Surprisingly Stochastic: Learning and Application of Emergent Behavior Using Interactive Simulations of Nano-Mechanical Biological Systems

Paul F. Egan (pfe@cmu.edu)
Department of Mechanical Engineering, 5000 Forbes Ave.
Pittsburgh, PA 15217 USA

Christian D. Schunn (schunn@pitt.edu)
Department of Psychology, 3939 O'Hara St.
Pittsburgh, PA 15260 USA

Jonathan Cagan (cagan@cmu.edu)
Phillip R. LeDuc (prl+@andrew.cmu.edu)
Department of Mechanical Engineering, 5000 Forbes Ave.
Pittsburgh, PA 15217 USA

Abstract

Emergent behavior is pervasive in complex systems, and often produces surprising phenomenon that are challenging to understand and apply usefully, especially with regards to inter-level causalities. Here, we study engineering undergraduates' capacity to understand and solve design problems concerning inter-level causalities in nanomechanical biological systems. To test user understanding and application of inter-level causalities in this system, we developed a GUI bolstered by an agent-based molecular simulation that calculates system performance and renders animations in real-time as users adjust design inputs. We randomly assigned undergraduate engineering students to design-based learning groups with support of either animated simulation rendering or charts. Both groups improved on a set of pre/post design problems. But on assessments of understanding of inter-level causal relationships, only the animation group demonstrated an understanding. Both groups were then presented contrasting animations of continuous and intermittent systems, resulting in about half of participants in each group demonstrating an understanding of inter-level causal behaviors. These findings demonstrate the difficulty in understanding inter-level causal relationships in emergent systems, the usefulness in interactive software tools in overcoming these difficulties, and that understanding of inter-causal relationships may improve performance in design applications.

Keywords: Emergence, Inter-level Causality, Learning, Graphical User Interface, Complex Systems Design, Engineering, Multiscale, Biomedical

Introduction and Motivation

Complex systems consist of many components and interactions that make them difficult to understand, with emergent behavior being cited as particularly troublesome (Chi, 2005; Hmelo-Silver, Marathe, & Liu, 2007). Emergent behavior, stated succinctly, is the system level behavior that occurs as a collection of components interact, and often refers to phenomenon with qualitatively distinct global and local behaviors (Bar-Yam, 2004). Although understanding behaviors at both a component and system level is indicative of deep understanding and expertise (Hmelo-Silver, Marathe, & Liu, 2007), understanding of causal relationships between these levels may also be necessary in forming a robust understanding of the system (Chi, Roscoe, Slotta, Roy, & Chase, 2012). Here we investigate how software tools can facilitate the learning of inter-level causalities and how this understanding extends to useful reasoning skills.

The effect of software tools in supporting users' understanding of emergence has been conducted in the context of science education, and includes examples such as fluid diffusion (Chi et al., 2012) and ecosystems (Jordan, Hmelo-Silver, Liu, & Gray, 2013). However, in complex systems engineering contexts, it is also necessary to apply learned knowledge pragmatically towards an application (Ottino, 2004). In particular, we focus on the design of nanomechanical myosin protein systems (Egan, Cagan, Schunn, & LeDuc, 2013a), because they are a prevalent in natural systems (e.g. muscle, cytoskeleton) and biomedical technologies (e.g. nanoactuator, contractile material). They are also unfamiliar to traditional engineering disciplines, thus providing an ideal system for investigating how engineers learn and demonstrate understanding of inter-level causality.

Myosin systems consist of individual motor proteins that stochastically attach and exert force to move protein filaments before stochastically detaching. This behavior is illustrated in Figure 1 panels of a graphical rendering of an agent-based myosin simulation (Egan et al., 2013a). The system is emergent (Huber, Schnauß, Rönicke, Rauch, Müller, Fütterer, & Käs, 2013) because local myosin cyclical behavior is qualitatively distinct from global filament translation behavior. Although individual myosins are intermittent and stochastic, the system is frequently continuous/predictable as a whole but sometimes also intermittent and stochastic because filaments only translate during periods when at least one myosin is attached (Harada, Sakurada, Aoki, Thomas, & Yanagida, 1990). Therefore, average filament velocity is generally faster

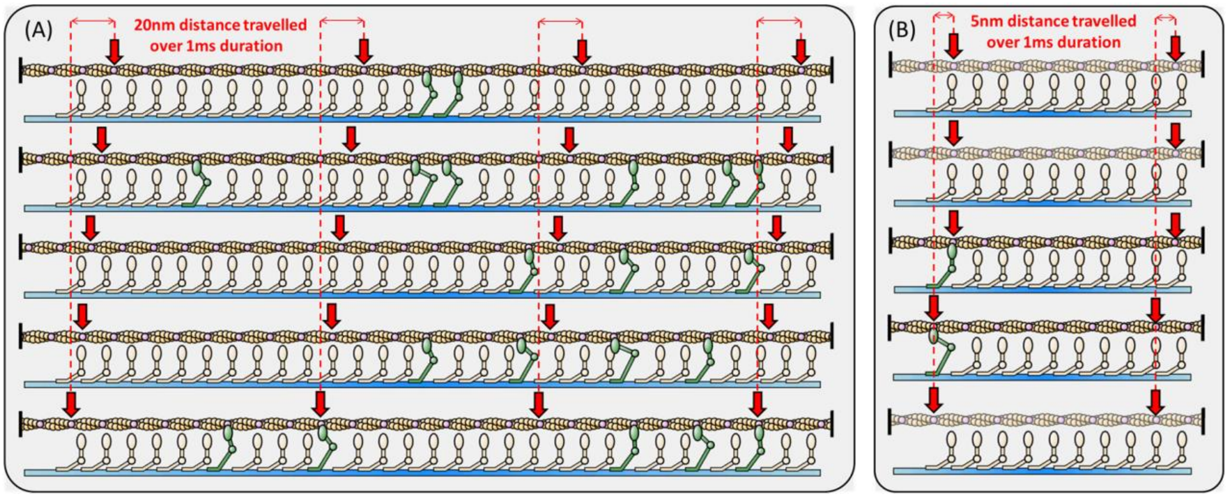

Figure 1: Simulated (A) continuous and (B) intermittent myosin filament translation. Each frame consists of an actin filament and myosins anchored on a microscope slide. The filament is shown as translucent when it is not moving, and the myosins are shown as white when not attached. Each frame from top to bottom demonstrates the translation of an actin filament by attached myosin heads over time. If no myosin head is attached, the filament remains static, resulting in lower average filament velocity in B. Red arrows track binding sites for 1 ms virtual time. Video: http://youtu.be/OvAYgchn0Bo

when there is continuous contact among myosins and the filament (Figure 1A), and lower with intermittent filament translation (Figure 1B). These continuous and intermittent filament translations are representative of different emergent behavioral regimes at a systems level, and a system's regime is dependent on the inter-level causal relationships of components and the system configuration as a whole.

In this study, we seek to understand how software tools could facilitate inter-level learning of the system, and whether understanding aids users in design tasks of optimizing myosin systems with a graphical user interface (video available at http://youtu.be/-14I3OSusgs) (Adapted from Egan, Cagan, Schunn, & LeDuc, 2013b). In this study, we have developed and tested three methods of software-aided learning that could promote understanding of inter-level causal relationships.

The first method is through supplying users quantiative feedback of system performance via charts, thus allowing the user to change myosin design inputs and see measured changes in performance. The second method is to allow users to recieve animated feedback of the agent-based molecular simulation while changing design inputs (video available at http://youtu.be/S8Fj67HeWps). The final method is to present users contrasting animations of a system in either the intermittent or continuous emergent behavioral regime, thus providing a clear distinction to the user for how systems configured in two different ways produce two different global patterns of behavior.

Our goal in this study is to demonstrate that these software tools aid users in understanding and designing these systems and to demonstrate that successes in understanding inter-level causalities aid in engineering design tasks. Our hypotheses are that 1) Learning via charts or interactive simulations will improve user design task performance, 2) Users exposed to animated renderings of agent-based simulation behavior will be able to demonstrate understanding of inter-level causal relationships, and 3) Users that demonstrate an understanding of inter-level causalities will perform better on design tasks.

Background

Studies in measuring student understanding of emergence have demonstrated that misunderstandings of emergent behaviors (e.g. diffusive fluid flow) are robust in comparison to misunderstandings of direct behaviors (e.g. blood flow in the circulatory system) (Chi, 2005). Students who lack understanding of emergent systems often have fragmented system understandings, such as being able to understand pieces of component behaviors but not how they relate across scales to promote a global system behavior. Most commonly, this fragmented understanding occurs because students try to explain emergent systems as direct processes, rather than distributed behaviors (Chi, 2005). Such findings are relevant to the myosin system, because (as we explain during the tutorial to participants) individual myosins are stochastic and propel filaments, but then (as participants must discover for themselves) these parts interact across levels to promote the varied emergent behaviors of continuous/intermittent filament translation.

Our approach in using an agent-based animation is supported by past studies that have had success in promoting system understandings of how aquarium systems work through agent-based interfaces (Vattam, Goel, Rugaber, Hmelo-Silver, Jordan, Gray, & Sinha, 2011) and understandings of inter-level causal relationships in diffusion (Chi et al., 2012).

Medical education has had success in utilizing animations for learning and found that successful approaches require consideration of cognitive load (Ruiz,

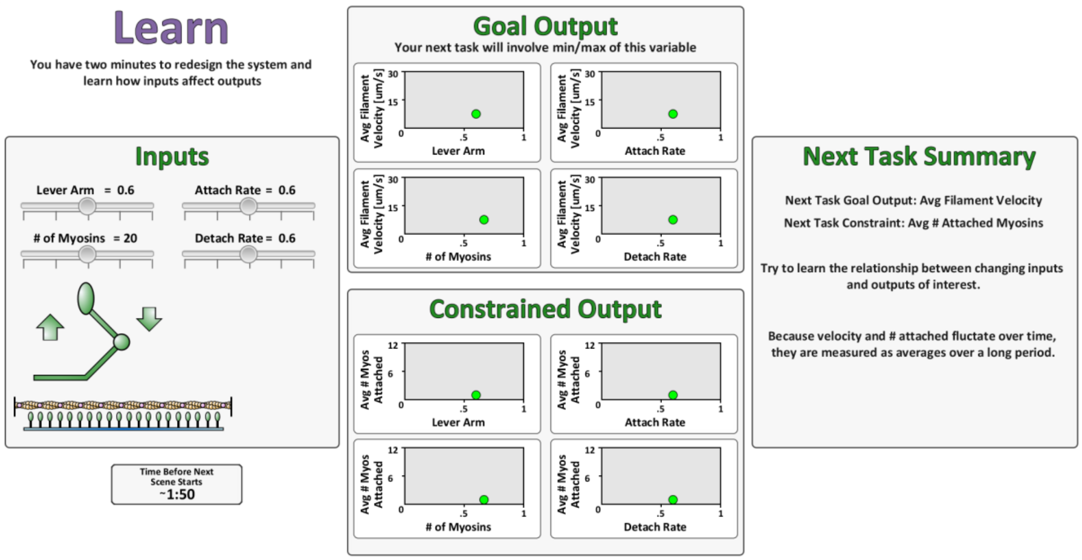

Figure 2: Myosin GUI configured with feedback via charts output. As users manipulate myosin design inputs with slider interfaces, calculations are performed and output in plots. Video: http://youtu.be/QXoIv48ntYk

Cook, & Levinson, 2009) to facilitate maximal learning (Van Merriënboer & Sweller, 2010). In our approach a single animation cannot convey all necessary information because emergent myosin system behavior varies with system configuration, meaning users must actively change the system to learn how different configurations lead to different emergent behaviors. However, to reduce cognitive load if users are unable to learn inter-level causalities through actively designing the system, contrasting animations (Alfieri, Nokes-Malach, & Schunn, 2013) have been shown as an effective teaching tool that could reduce cognitive load by focusing attention on critical features that are revealed through critical case contrasts.

Experimental Methods

Graphics User Interface

The graphical user interface (GUI) consists of a set of scenes that guide the user through a software tutorial, interactive design problems/learning sessions (Figure 2), and multiple choice/free response questions. During design/learning scenes, users are able vary the values of three independent myosin design parameters and one system design parameter; the influence of changing these design variables on system behavior has previously been empirically validated with agent-based simulations (Egan, Cagan, Schunn, and LeDuc, 2012). Once users configure a design, they evaluate it and receive feedback of how it performs with respect to a goal performance output.

There are also constraints on output performance, and if a design violates a constraint it is designated infeasible. Users are provided ten evaluations per design task, which is similar to past studies allowing for some benchmark comparisons (Egan, Cagan, Schunn, and LeDuc, 2014). Three difficulties of problems were created by increasing the number of output variables and constraints (only one output is graded as the goal output, additional constraints on secondary output variables inherently reduce the set of acceptable possibilities in the larger design space): one-variable-one-constraint, two-variables-one-constraint, or two-variables-two-constraints.

In this study, two different interactive learning scenes were created for users to explore inter-level causality among components and system behavior before design tasks. In the first configuration, termed 'Charts' style learning, there is an area for users to manipulate input sliders and then system behavior is provided in the form of plotted feedback that updates in real-time (center column of Figure 2). There are also static myosin images updated as design variables are changed. A second configuration of this scene is termed 'Animation' style learning. It is identical to Figure 2 except that the plots are removed and the static myosin images are replaced with the agent-based simulation rendering from Figure 1 in a continuous animated illustration of the movement, binding, and releasing of the system components. In both groups, information concerning the next design task is presented on the right to provide an idea of what information is important during a learning session.

Through one of these two interfaces, a user is expected to learn about the system by manipulating inputs and recognizing their effects on output performance. In the 'Animation' condition, a user could add more myosins to the system and notice that if the filament was originally intermittently moving, it would begin moving more often as there are more periods of at least one myosin being attached. In the 'Charts' condition, a user would see the filament velocity parameter increase on the y-axis as the number of myosins were increased in the system if it was changing from intermittent to continuous emergent behavior.

Procedure

Thirty-one mechanical engineering undergraduates in a senior design class participated for course credit. Participants were randomly assigned to either the 'Charts' or 'Animation' condition and groups followed different procedures as illustrated by Figure 3, which were developed to test the hypotheses outlined at the end of the Introduction.

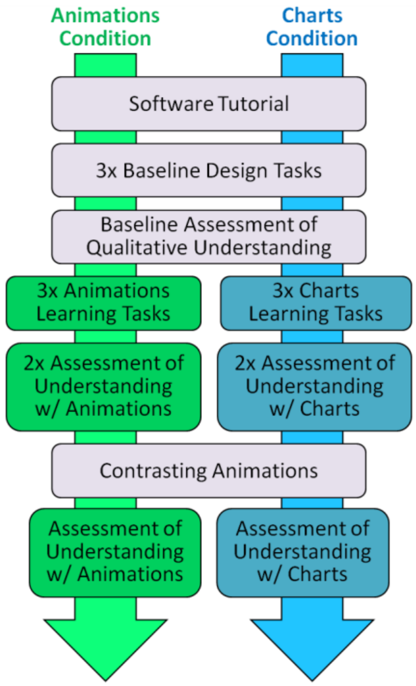

Figure 3: Flow chart of cognitive study protocol with left and right tracks representing different learning conditions.

Both groups are initially given the same software tutorial and then solve the same three baseline design tasks that have three levels of difficulty. In these baseline design tasks, myosin variable names are renamed generically to ‘Input A,’ ‘Output B,’ etc., which ensures minimal learning during this baseline measurement task.

Then, a baseline assessment of whether users could demonstrate inter-level causal reasoning was collected by asking questions about how changing design inputs might change system behavior. This was conducted by presenting users a system configuration clearly in the continuous or intermittent behavioral regime to a knowledgeable user, and then asking whether the filament velocity of the system would increase or stay about the same when a particular design input was changed. This question is an assessment of their understanding of inter-level causal relations because if a system is already behaving continuously, then adding more myosins would not improve its average velocity. However, if the system was initially behaving intermittently, adding more myosins would improve its average velocity.

Users are then provided one of the learning interfaces depending on their condition before solving their next set of tasks. Users interact with the learning interface for two minutes before each task, all baseline and learning design tasks were limited to four minutes, and users proceeded through all other GUI scenes at their own pace. Afterwards, their ability to describe system stochasticity is assessed a second time, followed by a third time where users are provided the correct quantitative relationship among system variables (e.g. filament velocity will/will not raise significant when a myosin’s attachment rate increases), and then must provide the correct reasoning. This third assessment therefore isolates a user’s ability to explain the inter-level causality without first having to assess what effect changing an input will have on the behavior of the system. For these assessments, users are also presented visualizations of the system according to their respective learning condition. Finally, users in both groups are presented contrasting animations before a final assessment.

Experimental Results

Learning Effects on Design Task Performance

The first hypothesis was: Learning via charts or interactive simulations will improve user design task performance. This was analyzed by aggregating data from each task separately, and then averaging the solution quality of a user’s best solution for that task with all other users in their condition. Solution quality was determined by first comparing a user’s goal output value to the global optimum for a task and providing it a relative score between 0 and 1 (all designs that did not meet constraints had a score of zero, the global optimum has a score of 1). The solution quality was then calculated by finding the difference between the user average relative objective function and a random solver to facilitate absolute evaluation and performance comparison across problem types. The average solution quality of users is presented in Figure 4 for each task and learning condition.

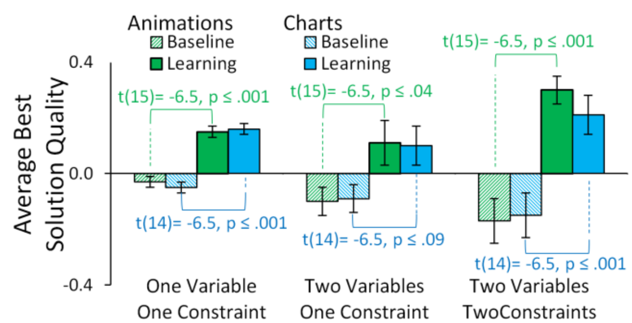

Figure 4: Average user solution quality in each learning condition for all tasks.

In comparing baseline and post-learning tasks, the average solution quality improved post-learning for all tasks, and supports the hypothesis. Each software tool improved performance about the same. The tools helped more as task complexity grew, thus motivating an ever increasing need for software tools as systems grow increasingly more complex. Interestingly, the charts and animations did not appear to aid design performance by changing design strategies, at least not with regards to the strategies previously shown to improve performance on these tasks (e.g., only changing one input at a time, searching near their current best design, min/maxing inputs) (Egan, Cagan, Schunn, and LeDuc, 2014). Thus, the benefit may occur through the knowledge used in existing strategies rather than via changing strategies.

Recognition of Stochastic System Behavior

After users had completed all the design tasks, their understanding of inter-level causalities was assessed to address the second hypothesis: Users exposed to animated renderings of agent-based simulation behavior will be able to demonstrate understanding of inter-level causal relationships. In each assessment phase, users were asked two questions of how filament velocity of a system would change if reconfigured. In each pairing, users were expected to recognize the emergent behavioral regime of the system on their own and questions were always paired such that the system would change regimes upon alteration in one question but not both.

Users indicated their answer via a multiple choice box (either filament velocity increased or about the same) and typed their reasoning in a free response box that was only analyzed for users that correctly answered the multiple choice question. Free response answers were tagged as demonstrating understanding of inter-level causality if users referred to the stopping/starting behavior of the filament being related to having at least one myosin attached. Example user responses that were tagged as correct were "*The increased number of myosins results in more time during which at least 1 myosin is attached and therefore the filament is being pushed forward*," and "*Average filament velocity increases because when there are more myosins there is less of a chance the filament will not be moving as a result of no present myosins.*" Some examples of answers tagged as incorrect were "*There are more myosin firing at any given time, giving more total force to the system and resulting in higher filament velocity*," and "*More myosins are in contact with the filament at any given time, which increases the velocity.*"

Understanding of inter-level causalities was assessed four times (Figure 3). No users indicated a proper understanding during the baseline assessment prior to learning sessions. The next two assessments occurred after the learning sessions (the first being directly after, and the second occurring once users were provided the correct quantitative relationship via the correct multiple choice answer but still had to provide an explanation). The final assessment occurred immediately after the contrasting animations were presented. The percentage of correct answers was aggregated for users in each condition and is presented in Figure 5.

Users in the animation group correctly demonstrated understanding of inter-level causality about 33% of the time directly after the learning tasks, while no users in the charts condition demonstrated understanding at this point, thus supporting the hypothesis that the agent-based simulation aids in learning inter-level causal relationships. However, despite supporting the hypotheses, only a small portion of users did demonstrate understanding. After the quantitative hint was provided, there was only a slight improvement; one user in the charts condition explained the stochastic system behavior correctly which suggests that it was not entirely implausible for users in that condition to formulate theories in line with the surprising stochastic behavior.

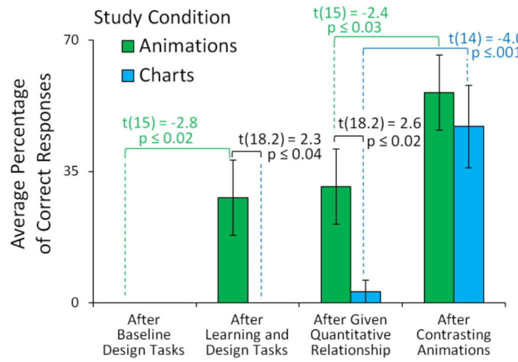

Figure 5: Percentage of correct responses in demonstrating an understanding of inter-level causality during each assessment phase.

After the contrasting animations were presented to users, approximately half of users in each condition correctly explained the inter-level causality. This pattern further supports the hypothesis that agent-based simulation renderings are effective in teaching inter-level causality, and are more effective when cognitive load is reduced via contrasting animations. Because average user score was only about 50%, our findings reinforce prior findings that emergent systems are difficult to understand (even for engineering students) and misconceptions about these systems are robust to learning interventions.

Does Recognizing Stochasticity Aid Design?

The final hypothesis tested was: Users that demonstrate an understanding of inter-level causalities will perform better on design tasks. This was investigated by separating the users in the animation condition among those that did and did not demonstrate an understanding during the second assessment. Group performance was then compared on the final design task. Only the final design task was selected because the assessment immediately followed it and thus was the closest measure of understanding during design. We also examined performance on the baseline task to rule out third variable differences among participants related to design ability (see Figure 6).

The results show that during the baseline, there was not a significant difference in design performance among the two groups. This demonstrates that when no users understood inter-level causality, design task performance among the groups was similar. Afterwards, users that had understood stochasticity via inter-level causal relationships performed better and found designs very close to the global optimum. This finding supports the hypothesis that users who demonstrate understanding of inter-level causality perform better in design applications related to that understanding.

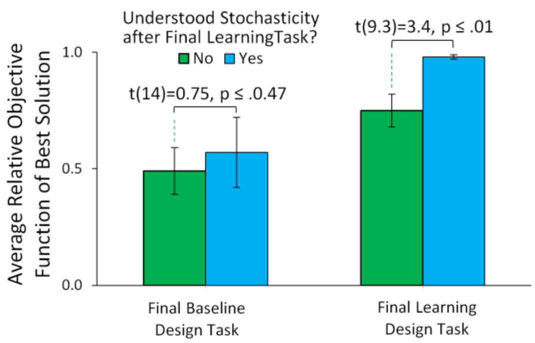

Figure 6: Design performance of users that did or did not demonstrate understanding of inter-level causality.

Conclusion

This study sought to investigate how different types of software tools could aid in user understanding and application of inter-level causality in complex emergent systems. It was found that learning about the system through visual charts and animations feedback improved user ability to find high quality designs in optimization problems. Animations were then demonstrated to improve user ability in describing inter-level causality that contributes to surprising stochastic behavior at the systems level. Providing users contrasting animations of the system configured in two different behavioral regimes resulted in users for all study conditions describing inter-level causality correctly about 50% of the time. Finally, users that demonstrated an understanding of inter-level causality immediately following a design task, performed better on it.

As a whole, these findings demonstrate the challenges in user understanding and reasoning about inter-level causal relationships in complex emergent systems, and that software tools can promote learning of these relationships. Gains in understanding can then promote better performance in complex systems design applications, where many complex systems operate on the same domain general principals as complex nanomechanical biological systems.

Acknowledgments

Partial funding was provided by the National Defense Science and Engineering Graduate Fellowship and by the National Science Foundation under grant CMMI-1160840. Tiffany Ho aided in developing early planning of the study and Kirstie Pomilio helped administer cognitive studies.

References

Alfieri, L., Nokes-Malach, T., & Schunn, C. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist, 48*(2), 87-113.

Bar-Yam, Y. (2004). A mathematical theory of strong emergence using multiscale variety. *Complexity, 9*(6), 15-24.

Chi, M. (2005). Commonsense conceptions of emergent processes: why some misconceptions are robust. *The Journal of the Learning Sciences, 14*(2), 161-199.

Chi, M., Roscoe, R., Slotta, J., Roy, M., & Chase, C. (2012). Misconceived causal explanations for emergent processes. *Cognitive Science, 36*, 1-61.

Egan, P., Cagan, J., Schunn, C., & LeDuc, P. (2013a). Design of complex biologically based nanoscale systems using multi-agent simulations and structure-behavior-function representations. *Journal of Mechanical Design, 135*, 061005.

Egan, P., Cagan, J., Schunn, C., & LeDuc, P. (2013b). *A modular design tool for visualizing complex multiscale systems*. Paper presented at the Internation Conference on Engineering Design, Seoul, South Korea.

Egan, P., J. Cagan, C. Schunn, and P. LeDuc. (2012). *Design of complex nano-scale systems using multi-agent simulations and structure-behavior-function representations*. Paper presented at the ASME International Conference on Design Theory and Methodology, Chicago, IL.

Egan, P., J. Cagan, C. Schunn, and P. LeDuc. (2014). *Cognitive-based search strategies for complex bio-nanotechnology design derived through symbiotic human and agent-based approaches*. Paper presented at the International Conference on Design Theory and Methodology, Buffalo, New York.

Harada, Y., Sakurada, K., Aoki, T., Thomas, D. D., & Yanagida, T. (1990). Mechanochemical coupling in actomyosin energy transduction studied by in vitro movement assay. *Journal of Molecular Biology, 216*, 49-68.

Hmelo-Silver, C. E., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: Expert-novice understanding of complex systems. *Journal of the Learning Sciences, 16*, 307-331.

Huber, F., Schnauß, J., Rönicke, S., Rauch, P., Müller, K., Fütterer, C., & Käs, J. (2013). Emergent complexity of the cytoskeleton: from single filaments to tissue. *Advances in Physics, 62*(1), 1-112.

Jordan, R. C., Hmelo-Silver, C., Liu, L., & Gray, S. A. (2013). Fostering Reasoning About Complex Systems: Using the Aquarium to Teach Systems Thinking. *Applied Environmental Education & Communication, 12*(1), 55-64.

Ottino, J. M. (2004). Engineering complex systems. *Nature*(427), 399.

Ruiz, J., Cook, D., & Levinson, A. (2009). Computer animations in medical education: a critical literature review. *Medical Education, 43*, 838-846.

Van Merriënboer, J. J., & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical education, 44*(1), 85-93.

Vattam, S. S., Goel, A. K., Rugaber, S., Hmelo-Silver, C. E., Jordan, R., Gray, S., & Sinha, S. (2011). Understanding complex natural systems by articulating structure-behavior-function models. *Educational Technology and Society, 14*(1), 66-81.