumping blood), and there is *a sequence of interactions* when blood flows through the chambers of the heart and then to the rest of the body, and so on. Similarly, if they are reading about historical political events such as the Holocaust, they may be applying their Direct-causal Schema to attribute such events to a single cause, such as the intentions of a controlling leader (Torney-Purta, 1994).

2.2. Seven common properties shared by all processes

Before we can understand how misconceptions occur, we need to first define some common properties of all processes. We do so using two examples here: a very familiar everyday process of a baseball game and the less familiar process of moths getting darker over generations throughout the industrial regions of England due to a number of interrelated factors. A common explanation for the less familiar latter example is that the abundance of coal smoke from factories killed the lichen-encrusted trees, and blackened the trees and building walls, making the lighter peppered-color moths more visible to hungry birds and, therefore, eaten more frequently compared to darker, less visible moths. These two examples will be used in the proceeding paragraphs to illustrate the seven common properties they share.

First, all processes can be conceived of as involving at least two distinct ‘‘levels’’ that can be uniquely described. We refer to one level as the *pattern* level. A *pattern* of a process is often what can be perceived and is often dynamic (although it can appear static). For example, the pattern of a baseball game is the changes in the distribution of players on the bases, changes in the score on the scoreboard, and so forth. In the case of moths, the pattern is the moths getting darker over generations. In the context of science processes, the pattern level is also referred to as the *macro* level.

Another level for all processes is the *agent* level. The *agents* are the individuated, separate elements of a process. In a baseball game, the players are the agents, and in the evolution of moths, bird and moths are the agents. In the context of science processes, the agent level is also referred to as the *micro* level.

Second, agents of all processes can cohere in some ways and be clustered into *subgroups*, through either visual similarity or functionality. For example, players in a baseball game typically wear one of two sets of identical uniforms, so the players are clustered into teams through visually similar uniforms. However, within a team, other coherent subgroups exist due to similarity in function or roles. In a baseball game, the players who catch high-flying balls far out in the field are called the ‘‘outfielders’’ as they have the similar role of catching balls that are hit into the outfield, versus the players who play the position of shortstop. Similarly, in the process of moths getting darker, the moths and birds can be partitioned as two subgroups. Thus, subgroups can be seen and conceived of as cohered components of processes with identifiable behavior, structure, or function, and serve as the aggregated components at the pattern level. Aggregated components at the pattern level can consist of either subgroups with visibly individuated agents (as in visually distinguishable team members or observable individual moths) or with invisible individuated agents. For example, in an everyday experience of pouring cream into coffee, after the cream has been poured into coffee, cream is an aggregate component of the pattern of cream diffusing in coffee (with swirling or flowing pattern), but one cannot see the individual cream molecules. Thus, regardless of whether the individual agents are visible or invisible in an aggregate component (such as members of a team versus molecules of cream), we will see that clustered subgroups or aggregated components play a significant (but incorrect) role in misconceived explanations of emergent processes.

Third, for all processes, the behavior at each level is different. For a baseball game, the behavior at the pattern level can be visualized as changes in the playing field, such as all the bases becoming loaded, and so on. The changes can be correctly described by referring to a subgroup as responsible for the pattern, such as ''the red team just struck out'' (that's why the players are going on and off the field). The changes at the pattern level of all processes can often be described qualitatively as either increasing or decreasing (either in number, as in an increase in the number of players on bases or the number of cells as they divide in the process of mitosis; or in color, as in the case of moths; or in intensity; or in speed; and so on). Thus, we assume that pattern-level behavior is more intuitively understandable to students, as most people can relatively easily understand the concepts of ''increasing'' and ''decreasing.'' At the agent level, the behavior consists of the interactions among the individual players, such as a player swinging at the ball that the pitcher pitches. Similarly, for moths, the behavior at the pattern level consists of changes in the color of the moths over generations and the behavior at the agent level consists of individual moths being eaten or not eaten by a bird. In short, behaviors at the pattern level are often more qualitative and intuitive, and similar across many processes and describable as increasing or decreasing in one of several dimensions, whereas interactions among agents are numerous and peculiar to specific processes (i.e., birds eating moths is uniquely different from batters swinging at a ball).

Fourth, for all processes, the behavior at the agent level can and should be considered as interactions between agents rather than as actions of individual agents. But in everyday language, one often refers to the agents' behavior as their individual actions rather than interactions with another agent. For example, in a baseball game, one often refers to the *action* of the batter (such as the batter *swings* at the ball), rather than the *interaction* of the batter with the pitcher (such as the batter *swings* at the ball that the pitcher *pitches*). This is because in a baseball game, the agents tend to interact with a constrained or restricted alternative set of agents (i.e., a batter always interacts with the pitcher); therefore, in everyday language, it is redundant to say that the *batter swings at the ball that the pitcher pitches* since the pitcher always pitches the ball and the batter only interacts with the pitcher when the batter comes to bat. In contrast, in the process of moths getting darker, one does refer to the interactions of birds and moths, as *birds eating moths*, in part because moths could be eaten by other creatures besides birds. Thus, even though everyday language tends to describe the batter's swinging behavior as the batter's *action*, it is in fact an *interaction* between the batter and the pitcher. Thus, agents in all processes interact.

Fifth, for all processes, there are distinct conditions or parameters controlling the behavior at each level. For example, the speed with which the moths get darker is determined by parameters such as *how many birds are in the area* and *how fast the coal smoke is being* *produced* by the factories. But these conditions that control the pattern level are distinct from the conditions that control each agent's behavior or interactions. At the agent level, whether a moth is eaten or not eaten by a bird depends on conditions such as: *how visible a moth is while resting on a tree trunk*, *how hungry a specific nearby bird is*, and *whether a specific hungry bird actually sees a specific moth*. Thus, these conditions governing the interactions at the agent level are distinct from the conditions that govern the behavior at the pattern level.

Sixth, for all processes, all levels may be visible or invisible. Although patterns are usually perceivable, this is not always the case. The pattern of moths getting darker over generations is obviously not visible (unless it can be seen artificially, as in a computer animation). These changes in patterns, even if not visible, are not difficult for students to imagine, since, as we indicated earlier, they often involve patterns of increasing or decreasing in number, size, color, speed, and so forth. The agents can also be visible or invisible. Moths and birds, and baseball players are all visible, but sometimes the agents are minute, such as molecules, which are not visually perceivable.

Seventh, there is an endless amount of information one can learn about both levels of processes. For example, one can learn about the structure, behavior, and function of the agents, whether the agents are moths or baseball players, such as their average size, height, how they interact, the frequency of interactions, and so forth. Similarly, at the pattern level, there are various relationships one can learn between the conditions or parameters and the pattern itself, such as the relationship between increasing the amount of coal smoke expelled from the factories and how quickly moths will get darker. However, lacking information about either level is not the source of misconceptions.

In summary, there are seven commonalities that seem to be shared by all processes. First, all processes can be described in at least two levels: the pattern and the agents. One can of course decompose agents further into another level; for example, each moth or bird can be further decomposed into his/her genetic composition, and so forth. However, we focus on the pattern and the agent levels because these two levels are sufficient for explaining the nature of misconceptions. Second, at both the agent and pattern level, agents can be clustered into subgroups or aggregated components. Third, for all processes, there are distinct behaviors at each level. Fourth, the behavior at the agent level can be conceived of as interactions among agents rather than as actions of individual agents. Fifth, for all processes, the conditions that control the behavior at each level are also distinct. Sixth, for all processes, the components and behavior of each level may be visible or invisible. Finally, there can be an endless amount of information to be learned about each level of a process, but as will be described later, lacking information about a specific level is not the same problem as having misconceptions.

Although all processes share these seven commonalities, processes can be differentiated into two kinds: sequential and emergent. These are the labels we use to refer to them, even though these labels may not be scientific or technical terms used to describe these processes. Other terms used in the literature to describe similar distinctions between processes are linear and non-linear. The next two sections describe these two kinds of processes.

2.3. Sequential processes and how they can be interpreted by a Direct Schema

Processes such as wolves hunting prey, a baseball game, or blood circulating in the human body may be referred to as ‘‘sequential’’ processes because these processes can be decomposed or reduced into a sequence of subevents. (We had referred to them previously as ‘‘direct’’ processes in Chi, 2005.) For example, a baseball game can be decomposed into a sequence of innings, and blood circulating in the body can be reduced to the sequence of components or organs to which blood travels. When children enter middle school, they will have to learn many processes that are similar to the sequential processes that they have encountered in their everyday environment. For example, in the eighth-grade science texts, they have to learn about processes such as digestion of food, circulation of blood, stages of development, phases of the moon, phases of cell division (or mitosis), photosynthesis (stages of transformation from sun energy to chemical energy), respiration, and so forth.

Because these cyclical, reducible processes are similar to the everyday sequential processes students already know, we assume that students will not have trouble interpreting these processes in the context of a Direct Schema. That is not to say that these stage-like science processes are easy to learn because there is still an endless amount of technical knowledge that has to be acquired about all levels of a process. In photosynthesis, for instance, many facts have to be learned about the agents of processes (such as the structure, behavior, and function of a chloroplast); in addition, students also have to learn about the pattern of the process of photosynthesis in terms of the sequence of stages.

Nevertheless, whether or not students understand the technical details of each level of a sequential process, students are often required to provide explanations about the pattern of processes and their explanations for the cause of a pattern for a sequential process tend to be correct in terms of the ‘‘structure’’ of the explanations. For example, suppose a question is asked about the pattern of a wolf hunt, such as ‘‘Why did the hunt take so long?’’ Several potential explanations that students might give are as follows: (a) ‘‘because the young male wolves were not doing a good job’’ (i.e., the explanation identified a *subgroup* of agents, young male wolves, as the cause); or (b) the hunt took so long ‘‘because some of the male wolves were distracted by other small animals around’’ (referring to some *specific interactions* among *subgroups* of agents as the cause); or (c) ‘‘the hunt took so long because the grass was too tall in the field for the wolves to see the prey’’ (referring to some *initial condition* or setting as the cause for the observed pattern). Notice that answers (a) and (b) refer to a subgroup of agents or their interactions, and answer (c) refers to a condition at the pattern level. In all three cases, regardless of whether the answers are true, these explanations are acceptable in the ontological sense because the ‘‘structure’’ of these causal explanations can be characterized as ‘‘direct’’ or ‘‘indirect.’’ That is, the explanations appeal to a specific component of the process (some interactions of either *individual agents* or a *subgroup of agents*, or some condition) as either *directly* causing the pattern, or *indirectly* causing it through a mediating agent. Indirect means that certain components or (inter)actions can be set up as a precondition for a subsequent (inter)action, and so forth. Thus, these explanations correctly appeal to characteristics of a Direct Schema, as we have broadly characterized for narratives and scripts.

Moreover, despite the fact that a variety of acceptable answers or explanations can be given to a question about the duration of a wolf hunt (as illustrated in the preceding paragraph), making the answers appear piecemeal (diSessa, 1993), there is in fact a set of five common and specific ''attributes'' that can be identified to *characterize* these explanations qualitatively. One such attribute is whether one can give credit to the interaction of a *single* agent (or a *subgroup* of agents) as the cause of the pattern. For example, it may make sense to say for a specific hunt that ''the hunt ended quickly because the alpha wolf (a *single* agent) was particularly fast at chasing down the prey,'' or that ''the distracting wolves (a *subgroup* of agents) were particularly cunning at distracting the prey.'' Thus, whether true or not, it makes sense (from an ontological perspective, Chi, 1997) to point to or claim that *a specific local interaction* or *agent*, or *a subset of local interactions or agents*, is directly (if not completely) causing the speed or success of the outcome. Second, because it makes sense to claim, more or less, that a specific agent or the agent's interaction (or subgroups of agents) as directly responsible for some outcome, one can give *special* or *controlling status* to the interactions of single agents or subgroups of agents. In this case, if one claims that the distracting wolves in a specific hunt were particularly cunning, then the interactions of these distracting wolves can have a special status as causing the desired outcome. Third, there appears to be a *correspondence* or *alignment* between what goes on among the agents' interactions and the pattern. For instance, if the wolves in a particular pack are all fast and strong and thus can chase and tire the prey quickly, then the *faster* the wolves are in a pack, the *faster* the hunt will end. Fourth, interactions at the agent level are undertaken to achieve local goals but also with the *intention* or *purpose* of fulfilling the *global* or pattern-level goal. For example, the wolves chase the prey (*local goal*) because they want to feed their clan (*global goal*). A *local goal* refers to the goal of a specific interaction, whereas the *global goal* refers to the goal of the entire process. Finally, the causal ''summing mechanism'' that produces the pattern from the agents' local interactions can be characterized as *additive*, *composite*, or *chaining*. That is, the pattern seen in a hunt is the chaining of all the subevents (consisting of events such as ''the distracting wolves blocking,'' followed by ''the alpha wolf charging at the prey,'' etc.). It can also be seen as *additive* in the following example: Suppose while driving, one sees a pile of trash getting taller each day because each passer by throws some trash at the pile. This pattern of a growing pile is caused by each new piece of trash being thrown at the pile, thus adding it to the existing pile.

The five properties defined above, consisting of (a) *a single* agent or *a subgroup of* agents directly or indirectly causing the pattern, (b) having *controlling or special status*, (c) having interactions *correspond or align* with the pattern, (d) being *intentional* toward a *global goal*, and (e) contributing to a *chaining* or *additive summing* mechanism that leads to the pattern, will be called ''inter-level attributes'' because they characterize the relationships between the local interactions at the agent level and the resulting pattern. Thus, relating the agent level to the pattern level makes them ''inter level.'' These five inter-level attributes for sequential processes are summarized in Table 1 (left column). We propose that these attributes are parts of the structure of a Direct Schema and explanations of the pattern of a sequential process appear to conform to this structure in terms of their characteristic attributes.

Table 1
Five qualitative attributes for distinguishing the inter-level causal relationships that characterize how the agents' interactions (micro level) relate to the pattern (macro level) for sequential and emergent processes

Inter-level Attributes	
Sequential Processes	*Emergent* Processes
1. The interaction(s) of a *single* agent or a *subgroup* of agents can (in)directly ''cause'' the observable pattern	1. The interactions of the entire *collection* of *all* the agents together ''cause'' the observable pattern
2. The interaction(s) of one or more agents may have a more *controlling* (or *special*) *status* with respect to the pattern	2. All the interactions have *equal status* with respect to the pattern. (There is no leader or a subgroup of agents whose interactions are more controlling than others.)
3. Agents' interactions and the pattern behave in a *corresponding* or aligned way	3. Agents' interactions and the pattern can behave in *disjoint* or non-matching ways
4. Because specific interactions can directly or indirectly cause some changes in the pattern, some interactions are undertaken *intentionally* to produce the pattern or to achieve the *global goal.*	4. Interactions are undertaken by the agents with the intention of achieving *local goals* only, without any intention of causing the (changes) in the pattern. The pattern *emerges* from the local interactions of all the agents
5. The pattern is caused by the *additive* summing or chaining of a sequence of subevents	5. The pattern is caused by the *collective* summing or *net effect* of all the interactions at each point in time

2.4. Non-sequential processes and how they can be interpreted by an emergent schema

Students encounter other processes in school, such as watching a video in class about ants searching for food by marching in single, orderly files, or experiencing a bottleneck (or jamming) at a doorway when the fire alarm rings, and these may be referred to as non-sequential or ''emergent'' processes. Students also encounter other emergent processes in their science texts, such as diffusion, osmosis, electrical current, and floating and sinking. We can define an *emergent* process by differentiating it from a *sequential* one, using the following two contrasting examples about ways of finding food: wolves hunting their prey and ants finding their sources of food.

As described earlier, in wolves hunting a prey, the perceivable pattern of the process consists of chaining of a sequence of interactions or subevents, such as ''some wolves block the escape paths,'' followed by ''the alpha wolf catches the prey,'' and so forth. In ants searching for their sources of food, likewise the pattern one sees is a sequence of ants acting in an orderly fashion as well, consisting of subevents such as ''ants looking for food,'' then ''ants filing in a straight line toward one source of food,'' then ''ants dispersing,'' then ''congregating in a single-file line toward another source,'' and so forth. Thus, ants also seem to execute a sequence of (inter)actions, as in the case of wolves, directed at the global goal of finding food. In short, at the pattern level, both processes look similar in that there appears to be a sequence of interactions that are goal directed. We could even assume that there is a queen ant that dictates or directs the worker ants' actions, playing a similar role to that of the alpha wolf. In reality, although there is a queen ant, her sole function is to reproduce;

therefore, she has no role or control over the process of finding food by the worker ants. Thus, the queen ant does not tell the worker ants what to do or where to go, in the process of finding food. All of the ants follow the same set of simple rules: wander around aimlessly and randomly until food is found; once food is encountered, emit pheromones; then, follow smelled pheromones and also emit pheromones. This mechanism of finding food, consisting of all ants following the same set of rules or carrying out a similar set of interactions (walking, following, and emitting), creates the appearance that ants know where the food source is and then march in single-file lines toward it (as if this was intentional and goal-directed, which it is not). Thus, the cause of this ants-finding-food pattern is very different from the cause of the wolves hunting pattern. The pattern in the wolf hunt can be explained by a chain of different interactions or subevents occurring serially in a contingent sort of way. For a complete description of ants finding food, please see Resnick (1994).

An emergent process, then, seems to require a different sort of schema to understand and interpret it. If the causal explanation by which ants find sources of food is different from the causal explanation by which wolves hunt prey, how can we characterize how these processes differ? We have proposed a diametrically opposite set of qualitative attributes to characterize the inter-level causal explanations of how the interactions at the agent level collectively cause the behavior at the pattern level for emergent processes, as compared to sequential processes (Chi, 2005). Using ants and wolves as contrasting examples again, first, one cannot attribute a single ant (such as the queen ant) to the cause of finding food in a single-filed pattern, whereas one could attribute the alpha wolf, as a particularly strong and fast wolf, to the quick ending of a hunt. Second, consequently, *all* ants have *equal status* when it comes to finding food, whereas the alpha wolf might have a more *controlling special status*. Third, some ants might be moving slowly but that will not affect the single-filed marches toward sources of food. Thus, the speed of some ants' movement and the resulting single-filed pattern may be *disjoint*, whereas the speed that the wolves chase the prey align or *correspond* more directly with the speed of the hunt. Fourth, this means that specific local interactions of the wolves can directly affect the pattern and be *intentionally* undertaken to affect the pattern or the *global goal*. For example, ''blocking all escape paths'' can allow the prey to be captured more quickly; therefore, ''blocking all escape paths'' are interactions aimed at achieving the *local goal* of blocking the paths, but these local interactions are undertaken *intentionally* for the purpose of directly affecting a faster capture of the prey, aimed at achieving the *global goal* of the hunt. For ants, however, there is no direct or intentional relationship between some specific ants' interactions and the pattern of marching in a single-file line. Ants walk around randomly and follow and emit pheromones, and they do not do it intentionally to cause single files to occur nor does the sidetracking of any individual ant have any bearing on the overall pattern of marching in single files. In fact, ants do not have a *global goal* of walking in single-file lines; the pattern of single files emerges without any intention. Fifth, the summing mechanism that explains the pattern of a wolf hunt is composed of the *chaining* or *adding together* of a sequence of the individual interactions (*additive summing*), whereas the pattern of ants marching in single files is the *collective* outcome of all the ants interacting simultaneously and independently at the local level (to be called *collective summing* or *net effect*). These five characteristic attributes are shown in the right column of Table 1.

In sum, Table 1 shows two sets of opposite (or mutually exclusive) inter-level attributes for the two kinds of processes, sequential and emergent. Recall that the inter-level attributes refer to the causal relationships between the agents' interactions and the pattern. Moreover, these are *characteristic* attributes, not *defining* attributes (Smith & Medin, 1981) in that not all attributes of each kind of processes need to be embedded in every process. In addition, these characteristics are *qualitative*, suggesting that in and of themselves, these characteristic attributes are not difficult for middle-school students to understand; they consist of common ideas such as *single* or *all*, *special* or *equal status*, *corresponding* or *not corresponding (i.e., disjoint)*, *intentional* or *not intentional (i.e., emerges)*, *chaining* or *additive summing*. The only idea that may be more foreign and challenging for students to understand is the attribute of *collective summing* and how patterns can emerge from it. Collective summing requires computing a ''net effect.'' We will further explain this difficulty in understanding collective summing through the lens of misconceptions.

2.5. Hypothesis of what misconceptions are: Applying the Direct-causal Schema to explain emergent processes

What is a misconception then? Our central thesis proposes that misconceptions are largely flawed inter-level causal explanations of the patterns of processes, and they are flawed in ''structure'' (and perhaps also incorrect in other ways, such as the technical details). By flawed in ''structure,'' we mean that the ''type'' of explanations might manifest an inter-level attribute that is ontologically inappropriate, rather than the actual ''token'' (the instantiated attribute) being incorrect. For example, an explanation that says that *It is the queen ant that tells the other ants where to find food* is flawed in ''structure'' because it is appealing to the attribute of a *single agent* as the cause, so that it is not merely incorrect in the token (that the agent is *the queen ant* as opposed to some other *worker ant*). Thus, misconceptions reflect the use of attributes of an alternative Direct Schema to explain non-sequential processes that ought to be explained by attributes of an Emergent Schema.

In short, our hypothesis for *what* a misconception is implies that (a) learning more about the structure, behavior, and function of each level of a process, or (b) learning more about the conditions and how they affect the behavior of each level, or (c) making each level more visible through simulations, will not impact (improve or remove) misconceived causal explanations, because misconceived causal explanations do not pertain to a lack of knowledge about any single level nor are they caused by invisibility of either the pattern or the agents. Rather, misconceived explanations result from an appeal to inappropriate inter-level relationships between the interactions at the agents' level and the pattern.

Why are there misconceptions if there are two kinds of processes? Our theory assumes that there are two reasons. The first reason that we stated earlier is that at the pattern level, sequential processes look very similar to emergent processes. We had suggested that ants marching in a single-filed line toward a source of food, then dispersing, then marching in a single-file line again toward another source of food, looks like a sequence of subevents similar to a wolf hunt. Another example includes Canadian geese that fly in a V-formation that looks very similar to airplanes that sometimes fly in V-formation. However, the former is an emergent process, in that the geese each respond only to the immediate (local) space around them by finding a pocket of least air resistance and these collective interactions form the V, whereas the latter is a sequential process, in that the captain may have told the other pilots to move into a V-formation. Because of their perceptual similarities, people activate their Direct Schema to interpret emergent processes. The second reason is that students may not necessarily know an Emergent Schema. Thus, lacking the Emergent Schema dictates that misleading cues in the patterns will automatically activate the alternative Direct Schema for interpreting an emergent process, resulting in misconceived explanations of the behavior at the pattern level from the interactions at the agent level.

3. Instructional implications and challenges

Our hypothesis that misconceptions arise from the activation-of-an-inappropriate-schema has several implications, for it can be tested in multiple ways. In this section, we first present seven ways that such a hypothesis can either be tested or is already consistent with existing findings, followed by a discussion of the five significant challenges of using an instructional intervention approach.

3.1. Seven ways to test or confirm our hypothesis

Our hypothesis that misconceptions arise from the activation-of-an-inappropriate-schema must be testable and be consistent with existing findings in the literature. Below, we consider ways that our hypothesis can be tested and whether it is consistent with some of the existing findings in the literature. First, this hypothesis predicts that the consequence of using a Direct Schema to assimilate, instantiate, or more generally interpret processes that are emergent results in persistent, robust, and tenacious misconceptions. There is a well-established literature showing overwhelming existence of tenacious misconceptions for emergent processes such as *diffusion*, *natural selection*, and *heat transfer*, even after classroom instruction and other forms of intervention (e.g., Clement, Brown, & Zietsman, 1989; Driver, 1987; Licht, 1987; Posner, Strike, Hewson, & Gertzog, 1982; White & Frederiksen, 1990). This huge set of evidence, collected over decades, can be taken as support of our hypothesis and, in particular, our assumption that an Emergent Schema might be missing for most students, leading to robust misconceptions.

The second prediction from our hypothesis that can be tested is that students should have less (resistant or) persistent misconceptions when learning about sequential processes that can be interpreted by a Direct Schema, as compared to learning about an emergent process. Such a difference would constitute evidence in support of the greater ease of understanding sequential processes. We cited some indirect evidence in support of this difference in Chi (2005), comparing the greater ease of learning about blood flow (in the human circulatory system) versus the greater difficulty of learning about diffusion flow by eighth-grade students. By learning about blood flow, we meant only learning about the sequential locations and organs along the circulation path to which blood has to travel (as is typically presented in an eighth-grade science text), not about how the dynamics of flow is determined. However, it is difficult to make meaningful direct comparisons between learning of two different process concepts when the processes differ in so many ways.

Third, if students' explanations of emergent processes contain misconceptions that have characteristics of a Direct Schema, then this would provide confirming evidence suggesting that students are applying their Direct Schema to generate explanations of emergent processes. Besides the evidence to be presented in the third section of this article, we have made one attempt at verifying this prediction by coding students' explanations of the process of *speciation* for characteristics of a Direct Schema, and the evidence does support our hypothesis (Ferrari & Chi, 1998).

Fourth, an extension of this approach would be to carry out a comprehensive meta-analysis of all processes that have been investigated and to show that sequential-like processes are learned more easily (with less persistent misconceptions) than emergent processes. Such a comprehensive meta-analysis has not been undertaken, but we have suggested concrete ways to segregate different types of misconceptions, which would greatly contribute to such an analysis (Chi, 2008). That is, such an analysis would need to segregate misunderstandings due to incorrect or missing knowledge, in contrast to misconceived knowledge.

Fifth, our activation-of-an-inappropriate-schema hypothesis is consistent with other evidence about the nature of misconceptions. One sort of evidence pertains to findings characterizing misconceptions as naïve explanations that are generated prior to formal instruction, supporting our assertion that inter-level explanations pertain to qualitative ideas such as *controlling agent* or *intentional* that are available to students even without any formal instruction on a specific concept. That is, misconceptions can be generated from a Direct Schema prior to learning any formal technical knowledge about a specific science concept, consistent with the findings that misconceptions are naïve qualitative explanations that are flawed.

Sixth, our hypothesis is also consistent with findings showing that students, upon confrontation about an incorrect prediction or explanation of a particular phenomenon, can revise their own specific explanation of the phenomenon, and yet their explanations will continue to be incorrect, making their explanations appear piecemeal (diSessa, 1993). Our hypothesis suggests that this is because the revised explanations are still generated from an inappropriate Direct-causal Schema.

Finally, the most powerful and difficult way to support our hypothesis is to see whether teaching an Emergent Schema might enhance students' learning and understanding of emergent processes. This article evaluates our hypothesis by using such an instructional intervention approach.

3.2. Five challenges of an instructional intervention approach

Our theoretical analyses of processes and their representations in terms of Direct- and Emergent-causal Schemas suggest that if students are missing an Emergent Schema, then an obvious instructional approach might be to teach students such a schema directly. However, even with detailed specification of the inter-level attributes, these inter-level attributes only characterize the nature of the causal explanations of the patterns; the attributes themselves do not directly help students generate an explanation nor help them recognize when an emergent explanation is required. We describe five additional significant challenges we face in our attempt to help students develop an Emergent-causal Schema using an instructional intervention approach, and how we have attempted to address three of them.

3.2.1. How to discriminate a sequential from an emergent process

The first challenge is in teaching students to recognize and differentiate a sequential from an emergent process. Correct recognition is an important first step since both kinds of processes are embedded throughout a standard school curriculum, and often co-mingled in the same curriculum theme (and sometimes even in the same chapter of a textbook). For example, sequential processes such as mitosis, meiosis, photosynthesis, moon phases, and blood circulation are often covered alongside emergent processes such as diffusion, osmosis, natural selection, extinction, erosion, floating–sinking, and so forth, in a non-discriminating way.

Discriminating that there are two kinds of processes requiring two types of causal explanations is not trivial since sequential and emergent processes look similar at the pattern level, as pointed out earlier. Visual similarities clearly often cause students to categorize two concepts or phenomena as similar, according to standard assumption in cognitive psychology (Smith & Medin, 1981), even if they do not categorize exclusively based on similarity (Goldstone, 1994). Thus, the fact that emergent and sequential process look similar at the pattern level obviously can cause students to treat them as the same kind of concepts. For example, a young child is misled by perceptual similarity when he/she categorizes a whale as a fish. Given that the patterns of emergent processes often look and feel similar to sequential processes, how can students discriminate them? For example, in a bathtub, hot water flowing from one part of the bathtub to another feels like water flowing downstream, in that it is directional, and seems to flow from one location to another. Therefore, students misconceive of heat flow as water flow. Thus, instruction must help students to differentiate sequential from emergent processes, but how? It is not sufficient to simply show the agents at the agent level (as invisibility is believed by some researchers to be the cause of misunderstanding), because the agents of many emergent processes are quite visible (e.g., the ants searching for food in a single-file line are quite visible).

Our proposal is that students have to focus not just on the agents alone, nor on the specifics of the agents' interactions, but on the agents' interactions *relative to other* agents' interactions. Thus, these are second-order relationships that will be referred to here as ''second-order interaction features'' to contrast them with the meaning of ''features,'' which usually refers to *intrinsic* properties of an object or entity (Barr & Caplan, 1987) when considered in isolation (such as ''has wings'' for a bird). ''Second-order interaction features'' also differ from ''relational features'' (Gentner & Kurtz, 2005). ''Relational features'' refer to noun categories whose members share the same core relationships, but different arguments. For example, *bridge* is a relational category in which membership must satisfy a specified relationship, such as ''connecting X to Y'' in which X and Y can be two concrete locations, two generations, or two entities. Thus, a wooden plank structure that

Table 2
Five second-order interaction features distinguishing the qualitative relationships between some agents' interactions with other agents' interactions in sequential and emergent processes

Second-Order Interaction Features	
Sequential Processes	*Emergent* Processes
1. Agents interact in *distinct* ways	1. All agents interact in the same *uniform* way in the sense of following the same set of rules
2. Agents can interact with pre-determined or *restricted* others	2. All agents can interact with any other *random* agents
3. Agents interact *sequentially*	3. All agents can interact *simultaneously*
4. Agents' interactions are *contingent* or *logically dependent* on other agents' interactions	4. All agents interact *independently* of one another
5. Agents' interactions *terminate* when the pattern-level behavior stops	5. All agents interact *continuously* regardless of the status of the pattern, whether it is changing or at equilibrium

connects one side of a canyon to another side is a *bridge*, and so is a metal wire that connects one tooth to another tooth, because the two bridges share the same specific relation of ''connecting X to Y'' even though they differ in scale, materials, shape, and so forth. We will refer to such a relationship as a ''first-order relational feature.''

Our analyses of sequential and emergent processes have specified two contrasting sets of five ''second-order interaction features'' about the *relative relationships* among the interactions at the agent level (Table 2) that will offer a way for students to learn to discriminate the two processes (Chi, 2005). We briefly describe these ''second-order interaction features'' now, contrasting the ants and wolves as examples again. We will illustrate for the first feature only in how it exemplifies a second-order interaction feature, and not an intrinsic feature nor a first-order relational feature.

First, one perceptually available feature about the interactions of ants that are searching for food is that they all follow the same set of rules in their (inter)actions (they all walk around, emit, and follow pheromones). Following the same set of rules implies that interactions among all the ants are relatively the same, to be referred to as *uniform*, whereas different wolves follow different sets of rules (depending upon whether a wolf is an alpha, a ''distracter,'' or a ''blocker''), and so these interactions will be referred to as *distinct*. Thus, the label *uniform* does not refer to the sameness of the individual agents themselves, either in their structure, behavior, or function, in that the individual ants can be different in size, speed of moving, weight, and so forth. *Uniform* also does not mean that each interaction of one pair of ants must be identical to the interaction of another pair of ants, since one pair of ants can follow each other closely whereas another pair of ants can follow each other from a greater distance. Rather, *uniform* refers to the fact that *all the ants* searching for food are following the same set of rules regardless of what any single ant's actions are, so that one pair of ants' interactions is more or less the same as another pair of ants' interactions. Thus, *uniformity* refers to the ''sameness'' *in the rules of interactions* that govern some ants' behavior relative to other ants. This makes *uniform* a second-order feature of interactions.

Second, ants can interact with any other *random* ant (an ant can follow any other ant), whereas the wolves who block the escape path interact mostly with other ''blocking'' wolves in coordinating the blocking and the alpha wolf interacts primarily with the prey. Thus, the interactions of wolves are more *restricted* than *random.* Third, at any moment in time, all ants can follow other ants *simultaneously* or without regard to any order, whereas different interactions among wolves need to follow some kind of *sequential* order. Fourth, whether one ant follows another ant is totally *independent* of whether another ant follows yet another ant, whereas the ''distracting'' wolves must set up the escape path for the prey so that the ''blocking'' wolves can block the path. Thus, the interactions of the wolves are *contingent* or *logically dependent* upon some other interaction, whereas this is not the case with the ants' interactions. Again, to clarify further, we use the terms *dependent* or *independent* as second-order interaction features, referring to the relative dependent or independent nature of the relationships among the agents' interactions at any moment in time. We are not referring to dependence or independence of any interaction upon some prior interactions occurring before this moment in time, which can be true, as Gupta, Hammer, and Redish (2010) have pointed out. For example, ants may not emit pheromones until they have found food (so there may be some contingency or dependency on the order of interactions), but whether one ant is emitting pheromones because it found food is *independent* of whether another ant is emitting/detecting pheromones at the same time. Fifth, unlike the wolves, whose interactions *terminate* once they have eaten their prey, ants can *continue* to wander around and search for food. These diametrically opposite sets of ''second-order interaction features'' characterizing the relative relationships among the interactions of sequential and emergent processes are shown in Table 2.

In sum, we have specified five pairs of perceptible ''second-order interaction features'' at the level of agents' interactions that hopefully will allow students to discriminate sequential and emergent processes. These second-order interaction features refer to the characteristics of one interaction compared to another interaction. Unlike first-order relational categories, whose members share the same specific relations but different arguments, members of the Emergent process category would share the same second-order relations, but the specific interactions can differ across members. For example, the specific interactions of the process of ants finding food are walking-following-emitting, but the specific interactions of Canadian geese flying in a V-formation is for each goose to fly behind another goose in a pocket of least air resistance. Thus, the interaction of walking-following-emitting is not the same as finding pocket of least air resistance, but both processes share similar second-order features, in that the interactions undertaken by all the agents are governed by sets of *uniform* rules, the interactions all occur *simultaneously* and *independently*, and so forth.

These second-order interaction relationships are referred to as ''features'' because they can be seen (e.g., one can observe and see that *all ants are doing the same thing* of following-emitting, or one can observe that all ants are following each other at the same time or *simultaneously*). Even if they are not perceivable, they can be made visible through simulations. However, even though simulations can make the second-order interaction features visible, it is still important to point them out either explicitly or to scaffold students to see them because they may be difficult to notice on one's own. For example, Gentner and Kurtz

(2005) have summarized studies showing that first-order relational categories are more difficult to learn than entity categories; this suggests that second-order interaction relationships would be even more difficult to learn than first-order relationships. Thus, we assume that showing a simulation of the agents' interactions per se may not help students understand the relative relationships among the interactions, unless these ''second-order interaction features'' are pointed out explicitly or students are scaffolded to see them. So, even though young students can easily understand qualitative relational terms such as same/different, sequential/simultaneous, dependent/independent, intentional/not intentional, and so forth, it is still important to point out that these qualitative relational terms refer to the relationships among the interactions, not just the interactions themselves.

As in the case of inter-level attributes, second-order interaction features are also *characteristic* in that they are not *defining* (Smith & Medin, 1981), again meaning that not all processes must exhibit all of these features. For some processes, some features will be more salient than for other processes.

Finally, the qualitative characteristics that we have identified for both ''second-order interaction features'' and ''inter-level attributes'' (shown in Tables 1 and 2) are specified in sufficiently concrete terms so that they can be translated into instruction, and more important, they are not described in mathematical terms. In contrast, emergent processes are often described in the scientific literature using complex terms such as ''linear versus non-linear,'' ''self-organizing,'' (Casti, 1997; Gell-Mann, 1994), ''chaotic,'' and so forth, with difficult-to-understand properties such as ''equilibration,'' ''small actions leading to big effects,'' or ''decentralized.'' For example, the characteristic ''decentralized'' includes at least three of our inter-level attributes (''all,'' ''equal status,'' and ''local goals'').

3.2.2. How to avoid pattern-level explanations only

The second instructional challenge is to help students learn to give a deeper inter-level causal explanation for the behavior at the pattern level, as opposed to giving a shallower single-pattern-level explanation. By a single-pattern-level explanation, we mean an explanation appealing to the conditions, constraints, parameters, or any static component *within the pattern level* only, without resorting to the dynamics at the agent level. Using crowding at a doorway as an example, a single-pattern-level explanation for why students end up jamming at the door might be that the ''doorway is narrow'' or that ''there are too many students.'' Such single-level explanations, even though correct, point to either a static component of the process (a narrow opening) or some parameters or conditions of the agents (that there are too many students) governing the pattern, rather than to the dynamic interactions of the agents. A deeper inter-level causal explanation would appeal to the interactions of the agents as a cause for the pattern, such as that ''*all* the students are rushing to the door *at the same time*, and they are all *walking at about the same pace*, causing a jam to occur.'' This latter explanation is appealing to several emergent features and attributes, such as that *all* the students are behaving with *uniform* interactions (walking at the same pace) and *simultaneously* (at the same time). One way to understand that such an explanation is correct and deeper than the typical explanation is that a jam will not occur at the door if *not all* the students are rushing to the door (but only *some* students are rushing to the door), or if they are doing so *not at the same time* (but *sequentially*), or they are walking *at different paces/speeds* (i.e., they are walking at distinct pace). Thus, a deeper correct explanation is one that draws upon the correct inter-level attributes and features of the agents' interactions, because if these attributes and features were not present, then the emergent phenomenon (the jam) would not occur.

A single-pattern-level explanation is different from a misconceived explanation in one important way, besides the fact that a single-level explanation is correct at the pattern level whereas a misconceived explanation is incorrect. The important difference is that a misconceived explanation tends to appeal to an attribute and/or feature of the Direct Schema rather than the Emergent Schema. For example, a misconceived explanation in the crowding phenomenon might be that ''some of the big kids ran faster toward the door than the smaller kids,'' thus attributing a ''subgroup'' of agents' interactions (the ''big kids'') as a cause (compatible with sequential attribute #1 in Table 1), rather than a correct inter-level explanation that *all* (emergent attribute #1) the kids were running at the *same time* (emergent feature #3 in Table 2) and at roughly the *same pace* (emergent feature #1), and a jam just emerged. In short, students can give incorrect deep inter-level explanations despite being able to give correct shallower pattern-level explanations. Thus, we predict that giving correct pattern-level causal explanations is not as difficult to learn as giving correct inter-level causal explanations based on the dynamic interactions of the agents. Our prediction is supported by evidence showing the success of teaching students to give single-level explanations, such as explaining which bar (metal or wood) will heat flow faster (Linn & Hsi, 2000). This type of pattern-level explanation is easier for students to learn because the instruction is not targeted at removing misconceived inter-level explanations. Our current solution to the challenge of eliciting correct inter-level explanations from students is for them to learn the set of five inter-level attributes (Table 1) needed for generating a correct inter-level causal explanation of the pattern.

3.2.3. How to create a novel schema when learning is via assimilation

The third instructional challenge has to do with the fact that learning is a process of assimilation. That is, we learn new information by assimilating it with existing knowledge, thus raising the problem of the learning paradox (Bereiter, 1985). The paradox is this: If we learn by assimilation, then how can we possibly learn anything completely novel, since all new knowledge must be assimilated into an existing category or schema? If we define assimilation as consisting of three subprocesses: activating a relevant schema, integrating new information in the context of the activated schema, and repairing the activated schema with new information (Chi, Kristensen, & Roscoe, in press), then how can we help students develop a new Emergent Schema without it being assimilated into the Direct Schema? To address this challenge, we propose that a new schema can be created not by focusing on similarities, but by focusing on differences. That is, it is possible that a new schema can be created by contrasting its features and attributes from an existing Direct-causal Schema, since the features and attributes of Emergent and Direct Schemas are diametrically opposite. Although contrasting cases have been shown to be a powerful method of learning when both cases are presented (Gadgil, Nokes, & Chi, in press; Medin, Goldstone, & Gentner, 1993;

Schwartz & Bransford, 1998), we are proposing that a new case (or schema) can be created by contrasting it with an existing schema using *opposite/different* relationships, as opposed to contrasting similarities and differences between two available cases. That is, we assume that an Emergent Schema can be taught by a modified assimilation process, such that the assimilation process might entail three subprocesses of *activating* a relevant schema, *contrasting* new information in the context of the activated schema, and *creating* a new schema. Thus, we believe that a missing schema situation is tractable and underlies our schema-based instructional approach.

3.2.4. Emergent processes are sequential at both the pattern and the agent levels

The fourth instructional challenge in fostering understanding of emergent processes is that the pattern itself is actually a sequential process (with initial conditions defining the sequence of behavior, final conditions defining when the pattern terminates, and intervening variables that affect the pattern), and the local interactions at the agent level are sequential as well. Using crowding as an example, at the pattern level, the pattern of crowding at the doorway is caused by the initial condition (having many versus few children seated at their desks and all trying to run toward the doorway), and a static constraint such as the ‘‘width of the doorway,’’ which will affect how quickly the children get out of the doorway (or how quickly the jam at the doorway disappears). At the local agent level, whether a particular student in the crowd pushes on the student in front of him or the student to his right depends on local conditions, such as whether the student to the right is smaller and weighs less than him. Thus, at both the pattern and the agent levels, the behaviors of the pattern and the specific interactions of the agents are sequential, with conditional dependencies. The *only* aspect of this process that is emergent is how the behavior of the pattern is produced by the *collective interactions* of the agents. We have not addressed this challenge directly in this first instructional effort, namely the challenge of differentiating among three sets of behaviors: the goal-directed local behavior within specific interactions at the agents' level, the goal-directed global behavior at the pattern level, and the non-goal-directed emergent inter-level nature of the entire process.

3.2.5. The difficulty of teaching collective summing mechanisms

The fifth instructional challenge is that the collective summing mechanism that explains the pattern of many emergent processes taught in middle school is often quite difficult to explain. For instance, with molecular diffusion, the flow pattern can only be explained correctly by a ‘‘net effect’’ or collective type of summation mechanism, rather than an additive type. An additive summing mechanism simply adds each effect sequentially over time. A net effect, on the other hand, must sum all the positive and negative effects at an instance of time. For a familiar everyday example, if you have two or more sources of income and two or more debts, then your net asset at the end of the week is the sum of all your incomes and all your debts. The summing mechanism for diffusion flow is complicated not only because it requires a net effect type of collective summing (to be detailed later), but moreover, the net effect has to be computed across time in order to explain the flow pattern.

In contrast to diffusion, for some common everyday emergent phenomena, the pattern seems self-explanatory so that there is no need to invoke a summing mechanism. For example, the common emergent phenomenon of a flock of Canadian geese displaying a perceptual V pattern seems explainable by pointing out the geese's interacting rule of proximity and least effort (that is, each goose instinctively flies behind another goose in a pocket of least air resistance), and together, they form a V. Thus, the geese's formation of the emergent V pattern seems self-explanatory because the summing mechanism is a perceptual one, of proximity and/or spatial contiguity of all the geese' interactions.

We have sidestepped this very difficult instructional challenge in our first intervention effort presented here. That is, our instruction (below) did not systematically address and differentiate the mechanisms of *additive summing* versus *collective summing* (attribute #5 in Table 1), and our examples used for instruction all relied on simple perceptual summing mechanisms, like ''proximity.'' We note that in other intervention studies in which students are given opportunities to either explore or construct NetLogo simulations of emergence, the collective mechanism is computed by the NetLogo system itself, thus opaque to the students, and students are generally not required to learn the summing mechanism (Wilensky & Resnick, 1999).

In sum, five significant instructional challenges exist for teaching an emergent process from the perspective of the student (in that the student is mistaking it for a sequential process). In our first instructional attempt (to be described below), we have attempted to overcome the first three of these challenges, but we have not addressed the last two directly.

4. A preliminary instructional intervention study

Our analyses of students' misunderstanding suggest that instruction should focus on helping students develop an Emergent-causal Schema, so that they can *recognize* an emergent from a sequential kind of process and be able to *give correct deep inter-level explanations* of the causes of a pattern without misconceptions. We tested our theoretical analyses on students' learning of the emergent process *diffusion.* Before describing our instructional intervention study, we will briefly explain why we chose *diffusion* as our test concept and used computer simulations as part of our training materials. We will also illustrate what we mean by *collective summing* in the context of *diffusion.*

4.1. The concept of diffusion

Diffusion has been recognized not only as a central concept in biology (LaBarbera & Vogel, 1982), but *diffusion* is misunderstood in much of the same ways as other emergent processes. Students' misunderstanding is manifested in (a) the existence of misconceptions, (b) giving shallow (but correct) pattern-level explanations, and (c) difficulty in understanding (collective) changes in proportion. We elaborate on each below.

4.1.1. The existence of misconceptions or failing to give correct inter-level explanations

Many studies have documented that *diffusion* is a tenaciously misconceived concept (Marek, 1986; Odom, 1995; Zuckerman, 1994). Although many lists have been compiled with respect to what exactly the misconceptions about *diffusion* are, we will illustrate the underlying structure of these misconceptions in the context of our framework, taking two examples from a list composed by Meir, Perry, Stal, Maruca, and Klopfer (2005). Let us say that the phenomenon is explaining why ink, dropped initially into a left beaker full of water, visibly appears to flow to the right beaker when the two beakers are then connected by a tube. One common misconception is that ''Molecules have a directional motion toward lower concentration,'' explaining incorrectly that the ink molecules move to the right beaker because there are fewer ink molecules there (i.e., lower concentration). Although this is true that at the macro (pattern) level (ink visibly appears to flow from the left beaker to the right beaker), students think that the ink molecules move there because they *want to* (sequential attribute #4 of intentionality, Table 1) move from an area of high concentration of ink to an area of low concentration (Meir et al., 2005). The interpretation we provide is not that students are simply unaware of how molecules move at the molecular level. We believe that students can learn about how molecules interact without robust misunderstanding, such as learning that molecules move around randomly and collide with each other. (We had referred to this as learning the structure and behavior of individual agents.) We postulate that, in addition to appealing to the notion of *intentionality*, students' misconception pertains to their assumption that ink molecules ''move in one direction,'' and, in particular, ''move in the direction of the flow,'' which is what *causes* the pattern (Meir et al., 2005). Thus, this incorrectly displays the inter-level attribute of alignment or *correspondence* between the micro and macro levels (sequential attribute #3 in Table 1). This *correspondence* attribute has also been referred to as ''slippage between levels'' (Wilensky & Resnick, 1999), and this slippage occurs in misconceptions for many other processes, such as electricity (Sengupta & Wilensky, 2009). In short, applying the *correspondence* attribute reveals that they are using an attribute of a Direct Schema to generate their inter-level explanation.

Similarly, students often possess another misconception about *diffusion*: that equilibrium is *static*. Again, this misconception arises from the attribute of *correspondence* because students assume that molecules must terminate their motion if the pattern appears static (as in the case of equilibrium at the pattern level). Thus, they are applying the feature *terminate* (sequential feature #5 in Table 2) to the motion of the molecules and the attribute *correspondence* (sequential attribute #3 in Table 1) to align the motion at the molecular level with the movement at the pattern level. In brief, misconceptions are explanations that manifest an appeal to attributes and/or features from a Direct Schema.

4.1.2. Texts presenting correct single-level explanations

It is easy to understand why students have these misconceptions about *diffusion*. We have examined over half-a-dozen descriptions of *diffusion* in middle-school and high-school texts. Even in good middle-school texts, the process of *diffusion* is described only by specifying the initial and final conditions, without saying anything about the process itself. Suppose we are describing the diffusion of ink in water from one beaker to another, as described earlier. A popular eighth- to ninth-grade (*Exploring Life Science*, Maton et al., 1995, p. 84) text description is:

> Molecules of all substances are in constant motion, colliding continuously with one another. The motion causes the molecules to spread out. The molecules move from an area where there are more of them (higher concentration) to an area where there are fewer of them (lower concentration).

Notice that a middle-school text does describe how molecules behave (''colliding continuously''), so it is not a failure to describe the nature of the molecular interactions themselves. However, the text only describes the outcomes at the molecular level (i.e., that molecules spreading out) and the conditional differences (i.e., from higher concentration to lower concentration). The text does not specify *how* the molecules appear to move from higher to lower concentration; therefore, students are left to assume, from using their Direct-causal Schema, that the movement of the molecules is either *intentional* and/or directional, and *aligned* with the appearance of the flow. Thus, movement *terminates* when the flow appears to terminate.

What is the correct inter-level explanation? Let us consider a similar scenario of a high concentration of ink initially contained in a beaker of water (the left beaker shown in Fig. 1) connected by a tube to a right beaker initially full of water only. After a while, *diffusion* gives the appearance of ink flowing from the left beaker to the right beaker. When asked questions about the pattern, such as ''Why does the ink appear to flow from the left beaker to the right beaker?'' students typically answer by pointing to the conditions at the pattern level (such as that ''there was ink initially only in the left beaker''). Or if asked ''Why does

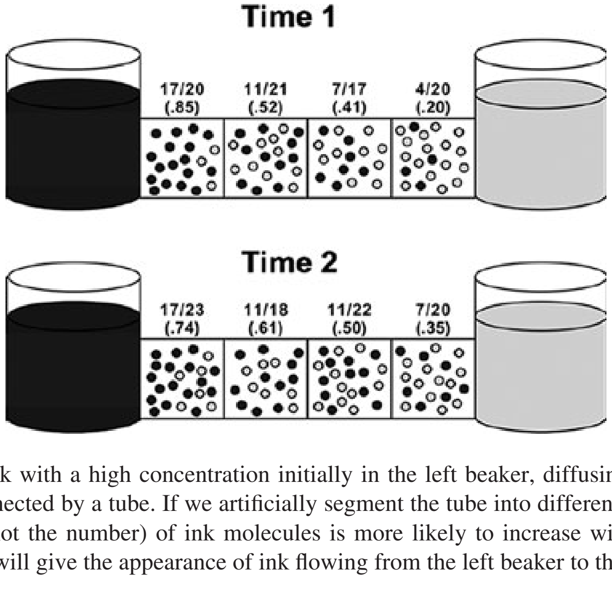


Fig. 1. Illustration of ink with a high concentration initially in the left beaker, diffusing to the right beaker, if the two beakers are connected by a tube. If we artificially segment the tube into different sections, what happens is that the proportion (not the number) of ink molecules is more likely to increase within each segment from Time 1 to Time 2. This will give the appearance of ink flowing from the left beaker to the right beaker.

the flow appear faster when two cups of ink is poured into the left beaker compared to one tablespoon of ink?'' a typical response might refer to the conditions or initial parameters (such as that ''the concentration of ink is greater with two cups of ink''). Although such explanations are technically correct (we had referred to them earlier as single-pattern-level explanations), the explanation is not inter-level in appealing to what goes on at the molecular level as a cause for the pattern. Being able to give correct single-pattern-level explanations does not address nor replace misconceptions, which can continue to persist since misconceptions reveal themselves when students give an inter-level explanation. Single-level explanations can occur at the agent level as well, but since students know less about the molecular level, this is not likely to occur as often. The correct inter-level explanation will be provided later.

We are not proposing that middle-school texts must explain college-level ideas, such as proportion changes to account for the pattern. We are suggesting that the notions of emergence can be taught to younger students so that they can start developing an Emergent Schema, setting the stage for correct interpretations of emergent processes when they encounter them later, without resorting to using a Direct Schema for assimilation.

4.1.3. Failing to understand the collective summing mechanism of proportion change

A third challenge for understanding *diffusion* relates to difficulty in understanding the collective summing mechanism that accounts for the flow pattern. In the scenario depicted in Fig. 1, what causes the perception of flow is that when molecules (both ink and water) collide randomly, by chance, some ink molecules will occupy the space that water molecules used to occupy, thus causing the proportion of ink molecules to generally increase from one segment of an area to the next. Fig. 1 shows two snapshots of the beaker situation at two time points: Time 1 and Time 2. If we divide the tube connecting the two beakers into four segments, then each segment at Time 1 has a different proportion of ink molecules, such as 17/20 (.85) in the first segment, 11/21 (.52) in the second segment, and so on. After some elapsed time in which all molecules collide with each other and bounce around, we take another snapshot at Time 2. Let us focus on the second segment. The second segment had a proportion of .52 at Time 1 but had a higher proportion of ink molecules at Time 2 (.61). The same increase occurs for the third and fourth segments. Thus, a flow is seen because of an increase in the proportion of ink molecules from Time 1 to Time 2 for each segment (this is difficult to see without it being shown dynamically), but the flow is not the result of the ink molecules moving from left to right from Time 1 to Time 2. Note that in the second segment, the number of ink molecules did not increase over time (they remained at 11) even though the proportion of ink molecules increased (giving a sense of flow). In addition, the proportion can also occasionally decrease from Time 1 to Time 2, as shown for the first segment. Thus, the perception of flow is caused by the general increase in the *proportion* of ink molecules, and not necessarily in the *absolute number* of ink molecules, as the case may be for the third and fourth segments (in which the number of ink molecules did increase from 7 to 11 in the third segment and from 4 to 7 in the fourth segment, but these increases in number corresponded to an increase in proportion, so it reflects an unfortunate spurious correlation (Cheng, 1997).

To summarize, the change in the proportion of ink molecules over time gives the appearance of flow. Therefore, it is not simply an increase in the *number* of ink molecules, but an increase in the *proportion* of ink molecules, as indicated in the second segment from Time 1 to Time 2. Students mistakenly assume, however, that flow is caused by the ink molecules moving from one location to the next (as in the way water flows downstream in a river), so that the *number* of ink molecules increases from one segment and time to another, which is only sometimes true (as shown in the third and fourth segments). The point is that this misconception indicates that students conceive of the summing mechanism to be an *additive* one. In reality, it is the proportion of ink molecules that increases, resulting in the perception of dark blue ink molecules moving to the right. Thus, the concept of *proportion* requires a consideration of all the molecules, both ink and water molecules. Thus, we refer to this summing mechanism as *collective summing* (emergent attribute #5, Table 1).

In short, *diffusion* is a complex non-sequential process that causes tenacious misconceptions because it gives the perception of a straightforward sequential flow in which ink molecules move from one location to another. But, in fact, some specific ink molecules may not be moving from one location to another, and yet flow can still be perceived.

4.2. Dynamic computer simulations for diffusion

Our training materials for the specific concept of *diffusion* included the use of dynamic simulations for many obvious reasons. The standard reasons are that dynamic simulations can enhance learning of complex science concepts because they can depict changes (Lowe, 2003), complex processes (Tversky, Morrison, & Betrancourt, 2002), and unobservable (micro level) components of processes (Sanger, Brecheisen, & Hynek, 2001). In short, dynamic simulation is a valuable tool in general, to use for instruction of concepts such as *diffusion.*

Aside from these general reasons, there are two additional reasons why we believe that dynamic simulations may be particularly appropriate for learning about emergent processes (Wilensky & Reisman, 2006). First, because of the simultaneous and collective nature of emergent processes, it can be difficult to visualize a macro-level pattern emerging from micro-level interactions. This is exemplified by the difficulty in visualizing the macro process emerging from the micro processes of liquid diffusion, as shown in Fig. 1 (one has to simultaneously scan across and down to get an idea of flow). Second, some attributes make more sense and are more believable when they are seen. In the case of *diffusion*, many students would not believe that some ink molecules move against (rather than along with) the direction of the ink flow (depicting the attribute of non-correspondence or ''disjoint,'' emergent attribute #3 in Table 1), unless they saw it for themselves, and a dynamic simulation makes that possible. We include dynamic simulations in our instructional materials because they are such powerful tools in the ways we have stated.

However, we also pointed out earlier that second-order interaction features may be difficult to notice unless they are specifically pointed out or highlighted somehow (as through scaffolding, for instance). Accordingly, we agree with others that attention guidance is helpful to students for learning from animations in general (de Koning, Tabbers, Rikers, & Paas, 2009; Moreno & Mayer, 2007), but we believe that it is particularly crucial for dynamic simulations of emergent processes since students must be scaffolded to notice the second-order interaction features and inter-level attributes. Thus, our simulations were supplemented with prompts as a scaffolding device.

In addition, because our theoretical analyses assume that misconceptions are a failure to give correct inter-level explanations, we deviate from other intervention studies by presenting simulations of both the macro-level pattern as well as the micro-(agent)-level behavior (i.e., to the best of our knowledge other interventions studies have not used the ''world view'' part of NetLogo type of simulations along with a micro view).

4.3. Our instructional approach

This section first describes the logic of our instructional approach, since it is unconventional and requires the development of a new set of instructional materials. (The details of our intervention study will be described in the next section.) Our theoretical account of misconceptions suggests that we should teach students about the ideas of an Emergent Schema, which then can be applied to many emergent processes that they have to learn. To that end, we need to design some sort of qualitative concept-general overview about processes (we will refer to it as the *Process Module*). What information should such a Process Module contain and what information should it not contain?

Our preliminary ideas are that such a Process Module should describe two different kinds of processes, illustrating the qualitative inter-level attributes and second-order interaction features of both sequential and emergent processes (as shown in Tables 1 and 2), so that the second-order features of agents' interactions might teach students how to recognize and discriminate one kind of process from another. In addition, the inter-level attributes might teach them how to generate a correct causal explanation for the observable patterns. This Process Module should use everyday examples (without any reference to scientific processes) in order to help students build an Emergent Schema that is applicable to many science processes.

The design of the Process Module was guided not only by our theoretical account of misconceptions, but it was also designed in ways that might overcome some of the challenges raised in the prior section. In addition, the design was also guided by several principles derived from cognitive science. For example, we used two examples in the Process Module because the cognitive science literature shows that two examples are significantly better for understanding than one example (Gick & Holyoak, 1983; Loewenstein, Thompson, & Gentner, 1999). Our instruction also adapted the contrasting cases technique, which has been shown to be a helpful instructional format (Bransford & Schwartz, 1999; Schwartz & Bransford, 1998), but we assumed that it might overcome the assimilation problem stated earlier. Moreover, given that students misconceive one kind of process for the other kind, contrasting differences may be particularly helpful in creating a new schema. Finally, we also embedded self-explanation prompts in the texts when appropriate, given that responding constructively to such prompts enhances learning (Chi, de Leeuw, Chiu, & LaVancher, 1994).

Since a Process Module describes ideas about emergence that are relevant to many science concepts introduced in a middle-school curricula, such as processes of diffusion, osmosis, extinction and other biological evolutionary processes, geological processes such as erosion and weather formation, floating and sinking, and forest fire spreading, it is important that descriptions of emergence are understandable to middle-school children. Moreover, because drop-out rate and lack of interest in science occurs at the early high school level (Hofstein & Welch, 1984; James & Smith, 1985), it is imperative that we can increase students' understanding of science concepts at the eighth- and/or ninth-grade levels, so that we can perhaps sustain their interest in science throughout high school. Therefore, our Process Module was tailored to the comprehension level of eighth and ninth graders.

Instructional interventions that attempt to teach a specific concept typically take the approach of improving explanations about the concept itself. We also take this approach. So, besides providing a concept-general Process Module, we also provided a concept-specific Emergent *Diffusion* Module consisting of an ''improved *diffusion* text,'' an ''instantiated *diffusion* text,'' along with macro and micro simulations. Details of how we designed the concept-specific Emergent *Diffusion* Module will be provided in the next section.

In the instructional literature, a standard way to assess the effectiveness of an instructional intervention is to compare how well students learn *diffusion* with our instructional intervention compared to regular classroom (the ''business as usual'') instruction, without any intervention. However, such a comparison is not that informative since intervention materials generally are improved, modified, or added to what students normally get in their classes. Therefore, it makes more sense to compare our intervention to an alternative intervention, consisting also of a concept-general module and a concept-specific *diffusion* module.

The alternative concept-general module could be conceived of as filler materials that students in the non-intervention group could study when students in the intervention group study the Process Module. So, in other words, in our intervention condition, our concept-general module is the Process Module, and in the non-intervention condition, the concept-general module is a filler module. This way, at least time-on-task and other variables (such as fatigue) are somewhat equated. But what should this filler module include and what should be our constraints? We imposed the constraint that a filler module should be an authentic concept-general module that might be introduced in many middle-school textbooks.

Accordingly, we searched for other general modules that students are exposed to in middle-school science courses. We found that typically, the introductory chapter of a science text contains a unit about the nature of science (NOS), covering materials such as the scientific method, the metric system, and so on. Such introductory chapters seem to appear in science texts for all grades and in all domains. A random sampling of several texts shows, for example, that Chapter 1 of a college biology text by Baker and Allen (*The Study of Biology*) is about *The Nature and Logic of Science* (24 pages long), and Chapter 1 of a high school chemistry text by Heath (*Chemistry*) is about *Activities of Science* (21 page long). In order not to pick a first chapter that might be tailored to chemistry or biology or physics, we picked the American Association for the Advancement of Science (AAAS) (1989) version (since, presumably, our Process Module is also applicable to many concepts in biology and chemistry, etc.).

4.4. Content of the intervention materials

This section provides specific details of all the training and assessment materials, which were entered into the Web-based Inquiry Science Environment platform (WISE, to be elaborated below) by two assistants at University of California, Berkeley. This allowed the instructional activities of the study to be carried out on laptops. The sequence of the presentation of the instructional materials and the assessments are depicted in Fig. 2. The intervention group of students used the materials depicted in the upper sequence (the Process Group) and the non-intervention group used the materials depicted in the lower sequence

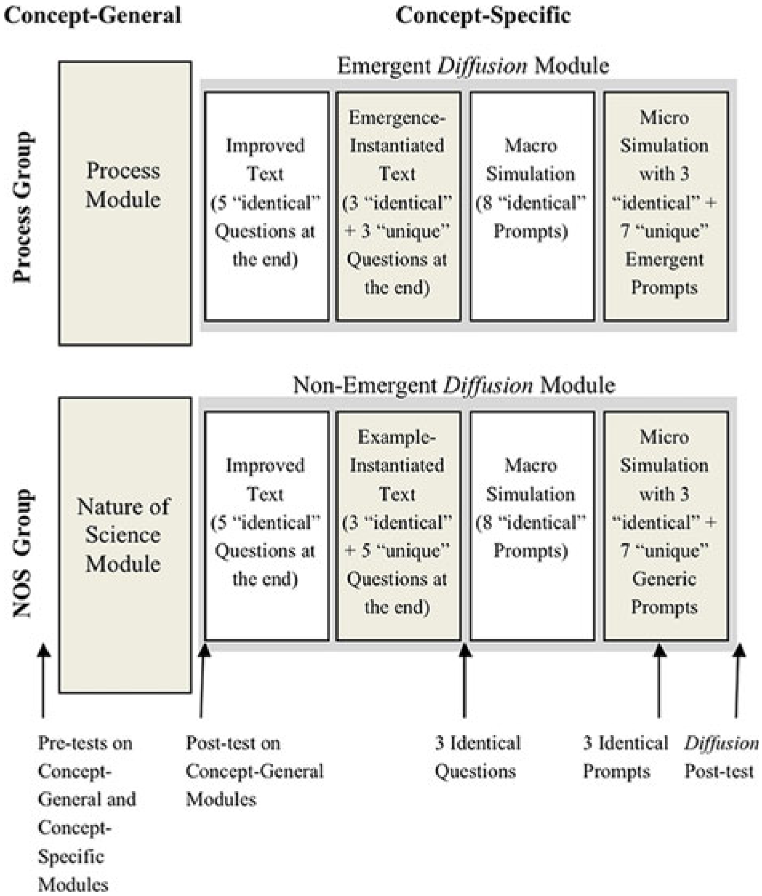


Fig. 2. A schematic of the sequencing of the instructional materials used by the Process Group (top) and the Nature of Science (NOS) Group (bottom), as well as indications by the arrows of when the assessment data were collected. Darkened boxes are instructional materials that differed substantially between the two groups.

(the NOS Group). We begin by describing the two concept-general Modules, shown as the first pair of large dark boxes in the first column in Fig. 2.

4.4.1. The Process Module

The Process Module consisted of a 6,479-word-long text passage, divided into three activity units. An activity unit is defined in the WISE platform as a coherent unit introducing a concept. The three activity units were Introduction, Definition of Processes, and Elements and Interactions. The Introduction unit began by pointing out that there are two kinds of processes that are often confused because they appear similar, just as whales look like fish even though they are actually mammals. It described the similarities and differences between the two kinds of processes, including ways to identify them and to understand what constitutes a causal explanation for the pattern-level behavior of each.

The Definition of Processes unit then described both sequential and emergent processes with two everyday examples of each. The two pairs of contrasting examples used were wolves hunting and building a skyscraper as sequential processes, and fish schooling and students crowding as emergent processes. Each set of examples included a natural event and an example involving people. We intentionally used everyday examples rather than scientific examples because they are familiar and generally easy to understand. This meant that students could focus on the underlying structure of the example concepts without being hampered by scientific terminologies and technical details, since the purpose of the Process Module was to build a qualitative concept-general understanding of the nature of the processes, and not to teach specific technical details about any specific science process. Thus, the Process Module described two kinds of everyday processes (sequential and emergent) that are totally unrelated to any science processes, in much the same way that they are described in prior sections of this article.

In the Elements (or what this article has referred to as agents) and Interactions unit, in the context of these everyday examples, we explicitly illustrated the specific second-order interaction features of the agents' interactions (Table 2) that would enable students to identify, recognize, and discriminate between sequential and emergent processes. In addition, we also illustrated the inter-level attributes (Table 1) that characterize the causal explanation of the pattern from the interactions of the agents, much in the same way as they are illustrated in this article. For the sequential building-a-skyscraper example, we described the elements as consisting of workers, with each worker behaving

> in its own way, depending on the specific job of the worker: architects design the plan for the building, foremen oversee the construction workers, welders connect steel beams… Notice that each type of worker interacts with another worker or some other element in a different way. The architect designs the plan, but the foremen oversee the construction workers…some workers' roles may be seen as more central than others. For example, you might say that the architect is the mastermind of the whole project.

Thus, in this example, both a second-order feature (e.g., that the workers have different/distinct roles, emergent feature #1, Table 2) and an inter-level attribute (e.g., that the architect may have a more controlling role, emergent attribute #2, Table 1) are introduced. An assessment post-test question relevant to this skyscraper example is: *Do certain elements of the micro level have a more important role than others?* Similarly, we described the elements of the emergent crowding example as:

> The elements in the crowd process …are the individual students who are all trying to leave the room. …All the students are behaving more-or-less the same way, which is to run toward the door at about the same speed, as each student is about the same age and size…

This description articulates that the elements/agents interact in a *uniform* way (emergent feature #1, Table 2). A post-test question relevant to this crowding example is: *Do certain elements of the micro level have a more important role than others?* These snippets show that the description of sequential and emergent processes can be done in simple and straightforward ways, so understanding our Process Module should not be overly difficult for eighth and ninth graders.

The Process Module text contained 12 illustrations and 20 text-embedded questions. These questions were designed to make students more attentive and reflective, requiring self-explanation to answer them. The questions were also meant to draw students' attention to the content without necessarily giving away more information. For example, one question was ''What are some of the elements of the process of building a skyscraper?'' Feedback was not given, but students could get feedback indirectly when they read the correct information embedded in the text, such as that ''the elements are the workers that are involved in the building's construction.'' To summarize, the design of the Process Module was guided by our theoretical account for the causes of misunderstanding an emergent kind of process, as well as challenges and other design principles from cognitive psychology.

4.4.2. The NOS filler Module

The NOS Module was taken directly from the first chapter of *Science for All Americans*, published by Project 2061 (AAAS, 1989). The chapter, titled *The Nature of Science*, was also divided into three activity units in WISE: the Scientific Worldview, the Inquiry Process, and the Scientific Enterprise. In addition, we added passages on models and the concept of scale from Chapter 11, titled *Common Themes*, in order to make the overviews for the two modules more comparable in length (approximately 6,016 words). The NOS text was interspersed with 17 illustrations and 11 text-embedded questions. We similarly inserted prompting questions for the purpose of directing students' attention toward important information.

4.4.3. Two concept-specific diffusion modules

We also created two different concept-specific *diffusion* modules, each containing four lessons. One version is referred to as the Emergent *Diffusion* Module because two of its four lessons instantiated the ideas of emergence in the context of *diffusion.* This meant that we illustrated the ways that *diffusion* is an emergent process. Such instantiation seemed necessary because the ideas introduced in the Process Module may have been quite foreign to students, as they were not related to *diffusion* or any other science processes. Therefore, we could not expect students to be able to apply the knowledge learned in the Process Module to the concept of *diffusion* without some scaffolding. This assumption is based on Gick and Holyoak's (1983) study, which showed that even when participants are able to solve one problem (the fortress problem), they cannot necessarily transfer that knowledge to solve a similar problem (the tumor problem), unless they are reminded to apply the previous solution. Likewise, here we expected students to be more likely able to apply ideas from the Process Module to learning *diffusion* if we ''instantiated'' *diffusion* as an example of an emergent process. To that end, we introduced *diffusion* in the context of the Emergent *Diffusion* Module in an ''instantiated text'' describing *diffusion* as an emergent process, and emergent attributes and features were also prompted in the context of our micro simulation.

In contrast, the concept-specific Non-emergent *Diffusion* Module did not instantiate or refer to emergence in its texts or simulations, since such scaffolding would not make sense to students studying the NOS filler Module. Instead, the non-emergent *diffusion* text was instantiated in the context of two examples. Similarly, for the micro simulation, understanding the attributes and features of the molecular interactions was scaffolded with ''generic'' (content-free) prompts, rather than ''emergent'' prompts.

Details of the four lessons for each *diffusion* Module are described below. The two *diffusion* Modules are shown by the two sets of smaller boxes in Fig. 2. Darkened boxes indicate that the instructional materials used for the Process and the NOS Groups are substantially different.

4.4.3.1. Improved Diffusion Text: A description of *diffusion* must include information about the molecular component or the agent (micro) level interactions; otherwise there is very little hope that students can ever understand how the interactions produce the pattern. In that spirit, we also improved the *diffusion* text that we used for both groups. The first section of our *diffusion* text (around 1,000 words, nine paragraphs, with some words also describing the figures) was improved by taking a composite of selected sentences from six existing middle-school grade-appropriate textbooks and edited for coherence. Selected sentences contained what we deemed to be important information, such as descriptions of atoms and molecules, how they move and collide, and how they diffuse by spreading. These were selected because they represented typical and well-written descriptions from the various middle-school texts that we included for our composite. Illustrative drawings and photographs were also included, to enhance interest and clarity. At the end of this improved text, five ''identical'' short-answer questions were posed for the students to answer (although the fifth question was phrased slightly differently).

The Improved *Diffusion* Text used for the Process Group conveys essentially the same content information as the one used for the NOS Group, except that some editing (unfortunately) resulted in minor differences in the two texts.[1] Given that the content was similar, we treated these two versions as essentially equivalent (but definitely not identical), and they are shown in Fig. 2 as the Improved Text (second column) for each of the *Diffusion* Modules.

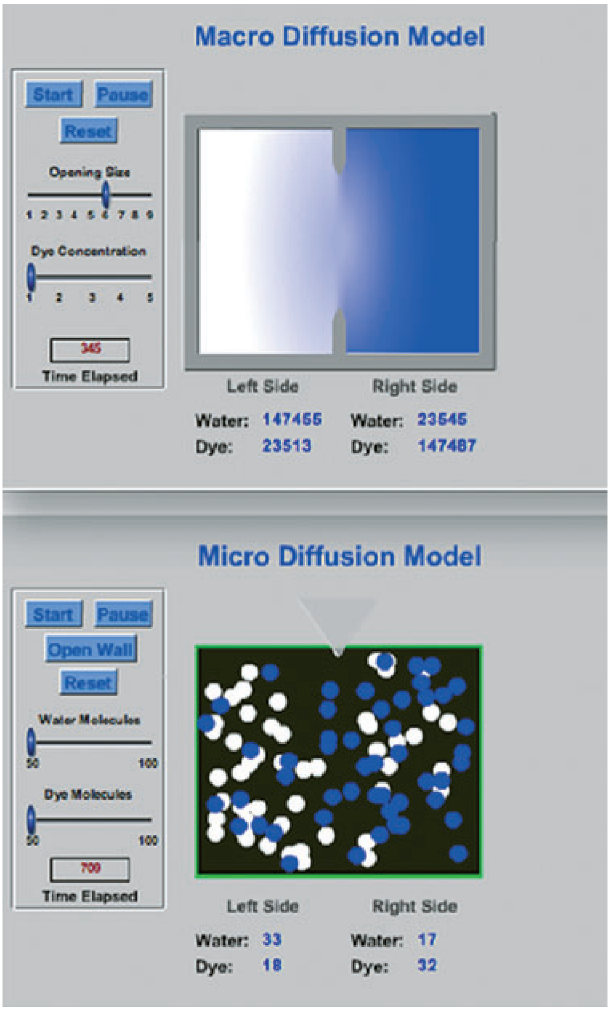
4.4.3.2. Instantiated Diffusion Texts (two versions): In addition to *improved* texts, two *instantiated diffusion* texts were created (both indicated in the third column of Fig. 2, darkened to show that they are different). A common assumption in the problem-solving literature is that the value of a general principle or theory (as entailed in the Process Module) will be useful to the extent that it is instantiated in a specific example (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Reed & Bolstad, 1991). Therefore, we created the Emergence-instantiated *Diffusion* Text that first described and explained a scenario of ink/dye diffusing in water, then instantiated *diffusion* in the context of emergence by describing ways in which *diffusion* is an emergent process. It illustrated the first four attributes in Table 1 and all the features in Table 2. For example, for emergent attribute #1 (in Table 1) and emergent feature #1 (Table 2), the text first says that, ''In an emergent process *the entire collection of elements* contributes to the pattern *in the same ways*.'' Then it illustrated this attribute and feature in the context of *diffusion* by stating that, ''In this beaker example, we see that *all of the water and dye molecules* are *doing essentially the same thing*: *bouncing around inside the beaker*, *sliding past other molecules*, and so on.'' This Emergence-instantiated *Diffusion* Text was 901 words long (containing two pictures showing ink flowing in clear water) and was followed by six short-answer questions that students had to answer.

The *Diffusion* text for the Non-Emergent *Diffusion* Module instantiated *diffusion* in the context of two example scenarios (to be referred to as the Example-instantiated *Diffusion* Text). It described diffusion of dye in a liquid solution in a scenario that was used in the assessment test (see Scenario 1 in Appendix A) and then described diffusion of oxygen and carbon dioxide between arteries and cells (another scenario also used in the assessment test). This *diffusion* text was 735 words long, containing two pictures depicting the two scenarios. Following this section were eight questions.

Among the six questions following the Emergence-instantiated Text and the eight questions following the Example-instantiated Text, three of the questions were identical for both versions. (These three identical questions, along with the non-identical questions, were administered after students completed the instantiated texts, as indicated by the third arrow in Fig. 2. The three identical questions are shown in *Material S1* as Questions 1, 2, and 3.)

4.4.3.3. Two simulations of diffusion: Diffusion was exemplified by two simulations: one for the macro level (showing the pattern and outcomes) and one for the micro level (showing the agents/molecules and their interactions). Both the macro- and micro-level simulations depicted a rectangular container separated into two halves by a wall. The right half contained ink (the macro version) or ink molecules (the micro version), and the left half contained water or water molecules. When a simulation was run, the wall opened and diffusion occurred. For each simulation, a separate window provided instructions and questions to guide students' explorations. The instructions asked students to manipulate one variable at a time, four separate times, and question prompts then required them to describe and explain a specific aspect of the outcome for each manipulation.

The macro simulation depicted liquids: ink and water (the aggregated components). In other words, we used Flash to visually portray the diffusion of ink in water in much the same way as one would see it physically. A screen shot of this macro-level simulation is



Left Side		Right Side	
Water:	147455	Water:	23545
Dye:	23513	Dye:	147487

Left Side		Right Side	
Water:	33	Water:	17
Dye:	18	Dye:	32

Fig. 3. The Macro (top) and Micro (bottom) simulation of diffusion.

shown in the upper half of Fig. 3. NetLogo usually depicts the macro-level pattern by abstractions, such as graphs, or in some cases as in flocking, the macro pattern visibly appears from the behaviors and interactions of the individual agents, whereas our macro simulation is visually isomorphic to the real physical pattern and is separate and distinct from the micro simulation. Students' exploration of the macro simulation was guided by four request steps: First they were asked to run the simulation without changing any variables, then to describe and explain the direction of the ink flow. Next, they were asked to increase the size of the opening, and then to describe and explain the pattern after the system reached equilibrium. For step three, they further increased the size of the opening, then described and explained the effect of doing so on the time required for the system to reach equilibrium. For the final step, students increased the ink concentration and, again, focused on the time required to reach equilibrium. While interacting with the simulation, they were prompted with eight questions. These prompts basically asked students to describe what they saw in order to insure that they were paying attention (e.g., ‘‘What direction does the blue liquid appear to flow?’’), as well as predict and describe macro-level behavior qualitatively, as a function of variability in the conditions (for instance, for widening the opening between the two sides they were asked, ‘‘When the opening is widened, does the system reach equilibrium more quickly or more slowly than before?’’). Both Groups of students saw the same macro simulation with eight identical prompts (see Fig. 2, Column 4).

The micro simulation showed the same scenario of water and ink in two separate sides of a container, except ink and water molecules were depicted, in ways that are similar to typical NetLogo simulations (a screen shot of it is shown in lower portion of Fig. 3). For the micro simulation, they were guided with only two exploration steps: (a) run the simulation without opening the wall, and (b) run it again with the wall open. While interacting with the micro simulation, they were prompted with 10 questions, five for each step.

The micro simulation itself was identical for the Emergent as well as the Non-Emergent *Diffusion* Modules. The only difference between the two Modules was with regard to the prompts, in that only three of the 10 prompts were identical for the two Modules. In the Emergent version, seven of the 10 prompts in the micro simulation were ‘‘emergent’’ in that the questions scaffolded students to focus on instantiating specific features and attributes of the Emergent Schema. For example, to instantiate the *random interactions* feature (emergent feature #2, Table 2), students were asked, ‘‘Can the water and dye molecules *interact with any other* water and dye molecules?’’ and to instantiate the *uniform* feature (emergent feature #1, Table 2), students were prompted with, ‘‘Do the water and ink molecules *behave and interact in more or less the same way?*’’

In contrast, seven of the 10 prompts in the Non-emergent version can be considered ‘‘generic prompts,’’ in that they were basically content-free open-ended questions that asked students to describe important aspects of the micro simulation to make sure that they are paying attention, but do not refer to specific emergent features or attributes. For example, students were asked to ‘‘Describe the movements of the water molecules. What are they doing?’’ and ‘‘Do the ink molecules go in any particular direction?’’ So the micro simulations used for the Emergent and Non-Emergent *Diffusion* Modules each had seven unique prompts, and they shared three identical prompts (as shown in *Material S1* as Questions 4, 5, and 6).

Based on our theoretical analyses, we predict that pattern-level information contained in the macro simulation was already intuitively familiar to students, since it resembles what students normally encounter in their environment (e.g., seeing cream diffusing in coffee), and it represents a kind of *sequential* process, as with initial conditions causing changes in the pattern (e.g., initial high concentration causing the appearance of flow to go faster). Moreover, as stated earlier, students often can give correct single-pattern-level explanations based on the conditions and parameters within the pattern level. Therefore, we would expect that the macro simulation would *similarly* benefit both the Process Group and the NOS Group. However, the micro simulation could *differentially* benefit each Group, since the Process Module might influence what students specifically learn from the micro simulation, especially since the micro simulation for the Process Module Group used emergent prompts (versus generic prompts). That is, we assumed that the Process Group were able to benefit and learn from the emergent prompts embedded in the micro simulation since they had learned the Process Module, which differentiated emergent from sequential processes.

4.4.4. Assessment

The arrows in Fig. 2 show when assessments were taken. In this section, we describe the content of each assessment. Understanding of the Process Module was assessed by multiple-choice questions (six on the pre-test and 20 on the post-test) and eight open-ended questions on the post-test. Similarly, understanding of the NOS was assessed by multiple-choice questions (six on the pre-test and 15 on the post-test) and five open-ended questions on the post-test. We routinely added extra questions for the post-tests in order to assess learning beyond the benefit of the testing and retesting effect (i.e., taking a test on studied material usually promotes subsequent learning and retention of that material on a final identical test; McDaniel, Anderson, Derbish, & Morrisette, 2007). Pre-test and post-test were administered before and after the concept-general Modules, as shown by the first and second arrows in Fig. 2.

Understanding of *diffusion* was assessed by 10 multiple-choice and two open-ended questions on the pre-test (administered prior to the concept-general Module, as shown by the first arrow in Fig. 2), and 15 multiple-choice and five open-ended questions on the post-test (with nine multiple-choice questions repeated from pre- to post-test). The post-test was administered at the end of the *Diffusion* Modules (indicated by the fifth arrow in Fig. 2). In the multiple-choice *diffusion* post-test, eight asked about either the micro or inter-level relationships that were related to the features and attributes listed in Table 1 and 2 (these will be referred to as emergent questions), and seven asked about the macro level. Thus, more than half of the post-test *diffusion* questions were about the micro-level and inter-level knowledge that may have been new to students. (The nine identical pre- and post-test questions are shown in Appendix A and are identified as either micro, macro, or inter level. Note that several questions often pertained to the same scenario.)

Although the same identical *diffusion* post-test questions were used with both Groups of students, the post-test questions could have been more difficult for the Process Group than for the NOS Group. This is because the post-test included two scenarios that were identical to the scenarios of the two examples in the Example-instantiated Text for the NOS Group. Thus, the post-test questions favored the students in the NOS Group because of their familiarity with these two scenarios, and so the post-test questions were stacked against the Process Group.

4.4.5. WISE delivery platform

The entire content of the intervention materials consisting of the texts, the simulations, assessment questions, embedded prompting questions, and pre- and post-tests were all delivered using the WISE platform on laptop computers. WISE provides tools and support for creating interactive, online student learning activities (Linn & Hsi, 2000). We chose online delivery for three specific reasons. First, in a pilot study undertaken in the prior year in which we tested the materials of the Process Module, the Emergent *Diffusion* Module, and an Emergent *Natural Selection* Module, we had the experimenter mimic an instructor by discussing the Modules with the students in small groups (Chi et al., in press). Based on our observation of these small group discussions, we realized that we could not make the discussions consistent across the separate groups, even though we had provided a script for the instructor to use. Second, we wanted the materials to be deliverable online because of the potential for scale-up in the future. Third, the use of simulations dictated some kind of computer use anyway. In the pilot study with small groups, the instructor projected the simulation on a screen. In addition, we chose WISE as our platform because WISE is not only easy to use and implement, but it is particularly suited for young students.

4.5. Participants and procedure

The study was carried out after the end of the year in a prestigious private school, because very few public schools at the time had individual laptops available to all students. The students in this private school carried out all of their homework assignments on laptops, so they were proficient at typing by eighth grade. Obviously, interpretation of the results has to be limited to this special population.

Twenty-nine eighth graders and 13 ninth graders volunteered to participate. One student served as a pilot so the study included a total of 41 students. Both the eighth and ninth graders had covered the topic of *diffusion* in their science class with the same science teacher, using materials from the McDougal Littell Life Science text (Trefil, Calvo, & Ms. Cutler, 2006), which was also supplemented by other materials. The study was supervised by the students' science teacher.

For the study sessions, all students came in at the same time everyday, and they read the materials and typed responses on their laptops at their own pace. In general, each session took around one and a half hour. Throughout the study, they were permitted to ask clarification questions or for help with the laptop. We refer to this as a quasi-classroom setting because it was similar to their regular science class (in the sense of knowing their peers, the environment, and the teacher), but it was different in that it was held during the week following the last day of classes for the year, no grades were given, and the students were paid to participate.

As pointed out earlier, the basic design of the study reflects the sequencing of the intervention materials shown in Fig. 2, with one group of students using the materials shown in the upper half of Fig. 2 (the Process Group) and the second group using the materials in the lower half of Fig. 2 (the NOS Group). We used a modified stratified random assignment procedure, based on the scores of a standardized test administered in the school (Educational Records Bureau Comprehensive Testing Program which assesses reading, math, and writing). After a median split on their standardized test scores, students within each block (higher vs. lower scores) were randomly assigned to either the Process or the NOS conditions. This resulted in 21 students in the Process Group and 20 in the NOS Group. All 41 students took the pre-tests and post-tests of the concept-general Modules, but one student's data from the NOS Group was lost for the *diffusion* sessions. Therefore, we had a total of 21 in the Process Group and 19 in the NOS Group.

The study took three sessions, administered across 3 days. The first session was the pre-test (during which all students completed four pre-tests) on materials in both of the concept-general Modules, as well as materials on two specific concepts: *diffusion* and *heat transfer*. A pre-test on *heat transfer* was included because we had anticipated assessing the students' learning of *heat transfer* on the fourth day, but because we did not have sufficient time to prepare the intervention materials properly, the learning results on *heat transfer* will not be reported here. We present the *heat transfer* pre-test results anyway, as they inform us about the student samples. The second session consisted of learning the concept-general Module materials and answering the prompting questions embedded in the concept-general Modules, followed by taking the post-test on the concept-general Module materials. Although only one session was allotted for the concept-general Module for each condition, ideally, information in the Module should be covered in perhaps four or five sessions. The third session consisted of learning the concept-specific *Diffusion* Module.

4.6. Results

4.6.1. Scores of four pre-tests

Although the participants were assigned using stratified random sampling, we assessed whether or not the Process Group and the NOS Group were, in fact, comparable. We compared their scores on all four pre-tests (on the Processes and NOS Modules, and on *diffusion* and *heat transfer*). (The first arrow in Fig. 2 indicates when the pre-tests were administered.) The mean percentage correct on the four pre-test scores for the Process and NOS Groups are shown in Fig. 4. There were no statistical differences between the two Groups on any of the pre-tests.

4.6.2. Learning the concept-general modules

After students in both Groups read the materials in their respective concept-general Modules and answered the embedded questions, they took the post-tests assessing their

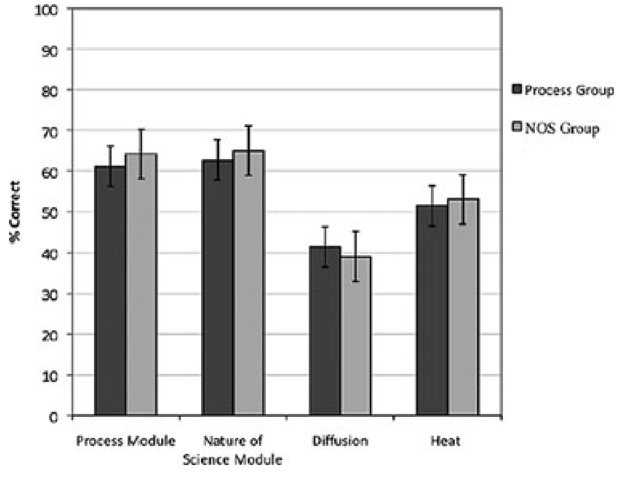


Fig. 4. Percent correct multiple-choice responses to all four pre-test questions for the Process Group and Nature of Science (NOS) Group.

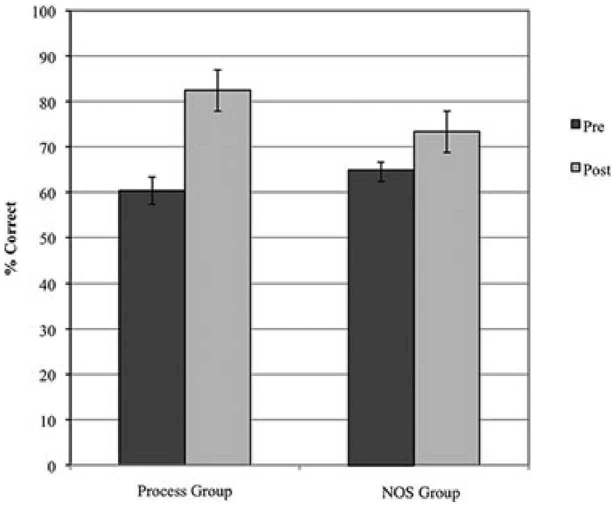


Fig. 5. Percent correct multiple-choice responses to all pre- and post-test questions on respective concept-general Modules for the Process Group and Nature of Science (NOS) Group.

knowledge of their concept-general Modules (as indicated by the second arrow in Fig. 2). Using only the objective measures, the six multiple-choice questions of the pre-tests for both Modules and the 20 (for Process) and 15 (for NOS) multiple-choice questions for the post-tests were scored and used as a manipulation check to make sure students did, in fact, learn the concept-general Modules. For the Process Module, of the six matched or identical questions on both the pre-test and the post-test, the scores were 61.11% on the pre-test and 71.91% on the post-test ($F(1,20) = 4.092$, $p = .05$), thus a significant gain of 10.80%. Using the entire post-test scores of all 20 questions, the proportion correct rises to 84.43%, indicating a significant difference ($F_{(1,20)} = 19.321$, $p < .001$, $d = 1.233$) based on a one-way repeated measure ANOVA (shown in Fig. 5). This result indicates that eighth- and ninth-grade students can somewhat understand that there are two kinds of processes, and that one kind is emergent.

Similarly, the means for the six matched pre-test and post-test scores of the multiple-choice questions on the NOS Module were 65.79% and 71.67% correct, indicating only a marginally significant gain ($F(1, 19) = 3.234$, $p = .09$). However, using the entire sample of 15 post-test questions, the learning difference as assessed by the pre-test and the post-test (73.33%) was again more substantial and significant ($F(1, 19) = 4.929$, $p < .039$, $d = 0.748$), using a one-way repeated measure ANOVA (see Fig. 5, second set of columns).

The results reported above for both the Process Module and the NOS Module show that students in both Groups did pay attention to the materials and demonstrated learning, although learning in the Process Group was more substantial (effect size $d = 1.233$) than learning in the NOS Group (effect size $d = 0.748$). However, it does not make sense to compare the *amount* of learning between the two Groups per se, because the two Modules differed in many ways. They differed in content, the number of prompts (20 for Process and 11 for NOS Groups), the number of illustrations (12 for Process and 17 for NOS Groups), the exact length (6,479 words for Process and 6,016 for the NOS Groups), the readability of the materials, their intrinsic appeal, whether examples were embedded, the writing style of the two passages, and so forth. Despite such differences, the intent was to provide a concept-general filler module that was relevant to grade-appropriate science concepts as indicated by their use in textbooks. The pre- and post-test results served as a manipulation check showing that students in both Groups were attentive and learning.

4.6.3. Analyses of objective measures on learning diffusion

Three objective measures are reported in this section: multiple-choice pre- and post-test questions, misconception-reductions, and correlations.

4.6.3.1. Pre-post-test learning of diffusion: Students took the *diffusion* post-test after having worked on the micro *diffusion* simulation (as shown by the last arrow in Fig. 2). Fig. 6 shows the mean proportion of correct responses for all 10 pre-test and 15 post-test multiple-choice questions on *diffusion* for the Process and the NOS Groups. Both Groups learned significantly (33.36% for the Process Group, $F_{(1,20)} = 57.270$, $p < .0005$, $d = 1.905$, and 17.22% for the NOS Group, $F_{(1,18)} = 22.290$, $p < .0005$, $d = 0.484$), based on repeated measures ANOVAS. Moreover, the Process Group learned more than the NOS Group, as demonstrated by the analysis of covariance (ANCOVA) performed with all 15 multiple-choice post-test questions as the dependent measure and 10 pre-test questions as the covariate. This stricter analysis showed that the amount learned by the Process Group significantly exceeded the amount learned by the NOS Group ($F_{(1,38)} = 9.391$, $p = .004$) by almost one standard deviation, with an effect size of $d = 0.967$.

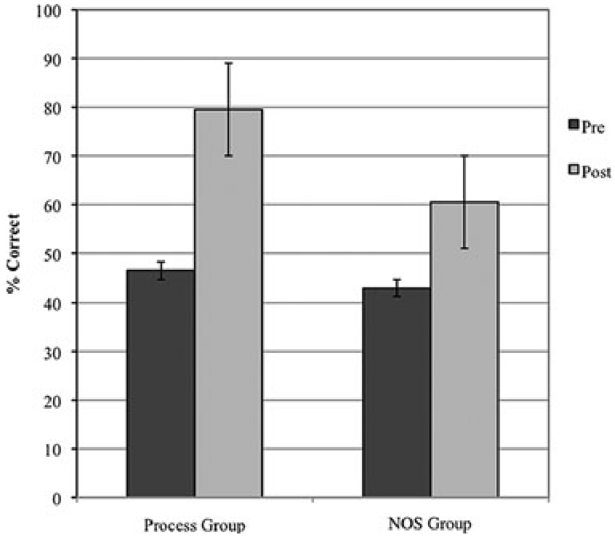


Fig. 6. Percent correct multiple-choice responses to all pre- and post-test diffusion questions for the Process Group and the Nature of Science (NOS) Group.

4.6.3.2. Misconception reduction: Six of the multiple-choice questions (Questions #4, #5, #6, #7, #8, #9 listed in Appendix A) that were repeated on the pre- and post-test tapped common emergent-relevant misconceptions. Each of these six questions provided four option choices, with one option being the correct answer, one or more options being a commonly misconceived answer, and one or two foils. Each question was scored as correct or incorrect to obtain measures of overall performance. Across these six questions, nine options targeted common misconceptions (e.g., ''Both the ink and water molecules have stopped moving because the solution is at equilibrium and so the process of diffusion is done.''), and the remaining nine were simply incorrect answers (e.g., ''The water has been transformed by the ink.''). Because there were an equivalent number of misconceived and incorrect options, the probability of choosing either one was equal so we could simply analyze the number of misconception options chosen on pre- and post-test. The Process Group chose, on average, 3.19 misconceived options on the pre-test and the NOS Group chose, on average, 3.75. On the post-test, the Process Group chose a misconceived option 1.10 times, whereas the NOS Group chose a misconceived option 2.05 times. Using an ANCOVA with the number of post-misconception answers chosen across the six questions as the dependent measure and the number of pre-misconception answers chosen as the covariate, there was a significant difference between the Process and the NOS Groups in the number of misconceived answers chosen (adjusted mean for the Process Group is 1.109 versus a mean of 2.036 for the NOS Group, $F_{(1,38)} = 7.200$, $p = .011$, $d = 0.8886$). This suggests that the Process Group reduced the number of misconceived answers to a greater degree from pre-test to post-test.

4.6.3.3. Correlations of the concept-general Modules with diffusion: The preceding analyses of learning, as assessed by the multiple-choice pre- and post-test questions on *diffusion*, showed that the Process Group excelled in their understanding of *diffusion*, compared to the NOS Group. The obvious question is: Did having learned the Process Module help students learn *diffusion* more easily? We can assess the contribution of the concept-general Modules by seeing whether there is a correlation between how well students in both Groups performed on their respective concept-general Module post-test and the *diffusion* post-test. Accordingly, their performances on the multiple-choice post-test concept-general Module scores were correlated with their multiple-choice post-test *diffusion* scores. The correlation was positive and overwhelmingly significant for the Process Group ($r = .782$, $p < .0005$), but not for the NOS Group ($r = -.218$, $p = .356$) (see the first row of Table 3).

Even though both the Process and NOS Modules were concept general, the highly positive and significant correlation between learning from the Process Module with learning from the concept-specific *diffusion* Module is consistent with the hypothesis that students could have built an Emergent Schema from the Process Module and then relied on this Emergent Schema to help them interpret information about emergent processes such as *diffusion*. In contrast, the lack of any significant correlation between the post-test scores for the NOS Module with the *diffusion* Module suggests that students' knowledge of the NOS did not have any bearing on their understanding of *diffusion*. This latter finding is not surprising given that the NOS Module is not directly related to emergent processes. Because the Process Module produced a significant, positive correlation between ''emergent-like''

Table 3
Correlations of post-test multiple-choice scores for concept-general modules with post-test multiple-choice scores for *diffusion*

	Process Group	NOS Group
Correlation of concept-general Module post-test scores with *diffusion* multiple-choice post-test scores	$r = .782$***	$r = -.218$
Correlation of scores of correct responses to identical Macro *diffusion* prompts with *diffusion* multiple-choice post-test scores	$r = .162$	$r = .086$
Correlation of scores of correct responses to identical Micro *diffusion* prompts with *diffusion* multiple-choice post-test scores	$r = .551$**	$r = .355$

Note. NOS, Nature of Science.
*Significant at $p = .05$; **significant at $p = .01$; ***significant at $p = .001$.

knowledge and *diffusion* knowledge and the NOS Module did not produce any significant correlation between ''non-emergent-like'' knowledge and *diffusion* knowledge, and given that both Groups of students performed similarly on all four pre-tests, we can assume that the correlation found for the Process Group was not due to other factors such as general ability and that the Process Module did, in fact, help students to better understand *diffusion*.

4.6.3.4. Summary of analyses of objective measures: If we consider the pre-test as a proxy for assessing students' regular classroom ''business as usual'' instruction (since they have been taught the concept of *diffusion* in class prior to this study), then the results shown in Fig. 6 indicate that both Groups of students only achieved at around 40% correct at the pre-test (with no significant difference between them, as shown in Fig. 4). The fact that both Groups learned significantly more from our instructional materials is not surprising since we included improved and instantiated *diffusion* texts, simulations that depicted both the macro and micro-level behavior, and text-embedded question prompts throughout the intervention that required student responses. Thus, overall, both Groups improved significantly over their classroom learning because the materials in the *diffusion* Modules were enhanced in many ways, and not necessarily only because we provided the concept-general Modules. However, the Process Group learned about *diffusion* significantly more than the NOS Group, as assessed by the pre- and post-tests and by the reduction in choosing misconceived answer choices. Finally, it appears that the Process Module may have mediated learning of *diffusion*, since the better the students learned the Process Module, the better they then learned *diffusion.* Such a strong, significant correlation was not obtained with the NOS Module. Moreover, as mentioned earlier, the post-test *diffusion* questions were stacked against the Process Group since two of the scenarios used in the post-test were already described and used as examples in the Example-instantiated Text for the NOS Group. Despite that fact that the post-test questions favored the NOS Group, the differences in learning between the two Groups provides further evidence for the advantages of the Process Module.

4.6.4. *Analyses of open-ended responses*

Open-ended questions and prompts occurred throughout the instructional materials. Although there were few open-ended questions and prompts, we coded them with an eye toward revealing findings that are important for our theoretical assumptions. Below, we describe which open-ended questions were coded and how they were coded.

4.6.4.1. Expressing emergent and non-emergent ideas in the post-test: Five of the post-test *diffusion* questions were open-ended. Four questions referred to specific scenarios used for the multiple-choice questions (one will be described below) and one was scenario independent. (*Dogs are well-known to have a tremendous sense of smell. They can find objects and people by their scent. Explain how a dog could find you by your scent. Hint*: *you should refer to the molecular behavior*, *if possible.*) They were coded in the following way. Of the students' responses, ideas that were relevant to *diffusion* were identified and segmented into idea units. Then each idea unit was coded as describing either emergent or sequential features and attributes (as shown in Tables 1 and 2). For example, Scenario 2 (see Appendix A) describes a beaker of dye and a beaker of water connected by a clamped tube. When the clamp is removed, the dye appears to flow into the other beaker. Students were asked in one of the open-ended questions to explain what causes the apparent flow. One student wrote,

> The molecules in both the liquids are constantly moving and colliding with each other.// When the clamp is released, the different molecules collide and pass by each other,// spreading out into different beakers.//

That response was coded as three relevant segments (separated by//), with the first segment coded as describing an emergent feature *uniform*, because it refers to the molecules of both liquids as following the same rules of ''moving and colliding with each other''; the second segment, when interpreted in conjunction with the first segment, was coded as *continuous*, because it indicates that the molecules move and collide before, during, and after the clamp is released; and the third segment was coded as a correct idea, but one that is not necessarily relevant to any emergent features and attributes. Another student answered the same question as follows: ''because the dye is given more space and wants to diffuse farther.'' That response was coded as *deliberate/intentional*, a sequential inter-level attribute.

Many idea units (like the third idea in the first student's response above) are not about the features and attributes as described in Tables 1 and 2, but they can still be coded as correct or incorrect. They could be correct ideas about molecular movement, such as:

> the motion of the molecules is random; or
> molecules mix together; or
> movement of one molecule from one beaker to the other occurs by chance.

Correct ideas can also be about macro-level behavior, such as ''the ink spreads out.'' Because the majority of the ideas could be coded as either characterizing an emergent or a sequential attribute and feature, few relevant segments were coded as incorrect per se. On average, the Process Group and the NOS Group generated 1.06 and 1.45 incorrect ideas, respectively, and these scores were not significantly different. Because of this low frequency of incorrect ideas, the analyses focused only on the correct ideas.

In short, each correct idea unit was assigned whatever features and attributes that were appropriate, similar to the way it was done in Ferrari and Chi (1998). Two researchers independently coded each idea unit in the segmented protocols across all 40 students' responses to the five open-ended *diffusion* questions on the post-test, and the inter-rater reliability for this coding was in virtually perfect agreement (Kappa coefficient = 0.99, $p < .0005$). The remaining open-ended responses to be presented in this article were coded by one researcher.

Those segments that were coded with emergent features/or attributes were combined together as ''emergent.'' For instance, for the first student's response in the above example, because the first two segments were coded as *uniform* and *continuous* (emergent features #1 and #5 in Table 2), they were categorized as emergent ideas. Segments that described either sequential features and/or attributes or were simply correct statements that described either the outcome and/or the pattern (such as the third segment of the example above that molecules were ''spreading out into different beakers'') were categorized as non-emergent. Other ideas pertaining to technical knowledge were also coded as non-emergent.

Fig. 7 shows only the correct ideas expressed in the five open-ended questions on the post-test and three basic findings should be noted. First, on average, there was no difference in the total number of correct idea units in the responses of the Process Group (9.83 idea units) and the NOS Group (10.08 idea units), suggesting that both Groups of students were equally articulate and verbose. That is, both Groups of students were actively trying to answer the questions. Second, on average, students in the Process Group generated a greater number of emergent idea units than students in the NOS Group (1.33 versus 0.60 emergent idea units), and also generated fewer non-emergent ideas than the NOS Group (8.50 versus

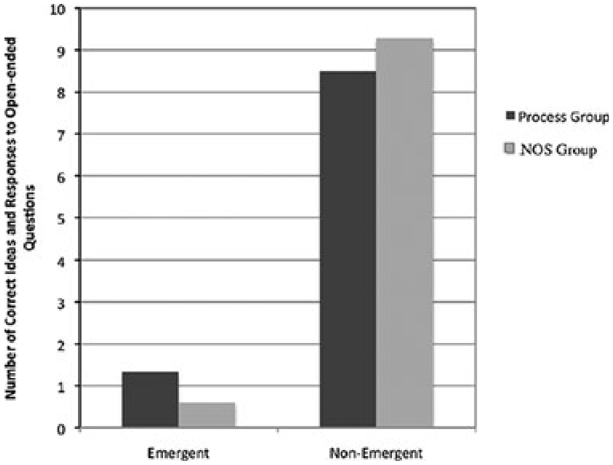


Fig. 7. Total number of correct emergent and non-emergent idea units expressed in the five open-ended post-test questions for the Process Group and the Nature of Science (NOS) Group.

9.28 non-emergent idea units). Although the pattern of results is compatible with the greater improvements of the Process Group in understanding *diffusion*, these differences are not statistically significant, in part because only three of the five open-ended questions required inter-level explanations.

Third, the most important finding we highlight in the results of Fig. 7 is the large discrepancy between the number of correct non-emergent ideas generated versus the number of correct emergent ideas. The roughly 9:1 difference in the amount of non-emergent versus emergent ideas generated reinforces an important assumption that we made earlier; namely that it is far easier for students to express (and presumably learn) non-emergent ideas and other technical knowledge than it is to express emergent ideas. Since these are all correct ideas expressed at the post-test, the large difference suggests that it is relatively easy for interventions to show success in learning (as many other studies have done), but often these successes cannot be attributed to an improvement in generating correct inter-level causal explanations (i.e., the successes cannot be attributed to correcting misconceptions). Thus, the data in Fig. 7 confirm the findings from several decades of research on the robustness and persistence of misconceptions, and they highlight the difficulty of understanding inter-level causal mechanisms of emergent processes.

4.6.4.2. Diffusion simulations: The question addressed here is, ‘‘Did the simulations contribute differentially toward learning of *diffusion* for the two Groups?’’ We will discuss the macro and micro simulations separately, since our theory predicts different benefits of each.

Our theory does not predict any differences between the two Groups in learning about the macro or pattern level since students have intuitive understanding of how patterns behave (since they often involve qualitative behavior such as increasing or decreasing) and can often give correct pattern-level explanations by citing the static conditions and parameters at the pattern level. Therefore, we would expect that the macro pattern-level simulation would be *similarly* beneficial to both the Process and the NOS Groups.

The macro-level simulation presented to the two Groups was identical, with an equivalent number (eight) of identical prompts embedded in the simulation. The prompts were questions such as, ‘‘When the opening is widened, does the system reach equilibrium more quickly or more slowly than before?’’ That is, these prompts asked students to predict and describe macro-level behavior as a function of variability in the conditions, as in this case, the condition of widening the opening between the two sides. We scored the students’ responses to the eight prompted questions in nearly the same way that we scored the open-ended responses in the post-test, as described in the preceding section. One slight modification was that all segments were coded more simply as correct or incorrect macro-level, micro-level, or inter-level ideas with respect to *diffusion* since we are interested to know how well students have learned about *diffusion* in the context of the macro-level simulation. (In contrast, in the preceding coding of the open-ended post-test responses, as shown in Fig. 7, incorrect explanations that specifically pertained to attributes/features of sequential processes were coded as non-emergent, since *diffusion* is an emergent process.) For the responses to these prompts, students’ non-emergent answers would be coded as incorrect. The total scores of all the correct ideas were then correlated with students’ scores on the *diffusion* post-test multiple-choice questions. Not surprisingly, as we predicted, the correlations were not significant for either Group and were of small magnitudes ($r = .162$, $p = .484$ for the Process Group, and $r = .0861$, $p = .717$ for the NOS Group, see Row 2, Table 3).

Based on our analyses of middle-school texts, we know that students are far less familiar with the agent-level behaviors and notions of collective summing. Thus, the micro-level simulation could be helpful to students' learning, but was it *differentially* beneficial to the two Groups and did the concept-general Modules influence what students could learn from the micro simulation? The micro-level simulation presented to the Process and the NOS Groups was also identical except for one difference: Both Groups were prompted to respond to 10 questions while viewing the micro simulation (as indicated by the fourth arrow in Fig. 2), but only three of these prompts were identical across the two Groups. We coded the responses to these three identical prompts (shown as questions #4, #5, and #6 in *Material S1*) in the same way that the responses in the macro simulations were coded. We correlated these scores with the *diffusion* post-test scores. The correlations showed that the correct responses made by the Process Group correlated significantly with their post-test performance on *diffusion* ($r = .551$, $p = .010$), whereas the correct responses articulated by the NOS Group did not correlate significantly with their *diffusion* post-test scores ($r = .355$, $p = .125$). (These correlations are shown in Row 3 of Table 3.)

To summarize the effect of our simulations, the lack of significant correlations between the ideas contained in the responses generated during interactions with a macro simulation for both the Process and the NOS Groups confirms our expectation that the macro-level simulation was similarly beneficial to both Groups, perhaps because students in both Groups were already intuitively familiar with the sequential process of ink flowing at the pattern level and can more easily understand the qualitative aspects of the pattern-level information (e.g., predicting whether the flow will increase or decrease, as the opening widens, etc.). Therefore, the macro-level simulation may not have any added benefits for understanding the inter-level causal explanations of the process of *diffusion*, and having been exposed to either the Process Module or the NOS Module had no differential effect. This finding suggests that understanding an emergent process is not in the correct understanding of the pattern-level behavior. In short, the macro simulation itself, being identical for the two Groups, could not have contributed to the Process Group's superior learning of *diffusion*.

However, what students expressed in response to the identical prompts in the micro simulation did seem to correlate with learning for the Process Group, but not for the NOS Group. Although we would like to infer that this pattern of correlations suggests that the Process Module mediated learning in the context of the micro simulation, we cannot rule out the possibility that the other seven emergent prompts (as opposed to the seven generic prompts of the NOS Group) that were embedded in the micro simulation for the Process Group might also have helped students learn (or that the Emergence-instantiated *diffusion* Text may have aided in learning, which was also unique to the Process Group).

In *Material S1*, we present another coding of the three ''identical'' prompts embedded in the micro simulation along with the three ''identical'' questions following the Instantiated *diffusion* texts to get greater power. In brief, the results of this combined coding showed that the Process Group expressed significantly more correct micro-level and inter-level ideas than the NOS Group. This second coding essentially confirms our prior (but, statistically insignificant) finding that the Process Group seemed to have learned more correct emergent ideas about *diffusion* than the NOS Group (as shown in Fig. 7).

To summarize, the intervention study described here attempted to teach students about emergent processes, in order to help them build an emergent schema. We tested our intervention idea by designing a concept-general Process Module that discussed and differentiated two kinds of processes, along with their features and attributes, in order to help students build a domain-general Emergent-causal Schema. To test the effectiveness of such a structure in understanding the concept of *diffusion*, we also designed a concept-specific *Diffusion* Module, using both text materials and dynamic computer simulations showing both the macro and micro levels of the process of *diffusion*. Given that the ideas embedded in the Process Module are likely very foreign to students, we assumed that students would not see the relationship or the relevance of the Process Module to the concept of *diffusion*, unless we made that relationship explicit. Thus, we created a *diffusion* text that instantiated the Process Module, and created a presentation of the micro simulation that was guided by ''emergent'' prompts. As a result, students' learning and understanding of *diffusion* improved substantially and was positively correlated with how well they learned the Process Module.

Learning about the concept *diffusion*, although significant, was smaller in magnitude for an alternative group of students who were exposed to (a) a filler concept-general NOS Module and (b) a Non-Emergent *Diffusion* Module that included an improved *diffusion* text, supplemented by an *instantiated diffusion* text that described two additional examples of *diffusion*, and who (c) had the same macro and micro simulations as the Process Group, but with generic attention prompts, rather than emergent prompts. Although students in the NOS Group also learned *diffusion* far beyond what they had learned in their regular science class, their learning was approximately one standard deviation below the students in the Process Group. Their improvements on the *diffusion* post-test also had no relationship to how well they understood the NOS Module.

On the whole, our intervention attempt is promising in allowing students to instantiate the knowledge that they have acquired about emergent processes to learn about and understand *diffusion*. However, we are not prepared to claim that the evidence is sufficient for showing that knowledge about emergent processes (as learned from the Process Module) enhanced their learning of *diffusion*, since their improved learning could have arisen in part from the Emergence-instantiated *Diffusion* Text and the emergence-prompted micro simulation. However, it is not immediately obvious how we can avoid instantiating *diffusion* in the context of the Process Module. However, many other studies need to be done before we can reach any definitive conclusions.

5. Discussion

This article described our work on misconceptions, which has evolved over the last two decades. The first section of this article proposed a theoretical explanation of *what* misconceptions are and *why* they exist for concepts of processes. The second section considered *how* we might overcome misconceptions through instructional intervention. The third section illustrated one attempt at designing and testing an intervention approach. Providing an account of misconceptions required that we make many assertions and assumptions, some of which are novel, some are implicit, and others are controversial and contradict some prevailing views. In this final section, we revisit and make explicit our contributions and assumptions (explicated below in no particular order) so that they may be tested in future work, in order to either confirm or refute them.

In this article, we introduced the idea that *all* processes, both sequential and emergent ones, can be conceived of in at least two levels, the pattern level and the agent level (along with six other properties that they all share). In contrast, the literature tends to describe only emergent processes as having two levels (Schelling, 1978; Wilensky & Resnick, 1999). Therefore, we suggest that having two levels per se is not the relevant property that differentiates sequential from emergent processes.

In addition, at the pattern level, if the agents are visible, they can often be clustered into subgroups on the basis of perceptual similarities, such as a baseball team with the red uniform. Grouping according to perceptual similarity is a classic finding in cognitive psychology (Goldstone, 1994). When the agents are not visible, they can still form coherent subgroups, and this can sometimes be referred to as an aggregate component, such as the ink in water or the cream in coffee. In essence, referring to either the ink or the ink molecules is treating it as an aggregated subgroup, and may be similar to Levy and Wilensky's (2008) notion of mid-level groups. However, contrary to Levy and Wilensky, we assumed that referring to any subgroup(s) or aggregated component(s) (rather than considering the entire collection) as a cause of the macro-level pattern in an emergent process is flawed, revealing misconception.

Despite the similarity across all processes in having these two distinct levels, we proposed that there are two different kinds of processes, sequential and emergent ones; and that many processes embedded in middle-school science curricula are of the emergent kind. This assumption was also held by others (Levy & Wilensky, 2009; Whitesides & Ismagilov, 1999). However, instruction in middle-school science curricula does not differentiate processes on this basis, and our instructional intervention study suggests that it might be feasible to help students differentiate the two kinds of processes with introductory materials inserted into science texts, in much the same way that ideas about the NOS are currently inserted as the introductory chapter of many science texts.

We assumed, along with Levy and Wilensky (2009), that pattern-level behaviors are more understandable to students because they are often familiar and more intuitively clear, involving qualitative ideas of increasing and decreasing, whereas agent-level behaviors are less familiar, but still somewhat understandable. This implies that viewing and exploring dynamic simulations of the macro patterns (such as by changing the parameters or conditions that impact the collective behavior at the macro level) should not be as helpful to students' understanding of emergent processes as viewing and exploring the dynamic simulations of the micro-level interactions (in fact, showing both simultaneously is probably best). These differential utilities were confirmed by our data, showing a lack of correlation between learning of the Process Module and learning from the macro simulation, in contrast to a significant correlation between learning of the Process Module and learning from the micro simulation (see the correlations in Table 3).

We also emphasized that learning any science process requires learning an endless amount of technical and other information, and that assessment needs to differentiate successful leaning about these other technical aspects of a science concept from learning that reveals a deeper understanding that addresses misconception. The results depicted in Fig. 7 reinforce this point.

Even though it is conceivable that young children may have picked up some ideas about emergence from their everyday experiences, as proposed by Levy and Wilensky (2008), we assumed instead that young children are more likely to have developed a Direct Schema from their everyday ordinary encounters with stories and other events; and moreover, that they interpret science processes embedded in their school curricula largely on the basis of such a Direct Schema. We based our assumption on the large body of psychological literature showing the existence of narrative schemas and scripts for stories and ordinary events.

We defined a misconceived explanation as an incorrect inter-level explanation, one that requires appealing to the collective summing of the agents' interactions as a causal explanation for the pattern that is perceived. A prevailing assumption in the literature on understanding emergent systems is that students fail to understand the agents' structure, behavior, and function. Although that may be true, such understanding of the local agents' structure, behavior (or actions and interactions), and function, will not increase students' understanding of inter-level causal explanations for emergence (the source of misconceptions), contrary to the assumption proposed by Hmelo-Silver and Pfeffer (2004). Moreover, students can more easily understand the structure, behavior, and function of components of systems than they can understand the inter-level causality of emergent processes. For example, eighth-grade students were able to induce the function of components of the human circulatory system when the text materials omitted mentioning such function information (Chi, de Leeuw, et al., 1994), as long as they self-explained, but eighth-grade students in this current study had great difficulty generating inter-level emergent ideas even when they were answering explicit questions about emergent processes (see Fig. 7). Thus, we suggest that learning to understand an emergent process is not the same as learning to understand the structure, behavior, and function of either the macro-level or micro-level components of emergent processes, even though learning the structure, behavior, and function may be very important for understanding all processes.

We examined students' misconceived explanations and captured their underlying structure in terms of a handful of inter-level attributes, embedded in what we called a Direct Schema. We similarly also analyzed the underlying structure of the processes for which students tend to generate misconceptions. These processes have a similar underlying structure, consisting of a handful of inter-level attributes that may be embedded in what we called an Emergent Schema. Because the inter-level attributes of Direct and Emergent Schemas seem to be mutually exclusive, we assumed that a Direct Schema and an Emergent Schema might be ontologically distinct. Thus, we defined misconceptions as causal explanations that would be correct if the to-be-explained process is a sequential one, but because the to-be-explained process is an emergent one, the generated explanations are robustly misconceived. In short, a misconceived explanation is one generated by applying a Direct Schema to interpret an emergent process.

Our prior theoretical framework focused primarily on differences between the ontologies of substances (or entities) versus processes (Chi, 1997; Chi, Slotta, & de Leeuw, 1994; Slotta & Chi, 2006; Slotta, Chi, & Joram, 1995; see also Appendix B for other differences between the 2006 Slotta & Chi study and the current study). In this article, processes are further differentiated into two kinds in order to account for the nature of misconceptions about processes. If Direct and Emergent Schemas are ontologically distinct, this implies that once a process is assimilated into an inappropriate alternative schema (i.e., mis-assimilated into the Direct Schema in this case) by *beginner* students, then all subsequent understanding will be flawed because students will continue to draw on knowledge from that inappropriate Direct Schema for further interpretation. Thus, mis-assimilating into an inappropriate schema has costs, mostly for beginner students while learning about a new concept. Clearly, experts can shift across ontologies flexibly and treat an emergent process from a Direct perspective if they wish to consider either its pattern-level behavior or its agent-level behavior, rather than the inter-level relationships.

Mis-assimilating into an alternative schema also accounts for the diversity of misconceived explanations when students are confronted, because students can always generate another misconceived explanation from the same inappropriate or alternative schema. Thus, this process of having initially assimilated a concept into an inappropriate schema will cause misconceptions to be persistent and robust even when confronted, and in addition, students will tend to continually generate alternative flawed explanations from the same inappropriate schema, thus making it look like the misconceived explanations are piecemeal. Contrary to diSessa (1993), we suggested that there is an underlying coherence to the apparent piecemeal aspects of misconceptions.

If mis-assimilation to an alternative schema is the cause for misconceptions, then instruction should be aimed at helping students shift assimilation into the correct schema. In this view, the assumptions of mis-assimilation and shifting have no bearing on issues such as whether the initial mis-assimilated schema (Direct Schema in this case) remains or should fade away. The initial schema would remain since it is useful for interpreting other concepts and processes.

We also assumed that the process of shifting itself may be problematic only because students may not be aware that they need to shift, especially across ontological categories, not because shifting per se is difficult or impossible. These ideas have been described in Chi and Roscoe (2002). When people have all the available ontological categories, they shift freely and readily, such as when generating metaphors. Our exposition on these issues ware elaborated in Chi and Hausmann (2003). Moreover, in the case of learning emergent processes, we assumed that shifting is difficult not only because students are not aware of the need to shift but also because the Emergent Schema itself may be unfamiliar to students (i.e., it is missing). Because it is unfamiliar to students and because the pattern level looks similar to sequential processes, it is not likely that young students would be able to construct such a missing schema themselves without some direct instruction. Therefore, the missing schema assumption naturally led to an instructional intervention approach of trying to teach students such a schema directly.

In order for students to correctly assimilate a process into the appropriate schema, they must be able to recognize which process is a sequential one and which process is an emergent one. Therefore, we further identified a set of features that can discriminate these two processes. Since sequential and emergent processes look similar at the macro level, such discriminations must rely on features at the agents' level. We developed the novel idea that agent-level interactions have important second-order features, and it is at this level of second-order interaction features that discrimination can be made. To date, categorization research has considered only first-order relational features (Gentner & Kurtz, 2005).

Many of us agree that when learning emergent processes, it is important to connect the micro level with the macro level. Although a variety of bridging methods across the two levels have been attempted (Levy & Wilensky, 2009; Sengupta & Wilensky, 2009), our attempt bridged the agent level to the actual pattern that is physically real in terms of being isomorphic to what students have experienced in their everyday environment. This is important because misconceptions are derived from their everyday experiences with pattern-level phenomena. In contrast, other intervention efforts attempted to bridge the agent level to an abstraction, such as a changing graph, which can add another layer of difficulty for students to understand.

Within our intervention, we adopted a concept-general approach, in that we taught students a new concept-general schema about emergent processes, using simple everyday language and everyday examples. The common approach to overcoming misconceptions is to teach more about the specific concept itself that students misconceive. One obvious advantage of a concept-general approach is that it has the potential of transfer across concepts (Goldstone & Wilensky, 2008). We have tested this Process Module in middle-school aged students' understanding of *natural selection* with modest success (Chi et al., in press), and college engineering students have also benefitted from this Process Module in improving their understanding of *diffusion* and *microfluidics* significantly (Yang et al. (2010). Thus, there is some encouraging evidence that this Process Module is generalizable.[2]

Emergent processes and phenomena are typically described in terms of their overall characteristics such as ''irreducibility,'' ''small actions—big effects,'' ''decentralized control,'' ''adaptive and non-adaptive systems,'' ''self-organizing,'' ''linear versus non-linear,'' and so forth. These broad-stroke characteristics cannot be as easily translated into instruction, for either recognition or explanation purposes. Conversely, complex systems are sometimes characterized by a multitude of specific characteristics. During a workshop held at the Santa Fe Institute in 2006, sixteen leading researchers in the field tried to characterize complex systems and came up with well over 44 concepts, such as attractors, fractals, dynamic equilibrium, scaling, power laws, and amplification, just to name a few (Cabrera, 2006). To be more concrete, one common property describing emergence is the composite term ''decentralized,'' which aggregates many of our features (''uniform,'' ''equal status,'' ''all,'' etc.)

into a single descriptor. Other terms such as ''non-linear'' versus ''linear'' can refer to several aspects of emergent and sequential processes as well. ''Linear'' can refer to the sequential dependency of the interactions, or it can refer to the additive mechanism that causes the pattern. However, we believe that such encompassing terminologies are difficult for students to grasp. In contrast, our work and analyses of misconceptions and emergent science processes are at the level of identifying specific inter-level attributes and second-order interaction features that can be taught to students explicitly.[3]

We also distinguished several meanings of the term ''random'' in both our theory and our instruction, and suggested which meanings are more understandable to students and which are not. *Random*, when it refers to a single agent's movement (Jacobson, 2001), such as that a molecule can move randomly, meaning that it can move in any direction in unpredictable ways, should be a relatively easy concept for students to learn. *Random*, when it refers to interactions, as the lack of restriction in terms of with whom/what an agent can interact (as defined in this article as emergent interaction feature #2, Table 2) is less familiar and more difficult, but nevertheless, still may not be impossible to understand. However, there is a third meaning of *random* as it is used in the complexity literature and in popular press, and that refers to the emergent phenomenon itself being random. In this case, *random* refers to the ''unpredictability'' or counterintuitive and unexpected pattern of an emergent phenomenon (Wilensky & Resnick, 1999), arising from interactions in non-obvious ways. In this last case, it is not the concept *random* that is difficult for students to understand, but the fact that scientists do not know how to predict and explain emergent phenomenon such as a tsunami, and therefore call it *random.*

Our analyses and results revealed one important challenge in understanding emergent processes introduced in science curricula, which has to do with understanding the ''summing mechanism.'' So, even though we have differentiated *additive summing* (a mechanism of sequential processes) from *collective summing* (a mechanism of emergent processes), the findings of this study made us realize that the collective summing mechanism may be complicated in and of itself. That is, for simple perceptually available Emergent processes such as a school of fish and others that we have illustrated in the Process Module, the collective summing mechanism is perceptually available such as proximity or spatial contiguity. However, for a slightly more complicated Emergent process such as *diffusion*, the collective summing mechanism is one of proportional change. But the collective summing mechanism can get increasingly more difficult for more complex science concepts, such as for *heat transfer* or *electrical current.* For example, for *diffusion*, the collective mechanism is a summing of all molecules (molecules are an entity concept), but for *heat transfer*, it is a summing of the speed of the molecules (speed is a dynamic concept), and for *electrical current*, it is a summing of the velocity of the electrons (velocity includes speed and direction). Moreover, in many instructional interventions that involve either exploring or constructing simulations of emergent processes (such as in NetLogo), the collective summing mechanism is opaque to the students since it is often computed by the simulation system itself rather than discovered or constructed by students. Our study revealed that understanding this collective summing mechanism is crucial in generating correct causal explanations for more complicated science processes.

Another novel aspect of our instructional approach proposed that one way to teach completely new ideas in a way that can overcome the ''learning paradox'' (Bereiter, 1985) is to use the method of contrasts, because contrasts allow one to relate something novel to something known, even if it is not a relation of similarity, but a relation of differences (see also Marton, 2006). Thus, we taught new ideas about emergent processes by contrasting them with sequential processes. We hypothesized that a contrasting approach, as opposed to the commonly preferred analogical approach (Clement, 1993), might allow us to help students create a novel schema.

We based as many design decisions as possible for the format of our intervention on robust findings in the cognitive science literature, such as using two examples, instantiating the theory with examples, using prompts for self-explanations, using contrasting cases, scaffolding attention guidance in animation, avoiding the test–retest effect, and so forth.

The design of our intervention was also different from a ''cognitive conflict'' approach (Limon, 2001). In a cognitive conflict approach, students are confronted with contradictory information (such as an unexpected outcome that they did not predict), so that conflicts become obvious to them and presumably, experiencing such conflicts might force them to change their initial (mis)conceptions. Although our analyses also identified conflicts in that students' understanding mismatched the correct understanding, this mismatch occurred between a narrative schema and an emergent schema. We had postulated that a conflict at the schema level is not apparent to the students themselves, nor can such a mismatch be easily conveyed in a confrontational way. Thus, although our intervention (i.e., the Process Module) contained contrasts, which could be interpreted as containing some aspects of a cognitive conflict approach, the conflicts that we pointed out in our instruction was at the level of the schema.

Finally, although our theory was grounded in a framework of schemas, we are not claiming that schemas are permanent and static kind of knowledge structures. Framing our ideas in terms of schemas, however, is a convenient way for us to conceive of the differences in how students represent sequential and emergent processes, and it leads to one method of instructional intervention. No doubt other ways of framing our ideas are possible, leading to other methods of intervention.

Thus, the goal of this project was ambitious. Along with ideas published in Chi (2008), we wanted to provide a comprehensive theory of misconceptions and create a stand-alone intervention unit that is concept-general so that it may be applicable to many concepts with the potential of scale-up. However, the work is only preliminary since our results, although encouraging, are not completely conclusive, and we also have uncovered many additional challenges.

Notes

1. This is an example of minor differences between the two improved *Diffusion* Texts caused by editing. In one version, the description of molecules in solids is: ''In solids the particles move back and forth, or vibrate in place, because the molecules in solids

do not usually have enough energy to break free from the solid and move around freely. Solids have a rigid structure and the particles are closely packed and held together by strong forces.'' The same paragraph was modified in the other version to state: ''When they are in solids, atoms or molecules wiggle rapidly back and forth, vibrating in place, because they are trapped in very solid bonds that allow them to vibrate but not to move out of place.''

2. Does our theoretical framework apply to concepts that do not involve a massively large number of agents? For example, can our Process Module explain why misconceptions are so robust in a physics domain, where misconceptions were first identified? We believe it does, although we have not spent much time examining this. Let us take the motion of an object as a pattern-level behavior of an object. Students' explanations of what causes this motion do have features and attributes similar to the sequential ones we have identified in Tables 1 and 2. For example, students might say that the initial push by the hand caused the motion (attributing it to a single causal agent, sequential attribute #1, in Table 1), when in fact motion is determined by the sum of all the (or net) forces (emergent attribute #5, in Table 1). Thus, misconceptions about the motion of an object appear to embody some the same features and attributes that we have identified for misconceptions about processes with multi-agents.
3. Linear processes and complex dynamic systems are defined in the literature in numerous ways. The literature refers to some non-sequential processes as ''non-linear'' processes, and they are defined simply as a process in which the whole is not the sum of its parts (Waldrop, 1992, p. 64). Mathematically, these can be represented by non-linear equations whose graph is curved rather than a straight line. The science of non-linear systems is enormously complex, ranging from chaos systems, to complex dynamic systems, to self-organizing systems (Jacobson, 2001), and numerous definitions abound. Our analyses here differed from these other treatments of complex dynamic systems because our processes are deterministic ones. That is, the pattern-level outcomes of the kind of middle-school processes we discussed can be predicted accurately by the initial conditions, and so forth. In that sense, the processes we analyzed were not complex in the chaos sense. Nevertheless, our processes did share many similarities with other complex systems.

Acknowledgments

The first author is grateful for funding and support provided by the Spencer Foundation (grant 200100305). In addition, we would like to acknowledge the contribution of several individuals over a span of several years in the development of this project. Agnieszka Kosminska Kristensen assisted in the design of the learning materials; Stephanie Siler collected the pilot data; Jeff Holmes created the Flash simulations; Britta Cheng entered the learning and assessment materials into WISE; Mary Sartoris edited the WISE versions of the instructional materials and coded the open-ended responses; Jenna Small provided inter-rater reliability coding; and more recently, Glenda Stump re-coded the raw explanation data for the results presented in *Material S1*, and Yuning Xu re-scored the pre-tests again. In addition, we appreciate the many recent discussions of the theoretical concepts in this article with Keith Dybvig, Kasia Muldner and Kurt VanLehn, final editings by Rachel Lam, Muhsin Menekse, Kasia Muldner, and Dongchen Xu.

References

American Association for the Advancement of Science (1989). *Science for all Americans*. New York: Oxford University Press.

Barr, R. A., & Caplan, L. J. (1987). Category representations and their implications for category structure. *Memory & Cognition*, *15*, 397–418.

Bereiter, C. (1985). Toward a solution of the learning paradox. *Review of Educational Research*, *55*, 201–226.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*, 177–220.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, *24*, 61–100.

Cabrera, D. (2006). *A roadmap for complex systems education: Concept mapping research on what educational and science leaders think are the components of a complexity science curriculum*. Presented at the Santa Fe Institute Meeting of Educational and Science Leaders. Santa Fe, NM.

Casti, J. L. (1997). *Would-be worlds: How simulation is changing the frontiers of science*. New York: John Wiley.

Cheng, P. W. (1997). From covariation to Causation: A causal power theory. *Psychological Review*, *104*, 367–405.

Chi, M. T. H. (1997). Creativity: Shifting across ontological categories flexibly. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Conceptual structures and processes: Emergence, discovery and change* (pp. 209–234). Washington, DC: American Psychological Association.

Chi, M. T. H. (2005). Common sense conceptions of emergent processes: Why some misconceptions are robust. *Journal of the Learning Sciences*, *14*, 161–199.

Chi, M. T. H. (2008). Three types of conceptual change: Belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *Handbook of research on conceptual change* (pp. 61–82). Hillsdale, NJ: Erlbaum.

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145–182.

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994a). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439–477.

Chi, M. T. H., & Hausmann, R. G. M. (2003). Do radical discoveries require ontological shifts?. In L. V. Shavinina (Ed.), *International handbook on innovation* (pp. 430–444). Oxford, England: Elsevier.

Chi, M. T. H., Kristensen, A. K., & Roscoe, R. (in press). Misunderstanding emergent causal mechanism in natural selection. In K. Rosengren, S. Brem, M. Evans, & G. Sinatra (Eds.), *Evolution challenges: Integrating research and practice in teaching and learning about evolution*. New York: Oxford University Press.

Chi, M. T. H., & Ohlsson, S. (2005). Complex declarative learning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 371–399). New York: Cambridge University Press.

Chi, M. T. H., & Roscoe, R. D. (2002). The processes and challenges of conceptual change. In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 3–27). The Netherlands: Kluwer Academic Publishers.

Chi, M. T. H., Slotta, J. D., & de Leeuw, N. (1994b). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, *4*, 27–43.

Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in physics. *Journal of Research in Science Teaching*, *30*, 1241–1257.

Clement, J., Brown, D. E., & Zietsman, A. (1989). Not all preconceptions are misconceptions: Finding 'anchoring conceptions' for grounding instruction on students' intuitions. *International Journal of Science Education*, *11*, 554–565.

Closset, J. L. (1983). Sequential reasoning in electricity. In *Research on physics education: Proceedings of the First International Workshop* (pp. 313–319). Paris: Editions du Centre National de Recherche Scientifique.

Confrey, J. (1990). A review of the research in student conceptions in mathematics, science and programming. In C. B. Cazdan (Ed.), *Review of research in education*, Vol. 16 (pp. 3–56). Washington, DC: American Educational Research Association.

Driver, R. (1987). Promoting conceptual change in classroom settings: The experience of the children's learning in science project. In J. D. Novak (Ed.), *Proceeding of the second international seminar on misconceptions and educational strategies in science and mathematics* (Vol. II, pp. 97–107). Ithaca, NY: Department of Education, Cornell University.

Duit, R. (2008). Bibliography – CTCSE: Students' and teachers' conceptions and science education [accessed on October 19, 2011]. Available at http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.html.

Dupin, J. J., & Johsua, S. (1984). Teaching electricity: Interactive evolution of representations, models, and experiments in a class situation. In R. Duit, W. Jung, & C. von Rhoneck (Eds.), *Proceedings of the international workshop of aspects of understanding electricity* (pp. 331–341). Kiel, Germany: Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel.

Ferrari, M., & Chi, M. T. H. (1998). The nature of naive explanations of natural selection. *International Journal of Science Education*, *20*(10), 1231–1256.

Gadgil, S., Nokes, T. J., & Chi, M. T. H. (in press). Effectiveness of holistic mental model confrontation in driving conceptual change. *Learning and Instruction*, *22*, 47–61.

Gell-Mann, M. (1994). *The Quark and the Jaguar: Adventures in the simple and the complex*. London: Little, Brown and Company.

Gentner, D., & Kurtz, K. (2005). Relational categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. W. Wolff (Eds.), *Categorization inside and outside the lab* (pp. 151–175). Washington, DC: APA.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*, 1–38.

Goldstone, R.L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition*, *52*, 125–157.

Goldstone, R.L., & Wilensky, U. (2008). Promoting transfer by grounding complex systems principles. *The Journal of the Learning Sciences*, *17*, 465–516.

Grotzer, T. A., & Sudbury, M. (2000). *Moving beyond underlying linear causal models of electrical circuits*. Paper presented at the Annual Conference of the National Association for Research in Science Teaching, New Orleans.

Gupta, A., Hammer, D., & Redish, E. F. (2010). The case for dynamic models of learners' ontologies in physics. *Journal of the Learning Sciences*, *19*(3), 285–321.

Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behavior, and functions. *Cognitive Science*, *28*, 127–138.

Hofstein, A., & Welch, W. W. (1984). The stability of attitudes towards science between junior and senior high school. *Research in Science and Technological Education*, *2*, 131–138.

Jacobson, M. J. (2001). Problem solving, cognition, and complex systems: Differences between experts and novices. *Complexity*, *6*, 41–49.

Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. *Journal of the Learning Sciences*, *15*, 11–34.

James, R. K., & Smith, S. (1985). Alienation of student from science in grades 4–12. *Science Education*, *70*, 39–45.

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review*, *21*(2), 113–140.

LaBarbera, M., & Vogel, S. (1982). The design of fluid transport systems in organisms. *American Scientist*, *70*, 54–60.

Levy, S. T., & Wilensky, U. (2008). Inventing a ‘‘mid-level’’ to make ends meet: Reasoning between the levels of complexity. *Cognition and Instruction*, *26*, 1–47.

Levy, S. T., & Wilensky, U. (2009). Students' learning with the Connected Chemistry (CC1) curriculum: Navigating the complexities of the particulate world. *Journal of Science Education and Technology*, *18*(3), 243–254.

Licht, P. (1987). A strategy to deal with conceptual and reasoning problems in introductory electricity education. In J. Novak (Ed.), *Proceedings of the 2nd international seminar ‘‘Misconceptions and Educational Strategies in Science and Mathematics’’* (Vol. II., pp. 275–284). Ithaca, NY: Cornell University.

Limon, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction*, *11*, 357–380.

Linn, M. C., & Hsi, S. (2000). *Computers, teachers, peers: Science learning partners*. Mahwah, NJ: Lawrence Erlbaum Associates.

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review*, *6*, 586–597.

Lohmann, S. (1994). Dynamics of informational cascades: The Monday Demonstration in Leipzig, East Germany, 1989–91. *World Politics*, *47*, 42–101.

Lowe, R. K. (2003). Animation and learning: Selective processing of information in dynamic graphics. *Learning and Instruction*, *13*, 156–176.

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, *9*(1), 111–151.

Marek, E. (1986). Understanding and misunderstandings of biology concepts. *American Biology Teacher*, *48*, 37–40.

Marton, F. (2006). Sameness and difference in transfer. *The Journal of the Learning Sciences*, *15*, 501–538.

Maton, A., Hopkins, J., Johnson, S., LaHart, D., Warner, M. Q., & Wright, J. D. (1995). *Exploring life science*. Englewood Cliffs, NJ: Prentice Hall.

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, *19*, 494–513.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100*, 254–278.

Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology*, *51*, 121–147.

Meir, E., Perry, J., Stal, D., Maruca, S., & Klopfer, E. (2005). How effective are simulated molecular-level experiments for teaching diffusion and osmosis? *Cell Biology Education*, *4*, 235–248.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, *19*(3), 309–326.

Odom, A. L. (1995). Secondary and college biology students' misconceptions about diffusion and osmosis. *American Biology Teacher*, *57*, 409–415.

Perkins, D. N., & Grotzer, T. A. (2005). Dimensions of causal understanding: The role of complex causal models in students' understanding of science. *Studies in Science Education*, *41*, 117–165.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*, 211–227.

Reed, S. K., & Bolstad, C. (1991). Use of examples and procedures in problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *17*, 753–766.

Reiner, M., Slotta, J. D., Chi, M. T. H., & Resnick, L. B. (2000). Naive physics reasoning: A commitment to substance-based conceptions. *Cognition and Instruction*, *18*, 1–34.

Resnick, M. (1994). *Turtles, termites and traffic jams: Explorations in massively parallel microworlds*. Cambridge, MA: MIT Press.

Rumelhart, D. E. (1978). Schemata: The building blocks of cognition. In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 38–58). Hillsdale, NJ: Erlbaum.

Sanger, M., Brecheisen, D. M., & Hynek, B. M. (2001). Can computer animations affect college biology students' conceptions about diffusion and osmosis? *American Biology Teacher*, *63*, 104–107.

Schank, R., & Abelson, R. (1977). *Scripts, plans, goals, and understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schelling, T. C. (1978). *Micromotives and macrobehavior*. New York: W.W. Norton and Co.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*, 475–52.

Sengupta, P., & Wilensky, U. (2009). Learning electricity with NIELS: Thinking with electrons and thinking in levels. *International Journal of Computers for Mathematical Learning*, *14*(1), 21–50.

diSessa, A. (1993). Towards an epistemology of physics. *Cognition and Instruction*, *10*, 105–225.

Shipstone, D. (1984). A study of children's understanding of electricity in simple DC circuits. *European Journal of Science Education*, *6*, 185–198.

Slotta, J. D., & Chi, M. T. H. (2006). Helping students understand the challenging topics in science through ontology training. *Cognition and Instruction*, *24*, 261–289.

Slotta, J. D., Chi, M. T. H., & Joram, E. (1995). Assessing students' misclassifications of physics concepts: An ontological basis for conceptual change. *Cognition and Instruction*, *13*, 373–400.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Stein, D. W., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *Advances in discourse processes: Vol. 2. New directions in discourse processing* (pp. 53–119). Norwood, NJ: Ablex.

Torney-Purta, J. (1994). Dimensions of adolescents' reasoning about political and historical issues: Ontological switches, developmental processes, and situated learning. In M. Carretero & J. F. Voss (Eds.), *Cognitive and instructional processes in history and the social sciences* (pp. 103–122). Hillsdale, NJ: Lawrence Erlbaum Associates.

Trefil, J., Calvo, R. A., & Cutler, K. (2006). *McDougal Littell science: Life science*. Boston: Houghton Mifflin Company.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, *57*, 247–262.

Waldrop, M. M. (1992). *Complexity: The emerging science at the edge of order and chaos*. New York: Simon & Schuster.

White, B., & Frederiksen, J. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, *42*, 99–157.

Whitesides, G. M., & Ismagilov, R. F. (1999). Complexity in chemistry. *Science*, *284*, 89–92.

Wilensky, U., & Reisman, K. (2006). Thinking like a wolf, a sheep or a firefly: Learning biology through constructing and testing computational theories – An embodied modeling approach. *Cognition and Instruction*, *24*, 171–209.

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology*, *8*(1), 3–19.

Yang, D., Roman, A. S., Streveler, R., Miller, R., Slotta, J., & Chi, M. T. H. (2010). *Repairing student misconceptions using ontology training: A study with junior and senior undergraduate engineering students*. Presented at the American Society of Engineering Education. Louisville, KY.

Zuckerman, J. T. (1994). Problem solvers' conceptions about osmosis. *The American Biology Teacher*, *56*, 22–25.

Supporting Information

Additional Supporting Information may be found in the online version of this article on Wiley Online Library:

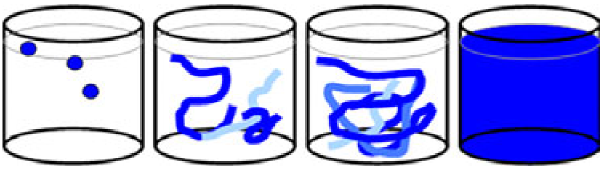
Material S1. Coding the three ''identical'' questions and three ''identical'' prompts.

Appendix A: Nine identical questions included on the *diffusion* pre- and post-test

Each question is indicated as asking about the Macro level, the Micro level, or the Inter-level, or both. Correct answers are italicized.

1. [**Micro**] What are molecules made of?
   a. The same or different types of atoms.
   b. Only different types of atoms.
   c. Different types of molecules.
   d. Protons, neutrons, and electrons.
2. [**Macro**] When you drop a cube of sugar into water, the diffusion of sugar molecules refer to:
   a. Dissolving of sugar molecules through the entire solution.
   b. Spreading of the sugar molecules through the entire solution.
   c. Binding of the sugar and water molecules in the entire solution.
   d. Spreading and dissolving of the sugar molecules through the entire solution.
3. [**Macro**] Suppose you have two beakers connected by a short tube. One beaker contains a highly concentrated solution of dye and water and the other beaker contains no dye, only water. Thus, there is a high concentration gradient between the two beakers. The term ''concentration gradient'' refers to …
   a. The difference in the concentration of water in the two beakers.
   b. The difference in the concentration of dye in the two beakers.
   c. The pressure caused by the uneven concentration of dye in the two beakers.
   d. The pressure caused by the uneven concentration of water in the two beakers.

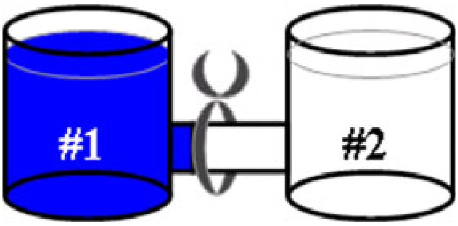
Scenario 1. Suppose there is a large beaker filled with clear (pure) water. Into this water we add several drops of dark blue dye liquid (blue ink, see the first beaker). What you will see is that the dye seems to swirl and spread through the water, as shown in the second beaker.



4. [**Macro**] In this scenario, why does the ink move away from where it was originally added to the water? [B & D are misconceived choices]

   a. The ink is forced to move away from where it is more concentrated because of the concentration gradient.
   b. The ink wants to spread out from concentrated areas in the water to where there is little or no dye.
   c. Because of diffusion that occurs from where the ink is more concentrated to where it is less concentrated.
   d. The ink needs to create equilibrium, so it must move from an area of higher concentration to an area of lower concentration.
5. [**Inter-level**] In Scenario 1, how does the ink spread out from where it was originally added to the water? [A is a misconceived choice]
   a. By the molecules being forced to move from an area of higher concentration to an area of lower concentration because of the concentration gradient.
   b. By spreading out where there is more room in the water, which has a low concentration of dye.
   c. By spreading out into the water because the water doesn't contain any ink.
   d. By all of the molecules colliding with each other, and purely by chance, the ink molecules will move into other areas.
6. [**Inter-level**] When the solution has reached an equilibrium state and has turned blue in color, [D is a misconceived choice]
   a. The ink molecules will keep moving but the water molecules will not.
   b. The water molecules will keep moving but the ink molecules will not.
   c. Both the ink and water molecules will keep moving.
   d. Both the ink and water molecules have stopped moving.

Scenario 2. Suppose you have two beakers connected by a short tube with a clamp. Beaker #1 contains a highly concentrated solution of darkly colored blue dye (and water), and Beaker #2 contains no dye, only water. We will refer to Beaker 1 as dye and Beaker 2 as water. Thus, there is a high concentration difference between the two beakers. At first the tube is clamped shut and nothing can flow between the two beakers. When the clamp is removed …



7. [**Macro**] In Scenario 2, when the clamp is removed, what seems to happen with the dye and water liquids? How is it the same or different from what is actually happening to the dye and water liquids? [D is a misconceived choice]
   a. It looks like the blue dye is flowing from Beaker 1 to Beaker 2, but in fact the water is also flowing in the opposite direction, from Beaker 2 to Beaker 1.

However, because the water is clear, you can only see the movements of the visible dye.
   b. It looks like the blue dye is flowing from Beaker 1 to Beaker 2, but in fact that is exactly what is happening. The water doesn't move very much.
   c. *It looks like the blue dye is flowing from Beaker 1 to Beaker 2, but in fact the water and dye are both flowing in every direction at the same time. However, because the water is clear, you can only see the movements of the visible dye.*
   d. It looks like the blue dye is flowing from Beaker 1 to Beaker 2, but in fact the dye and water are exchanging places. However, because the water is clear, you can only see the movements of the visible dye.
8. [**Inter-level/Micro**] In Scenario 2, as the dye appears to flow from Beaker 1 to Beaker 2, is it possible for a dye molecule in Beaker 2 to move backwards into Beaker 1? [A, B, C are misconceived choices]
   a. No. Once a dye molecule has moved to Beaker 2, from an area of higher concentration to an area of lower concentration it can never go back.
   b. Yes. The dye molecules need to create equilibrium, and so one or more of them may need to go back into Beaker 1 to maintain a balance.
   c. No. The dye and water molecules are linked together. After the dye molecules have moved into Beaker 2, the dye and water molecules must move together. So a dye molecule cannot just move into Beaker 1 by itself.
   d. Yes. All molecules move around randomly and can collide with each other and any dye molecule can go anywhere in either beaker.
9. [**Inter-level**] In Scenario 2, after the solutions in both beakers have turned a uniform light blue (the equilibrium state), will we start seeing a flow of the dark blue dye from Beaker 2 back into Beaker 1 then? [A is a misconceived choice]
   a. No because once equilibrium is reached, the process of diffusion is finished and the dye molecules stop moving.
   b. No because once equilibrium is reached, the concentration of dye in both beakers is the same so that there will be no visible net flow of dye.
   c. Yes, because once equilibrium is reached, the process of diffusion is finished but the dye molecules keep moving.
   d. Yes, because once equilibrium is reached, the flow will start reversing, so that the blue dye will start flowing from Beaker 2 back into Beaker 1.

Appendix B: Differences between our current study and Slotta and Chi (2006)

The current study is analogous to a study Jim Slotta and Micki Chi carried out many years ago and published in 2006 (Slotta & Chi, 2006). In that earlier study, we also found that providing training on a schema about emergence improved students' ability to solve problems, but in the target domain of electrical current. Here, we want to point out several important differences between the current study and our prior study.

Foremost, in the prior study, the intervention schema described the features and attributes of emergent processes very broadly. In particular, it described emergent processes in terms of four properties: *system-wide*, *equilibrium-seeking*, *simultaneous and independent*, and *on-going*. Note that these properties do not differentiate our intra-level second-order features of interactions and the inter-level attributes of causal relationships. We believe that a differentiation between the two is necessary since the features can help students identify one kind of process from another kind, whereas the inter-level attributes serve the purpose of correcting their misconceived explanations.

A second difference between the current and prior study is that the alternative (control) condition in this current study is more stringent, in the sense that both conditions received as much of the same intervention materials as possible. For instance, both Groups in the current study received dynamic simulations, whereas in the prior study, only the Process Group did.

The third and most important difference between the two studies is in the assessment. In the prior study, the post-test questions about electrical current were all macro-level problems. That is, they were all concerned with the behavior of electrical circuits, and not with understanding the mechanism of how electrical current is produced by interactions of the electrons. Moreover, the explanations for these problems were coded in terms of students' use of (*process* versus *substance* or *entities*) predicates, rather than in terms of emergent and non-emergent ideas, as was done in this study.