Michael Helms, Svaroop Vattam, Ashok K. Goel, & Jeannette Yen (2011). Enhanced Understand of Biological Systems Using Structure-Behavior-Function Models. ICALT, *IEEE 11th International Conference on Advanced Learning Technologies* (pp 239-243). Athens, Georgia. July 2011.

Enhanced Human Learning Using Structure-Behavior-Function Models

[1,2]Michael Helms & [1]Swaroop Vattam,
[1]Design & Intelligence Laboratory
Georgia Institute of Technology
Atlanta, GA
{mhelms3, svattam}@cc.gatech.edu

[1,2]Ashok K. Goel & Jeannette Yen
[2]Center for Biologically Inspired Design
Georgia Institute of Technology
Atlanta, GA
goel@cc.gatech.edu, jeannette.yen@biology.gatech.edu

Abstract— An important issue in teaching interdisciplinary biologically inspired design is the external representations we use to foster understanding of biological systems. In this study we explore if functional models of biological systems, and in particular Structure-Behavior-Function (SBF) models, enable humans to better understand complex biological systems. The study compares the use of SBF models in answering questions about biological systems versus the use of textual, tabular and graphical representations. The results indicate that while no one representation is best for answering all types of questions, SBF models enable more accurate answers to questions entailing abstract and complex inferences.

Keywords: biological systems, functional models, learning, question answering, understanding.

I. INTRODUCTION

Biologically inspired design is an important and growing movement in design [1-4]. The success of biologically inspired design is leading to a growing number of academic courses. Georgia Tech's Center for Biologically Inspired Design (http://www.cbid.gatech.edu/), for example, offers a senior-level interdisciplinary course on biologically inspired design.

An important issue in teaching interdisciplinary biologically inspired design is what external representations foster understanding of complex biological systems as measured by the types of inferences the understanding enables. In this paper, we focus on the question of which external representations, such as text, diagrams, or structured knowledge representations, best help designers develop deep understanding of complex systems.

Cognitive studies of biologically inspired design have focused on representation with respect to analogical retrieval and use-in-design, but not understanding of the biological systems per se [5, 6]. Our own earlier cognitive studies focused on the computational processes of biologically inspired design [7], and the nature of analogies in biologically inspired design [8].

We have been investigating the use of Structure-Behavior-Function (SBF) modeling of complex systems [9-10] to enhance understanding of ecosystems in middle school science education. Empirical research in the SBF conceptual framework suggests that while experts understand a complex system in terms of its interrelated structure, behaviors and functions, novices express primarily its isolated structure, demonstrate minimal understanding of its functions, and largely miss its behaviors [11]. Additional empirical research on the use of SBF models demonstrates deeper understanding as measured by question-answering on pre- post-tests [12].

In this study we examine whether these SBF models may also lead to deeper understanding of complex biological systems among college-level biologists and engineering students in the context of biologically inspired design. Our pilot cognitive study attempts to answer the following questions: (1) Do SBF models provide any inferential capability beyond that provided by text and diagrams? (2) If so, how does the capability vary by the type of inference task, e.g. fact finding or spatial inference?

II. STRUCTURE-BEHAVIOR-FUNCTION (SBF) MODELS OF COMPLEX SYSTEMS

SBF models of complex systems originate in Chandarsekaran's functional representation scheme [13]. An SBF model of a complex system explicitly represents its structure [S] (i.e., its configuration of components and connections), its functions [F] (i.e., its intended output behaviors), and its behaviors [B] (i.e. its internal causal processes that compose the functions of the components into the functions of the system). The SBF language provides a vocabulary for expressing and organizing knowledge in an F → B → F → B … → F(S) hierarchy, which captures functionality and causality at multiple levels of aggregation and abstraction.

In Figure 1 we illustrate an SBF model of the self-cleaning function of the lotus leaf. The lotus leaf is interesting to engineers and others because it maintains a clean surface, despite being in otherwise dirty environments. It does this through nano-structures on the surface of the leaf that interact with water to cause it to bead up and roll off the leaf, carrying debris particles away with it.

System states are represented as shaded boxes, within which are described the components (e.g. contaminants, water droplets), the properties (e.g. location, shape, mass) and values (e.g. on leaf, spherical, or the variable value M). For each state, we include only those components, properties and values relevant to the particular state change that is occurring. The entire series of state changes along with annotations about why the states change constitute the behavior of the system. Connections between states are called transitions, and include a variety of explanation annotations that provide information about why the change occurs. One type of transition, called transition-by-function, gives rise to the hierarchical organization of SBF models as we demonstrate in the following model.

The Self-Clean function (Figure 1a, on the left in Figure 1) of the lotus leaf is the result of a Self-Cleaning behavior consisting of four states. In the first state, contaminants are at rest on the lotus leaf. In the second state, when a drop of water falls on the surface of the leaf, the leaf exhibits a super-hydrophobic effect, which causes the water droplet to take the shape of a sphere. Figure 1b (on the top right in Figure 1) illustrates the super-hydrophobic sub-function; note the by-function annotation on the first transition in the Self-Cleaning behavior. The arrow between the states is the transition, while the annotation on the arrow is the explanation. These annotations provide causal explanations for why state changes occur in the system. The by-function annotation includes a pointer to a function that is represented by another SBF model. In this way SBF models inherently provide function /sub-function decomposition. In the third state, after the water drop falls on the surface of the leaf, the drop rolls over the contaminants using the principle of motion of a spherical body on an inclined plane, subject to the structural constraint that the leaf is inclined and not horizontal. Figure 1c (on bottom right of Figure 1) illustrates this sub-function. In the fourth state, the drop of water rolls off the leaf, carrying the contaminants with them and leaving the leaf clean.

The Cause Superhydrophobic Effect sub-function of the leaf (illustrated in Figure 1b) has its associated behavior which is enabled by the nano-scale “bumps”, which is structural constraint present on the surface of the leaf, by the principle of interacting surface tensions captured by Young’s equation, and by the sub-function (not detailed in this model) of the nano-bumps of making the surface non-wettable. The Make Water Droplet Roll function of the leaf (illustrated in Figure 1b) too has its own causal behavior. When the water moves over the contaminants, it absorbs them subject to the constraint that the force of absorption is greater than the static forces between the contaminants and the surface of the leaf. Note that the SBF model enables access to the physical laws and mathematical equations.

The lotus leaf model presented in Figure 1 provides a representative example of an SBF model. This characterizes only one way of visualizing an SBF model. The model itself could be expressed in plain text or even computer code (LISP!) formats, although such a format may be difficult for a human to parse. The essential content of the SBF model is the emphasis on function (the intended end state) and behavior (intermediate states and explanations).

It is important to recognize that these models are qualitative. They do not provide mathematical descriptions of a system per se, but rather to capture a conceptual understanding of how a system works. Because of their flexibility, it is not uncommon to see many differences between models developed by different individuals.

III. STUDY METHOD

A. *Study context and participants*

This study was conducted as a classroom exercise for a group of 37 undergraduates enrolled in a biologically inspired design class at Georgia Institute of Technology. Of the 37 participating students, 16 self-identified as biologists and 21 as engineers. The participants were all junior and senior level undergraduates familiarized with the concept of biologically inspired design through four weeks of classroom training.

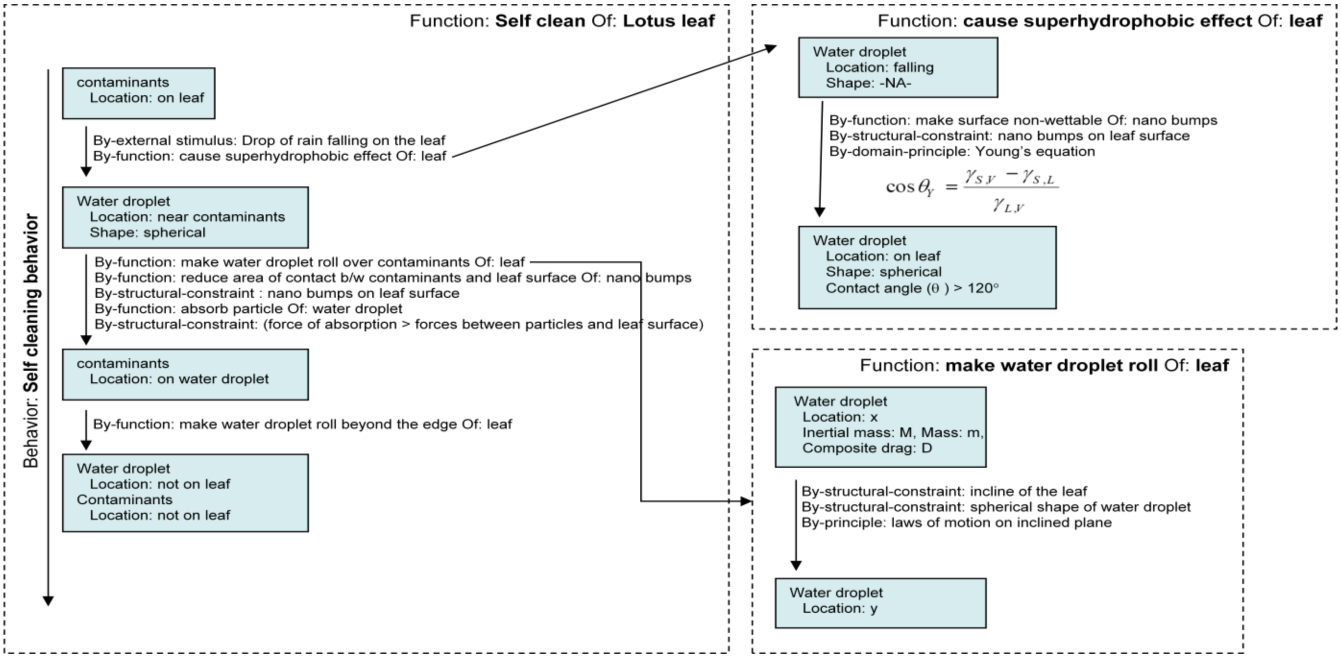

Figure 1: SBF model of the self-cleaning function of the lotus leaf.

The classroom exercise had both research and pedagogical goals. As a pedagogical device, the exercise served to (1) educate students on biological systems that might be useful to their design projects, (2) familiarize students with differences in inferential capabilities afforded by different representations, and (3) help students recognize patterns in communication and representation preferences among the different disciplines represented in the class. The pedagogical goals were realized both by participation in the exercise and by a reflective post-exercise discussion conducted after the exercise. The pedagogical goals and context served as incentive for the students to participate fully in the exercise.

One week prior to the exercise, the students received 90 minutes of classroom instruction in SBF models. In addition to the pedagogical benefits, this ensured that students were familiar with the SBF models presented during the study. Furthermore, a five minute primer was provided to the students prior to the exercise, explaining the state representation schema for SBF models used in the models provided.

The cover page of each packet asked students to self-report on whether their major was biology or engineering, and how familiar they were with respect to four concepts: the lotus leaf, the lotus effect, the basilisk lizard itself, and the basilisk lizard's water walking ability. Students were instructed to score their familiarity with each concept on a scale from 1 to 5, where one is unfamiliar, and five is familiar.

B. *Study methodology*

Students were provided one of three different modalities of representations of a single biological system, and asked to answer questions about the system along four dimensions:

a) fact finding, the ability to find and return a single fact;

b) spatial inference, the ability to reason about or recall the shape or metric relationships among components;

c) complex reasoning, the ability to reason about causal and functional relationships among various components and interactions; and

d) abstract problem solving, the ability to answer complex questions related to the systems behaviors, but that were not explicitly present in the representation(s).

The treatments for each model were (1) text only, (2) text plus graphical and tabular representations, and (3) text plus SBF models (see Figure 1). The students had fifteen minutes to review the new information and answer the questions with a five minute period offered at the end for students who were not yet finished. The exercise was conducted twice, for two different biological systems, the lotus leaf and the basilisk lizard. These two systems were selected because each was often cited by instructors in previous instances of the class.

Table 1. Self-reported familiarity scores

	1	**2**	**3**	**4**	**5**
Lizard	22	23	15	9	5
Lotus	37	22	12	3	0

For the basilisk lizard, seven questions were asked: two fact finding, two spatial reasoning, two complex reasoning, and one abstract problem solving question. For the lotus leaf, five questions were asked: one fact finding, one spatial, two complex, and one abstract.

Students that finished the first exercise early were instructed to close their packets, and not to look ahead to the second exercise. All students finished both exercises within the allotted time.

Exercise packets were arranged so that each student received two different modalities. If a student had text-only modality for the basilisk lizard, they received either text-plus-graphics or text-plus-structured-representation for the lotus leaf. Pedagogically this enabled student reflection on differences in their own experience with the different modalities. During the first round of exercises, some students did not look ahead in their packets at all, and were unaware that they were given more than just the text representation. When students alerted facilitators to this fact post-test, the test facilitators asked that anyone who was unaware of the graphical representation during the exercise record this fact on their answer sheet. All answer sheets thus noted were considered text-only in terms of the analysis. As a result, overall 20 students received diagrammatic representations, 24 received SBF representations, and 30 received text-only.

At the end of the exercise on the last page of the packet the following question was asked: "In each case you were provided with different representations (either text with SBF, text with graphs/tables, or text only.) Which representations did you prefer? Why?" Students were allowed as much time as required to answer this question.

C. *Materials used*

Text descriptions of the systems were extracted from papers describing the relevant details of their respective systems [14, 15]. The original papers were technical and difficult to quickly read, and so were paraphrased to a level appropriate for undergraduates. No mathematical formulae were present in the text descriptions.

We used SBF models that explicitly stated the relationships between states and state properties (see Figure 1). The SBF models used were prepared earlier by the authors as sample SBF models for demonstration purposes. Graphical annotations present in these original SBF models were removed. All other content of the SBF models were preserved.

Graphic representations were taken either directly from the corresponding academic papers [14, 15], or from diagrams previously developed in our lab for use in augmenting SBF models.

Each student was asked the same set of questions for each system. Following are a list of sample questions for both (a) basilisk lizard and (b) lotus leaf:

Fact finding

(a) Which provides more lift, the slap phase or the stroke phase of the basilisk lizard's movement?

(b) What physical properties of the lotus leaf account for it being clean?

Spatial Inference
(a) In which phase, slap or stroke, does the moving leg cover a greater total distance?
(b) What shape does the water droplet form on the leaf of a lotus leaf?

Complex Inference
(a) Which provides more thrust, the slap phase or the stroke phase of the basilisk lizard? Why?
(b) How does the water droplet move on the lotus leaf?

Abstract Inference
(a) How could you estimate the thrust and lift generated by the basilisk lizard, without measuring anything about the lizard itself?
(b) How is this different from how water might move over a surface without the properties of the lotus leaf?

D. *Grading method*

Answers were graded by the first author of this paper (Helms), a computer scientist, with neither biology nor engineering training. His knowledge of both the lotus leaf and basilisk lizard systems is derived from scientific research articles, developing SBF models of the systems, observing the biologically inspired design class, and from discussions with biology and engineering instructors in the class.

The correct answers to fact finding and spatial inference questions were unambiguous. The answers to complex questions, and abstract inference questions were subject to some interpretation, as discussed in the following section.

IV. DATA

The self-reported familiarity scores are presented in Table 1. The self-reported mean familiarity for the basilisk lizard system was 1.74, for the lotus leaf, 2.35.

Answers to questions were categorized as either correct or incorrect. For complex and abstract questions, some unanticipated answers were received. For instance, when asking how the lotus effect is accomplished, a student might cite the (anticipated) underlying property accounting for the behavior (for instance hydrophobicity), or might describe the motion of the drop of water as it rolls down the leaf and pick up particles. Both are legitimate correct answers to the question. For such questions, any rational answer citing facts and following a logical account were coded as correct. For the purposes of this study only the correctness of each answer was analyzed. Only obviously wrong answers were coded as wrong. For example, for the complex question "How does the water droplet move on the lotus leaf?" the answer "by spreading" was considered incorrect because it is the opposite of the correct answer (the water maintains a spherical shape and specifically does not spread.) Non-answers (blanks), accounted for 4.7% of the total answers, and were provided a unique code but were considered incorrect for purposes of the analysis.

For the basilisk lizard based questions, table 2 shows the percentage of correct answers for each question, by treatment type; table 3 provides the same information for the lotus leaf based questions.

With respect to the final question, preferred representation, interestingly some students felt strongly enough to not only comment on their preferences, but also to comment on their dislike for the SBF modality. Table 4 summarizes student preference by major, where the row heading Not SBF represents the number of students that reported a dislike for the SBF modality.

V. ANALYSIS

A. *Familiarity Scores*

Although the mean reported familiarity with lotus was greater than that for the basilisk lizard, and scores were generally higher for the lotus questions than for the basilisk lizard questions, correlation analysis between the self-reported understanding of a system and the number of correct answers show close to zero correlation (r-squared = .015 for basilisk, r-squared = .047 for lotus). Thus, self-reported prior knowledge of a system does not appear to be an important factor for this study. This is likely a result of the level of detail of the questions being asked relative to a student's perception of their own familiarity.

B. *Question Scores*

The mean score for the basilisk lizard was 4.27 out of 7 (61%), with a standard deviation of 0.87 (12.4%), while the mean score per student for the lotus leaf was a 3.7 out of 5 (74%), with a standard deviation of 0.66 (13.2%).

When assessing the significance of including SBF and diagrammatic modalities, we test the hypothesis that the proportion of questions answered with SBF or diagrams is greater or less than the proportion answered for the base rate for text only for the same question, assuming standard normal distribution. We note that for the basilisk lizard questions, the number of students n = 17 for text only, n = 11 for text plus diagrams and n = 9 for text plus SBF. Diagram plus text results are statistically different at a confidence interval of .01 for complex 2 (z = 2.68), and are statistically significant at a confidence interval of .10 for spatial 1 (z =

Table 2. Percentage correct for Basilisk Lizard questions, by treatment type

	FF1	**FF2**	**Sp1**	**Sp2**	**Cpx1**	**Cpx2**	**Ab1**
Text	94.1	88.2	82.4	58.8	35.3	17.7	35.3
Diagram	100	66.7	100	88.9	66.7	66.7	55.6
SBF	100	90.9	81.8	45.6	90.9	63.6	81.8

Table 3. Percentage correct for Lotus Leaf question, by treatment type

	FF1	**Sp1**	**Cpx1**	**Cpx2**	**Ab1**
Text	100	92.3	84.6	92.3	69.2
Diagram	100	72.7	90.9	90.9	54.5
SBF	92.3	84.6	92.3	84.6	76.9

1.34), spatial 2 (z = 1.54), and complex 1 (1.56). SBF + Text findings are significant at the .01 level for complex 1(z = 2.88), complex 2(z = 2.68) and abstract 1(z = 2.41) questions. For the lotus example, no significant differences were detected for any of the questions.

Likewise tests of significance between number of correct answers for each question were run between engineers and biologists. Statistically significant differences were detected between engineers and biologists for the complex 2 question for the basilisk lizard (z = 1.34) and for the abstract 1 question for the lotus (z = 2.55).

While not statistically significant overall, it is interesting and counterintuitive that for some questions, the additional graphical or functional information resulted in worse average performance. This can be seen in fact finding question 2 for the basilisk model, and for spatial question 1, and abstract question 1 for the lotus leaf model.

VI. CONCLUSIONS

There is a growing need to determine what kinds of representations of biological systems facilitate understanding. In this paper we described a pilot study to determine whether SBF models enable deeper understanding of complex biological systems. We draw three preliminary conclusions from the study. First, for some cases SBF models do enable more accurate inferences about biological systems for complex and abstract questions. When the inference tasks require knowledge about causality, for instance understanding the locomotion of the basilisk lizard on water, SBF models provided a deeper understanding than textual or diagrammatic representations. Second, no single representation is best for all different types of inferences. For spatial inferences, diagrammatic representations appear to be better than SBF models. This leads to our final conjecture: for supporting the understanding of biological systems in the context of biologically inspired design, it may be best to provide access to multiple external representations, including text, diagrams, and SBF models.

It is important to note that this paper describes a pilot study that is limited in many ways. The in situ study was conducted in a real classroom. Studies of this kind do not easily allow formal controlled experiments that isolate independent and dependent variables. Further, the study was conducted using pencil and paper. We have since built an interactive tool called DANE (for Design by Analogy to Nature Engine) for supporting biologically inspired design, using on SBF, text and graph representations. We have also implemented DANE in the classroom, and conducted a preliminary study in our laboratory to evaluate whether DANE enhances understanding of complex biological systems, and [16].

ACKNOWLEDGMENT

We thank the instructors of ME/ISyE/MSE/PTFe/BIOL 4740 in Fall 2007 as well as the student designers in the class for their input in this study. This research was supported by an NSF CreativeIT grant (#0855916) entitled "Computational Tools for Enhancing Creativity in Biologically Inspired Engineering Design."

REFERENCES

[1] Benyus, J., 1997, Biomimicry: Innovation Inspired by Nature. New York: William Morrow.

[2] Bonser, R., & Vincent, J., 2006, "Technology Trajectories, Innovation, and the Growth of Biomimetics," Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, pp. 1177-1180.

[3] Vincent, J., Bogatyreva, O., Bogatyrev, N., Bowyer, A., & Pahl, A., 2006, "Biomimetics: Its Practice and Theory", Journal of the Royal Society, Interface 3, pp. 471-482.

[4] Yen, J., & Weissburg, M., 2007, "Perspectives on Biologically Inspired Design: Introduction to the Collected Contributions," Journal of Bioinspiration and Biomimetics.

[5] Linsey, J.S., Wood, K.L., Markman, A.B., 2008, "Modality and Representation in Analogy," Artificial Intelligence for Engineering, Design, and Manufacturing, 22, pp. 85-100.

[6] Mak, T.W., Shu, L.H., 2008, "Using Descriptions of Biological Phenomena for Idea Generation," Research in Engineering Design, 19/1, pp. 21-28.

[7] Helms, M., Vattam, S., & Goel, A., 2009, "Biologically Inspired Design: Process and Products," Design Studies 30(5): pp. 606-622.

[8] Vattam, S., Helms, M., & Goel, A., 2010, A Content Account of Creative Analogies in Biologically Inspired Design. *AI for Engineering Design, Analysis and Manufacturing, Special Issue on Biologically Inspired Design* , 24: 467-481

[9] Goel, A. K., Gomez de Silva Garza, A., Grué, N., Murdock, J. W., Recker, M., & Govinderaj, T., 1996, "Towards designing learning environments -I: Exploring how devices work," In Proc. Third International Conference on Intelligent Tutoring Systems: Lecture notes in computer science 1086, Springer.

[10] Goel, A., Rugaber, S., & Vattam, S., 2009, "Structure, Behavior & Function of Complex Systems: The Structure, Behavior, Function Modeling Language," International Journal of AI for Engineering Design, Analysis and Manufacturing, Special Issue on Developing and Using Engineering Ontologies, 23, pp. 23-35.

[11] Hmelo-Silver, C., Jordan, R., Demeter, M., Gray, S., Liu, L., Vattam, S., Rugaber, S. & Goel, A., 2008, "Focusing on Function: Thinking Below the Surface of Complex Natural Systems," Science Scope, Summer 2008, pp. 27-35.

[12] Vattam, S., Goel, A., Rugaber, S., Hmelo-Silver, C., Jordan, R., Gray, S., & Sinha, S., 2010, "Understanding Complex Natural Systems by Articulating Structure-Behavior-Function Models," Educational Technology & Society, Special Issue on Creative Design, 14(1): 166-181.

[13] Chandrasekaran, B., 1994, "Functional Representation: A Brief Historical Perspective," Applied Artificial Intelligence, 8(2), pp. 173-197.

[14] Barthlott, W., & Neinhuis, C., 1996, "Purity of the Sacred Lotus, or Escape From Contamination in Biological Surfaces," Planta, 202, pp. 1-8.

[15] Hsieh, S.T. & Lauder, G.V., 2004, "Running on Water: Three-dimensional Force Generation by Basilisk Lizards," PNAS 101(48): pp. 16784-16788.

[16] Vattam, S. Wiltgen, B., Helms, M., Goel, A, & Yen, J., 2010, "DANE: Fostering Creativity in and through Biologically Inspired Design," In Proc. First International Conference on Design Creativity, Kobe, Japan, pp. 127-132.