# Making mechanistic sense: are we teaching students what they need to know?

**Michael W. Klymkowsky**
Molecular, Cellular & Developmental Biology,
University of Colorado Boulder, Boulder, CO. 80309

**Summary**: Evaluating learning outcomes depends upon objective and actionable measures of what students know – that is, what can they do with what they have learned. In the context of a developmental biology course, a capstone of many molecular biology degree programs, I asked students to predict the behaviors of temporal and spatial signaling gradients. Their responses led me to consider an alternative to conventional assessments, namely a process in which students are asked to build and apply plausible explanatory mechanistic models ("PEMMs"). A salient point is not whether students' models are correct, but whether they "work" in a manner consistent with underlying scientific principles. Analyzing such models can reveal the extent to which students recognize and accurately apply relevant ideas. An emphasis on model building, analysis and revision, an authentic scientific practice, can be expected to have transformative effects on course and curricular design as well as on student engagement and learning outcomes.

## Introduction

> *If the book is to remain manageable in size, it is inevitable that some favorite topics of the reader might be glossed over. However, despite the admirable emphasis on principles and concepts, I occasionally felt shortchanged. With these authors, I might have expected a deeper treatment of what surely must be one of the most important principles: the existence of threshold responses to morphogens – molecules that diffuse from a source and set up a graded concentration. Instead, when it gets to the nitty gritty of boundaries, the activating and repressing activities of one gradient (such as hunchback or dorsal) are mentioned, but how one protein both activates and represses is not explained.*
> – Richard Harland

Biology education has long struggled with the impression that it is more about memorization than the thoughtful consideration and application of widely relevant ideas (Lewin, 1982). Many teachers and academics often see and present areas of biology as distinct disciplines (Nehm, 2019; Nehm et al., 2009). The result is that some introductory biology courses read as surveys of implicitly unrelated topics, rather than a consideration of evolutionary processes acting on systems of interacting entities, whether molecules, cells, or organisms (Klymkowsky, 2010). To determine whether teaching is effective, that is, whether learning has occurred, there is often an emphasis on whether students know or can recognize the correct answer, rather than whether they can explain, in mechanistic terms, biological processes. Failing to emphasize the need to think about biological systems in a critical and mechanistic perspective contributes to a Dunning-Kruger effect (unwarranted confidence in one's understanding) in students, citizens, and politicians called on to make decisions related to a range of biomedical subjects (e.g., vaccine safety and the efficacy of homeopathic, naturopathic, and ineffective (and ecologically destructive) "folk" remedies). The ability to recognize correct answers in the context of a multiple-choice test is quite different from the ability to construct a relevant, plausible, and verifiable explanation, to justify the assumptions that support it, and to recognize the predictions that it implies. Various reform efforts in science education have focused on the importance of supporting learning within a coherent narrative that engages students, in part through a greater emphasis on research processes and common, cross-disciplinary principles (NRC, 2012).

As discussed by Beatty (1995), biology is distinctly different from chemistry and physics. From both a practical and a theoretical perspective, there are no "laws" in biology. While constrained by the laws and general principles of chemistry and physics, biological systems and processes are contingent upon their evolutionary history. As I will discuss, even closely related organisms can vary from one another in significant mechanistic detail. Typically, explanations of biological systems take two complementary forms: the mechanistic (how) and the evolutionary (why) (Mayr, 1961; 1985). At the cellular and molecular levels, mechanistic explanations address how a process occurs and involve the behaviors of molecular machines, a point made explicitly by Bruce Alberts (1998). The properties of molecules, the thermodynamics governing their interactions with one another, and the various chemical reactions in play, particularly the coupling of favorable reactions to drive unfavorable processes, contribute to and constrain such behaviors. "Why" explanations involve evolutionary adaptations, population behaviors (bottlenecks, founder effects, and genetic drift), and the ecological and environmental considerations driving speciation and species-specific behaviors. Both how and why mechanisms involve emergent behaviors, explicable only in terms of systems of interacting processes. While there are no biologic laws analogous to Newton's laws of motion or the laws of thermodynamics, the outcomes of biological processes reflect various, sometimes contradictory tendencies (see Beatty, 1995). While the complexities of biological systems complicate "how" explanations, the unobservable nature of evolutionary events that occurred in the distant past inherently constrain "why" explanations. When approaching the teaching and learning of biology, there are therefore two related questions: what is important to teach so that students can make sense of and appreciate the limits of our current knowledge, and how do we determine whether students are building an effective understanding, can they use their knowledge critically and constructively? How these questions are to be answered is rarely explicitly discussed within the (developmental) biology education community.

One approach to determining what one, and one's students know can be "borrowed" from working scientists and reflects the Socratic tradition. It is to have students engage in the process of building and defending plausible explanatory mechanistic models (“PEMMs”). Building a PEMM involves asking ourselves what combination of processes can produce the behaviors observed and then responding to questions raised by ourselves and others (fellow students, instructors, or peer reviewers) in order to determine:

(i) whether the model's underlying assumptions are consistent with established chemical and physical principles;
(ii) whether the model produces the expected behavior(s); and
(iii) how well the model predicts the response of the system to various perturbations.

Such model building and testing exercises can reveal misunderstandings about basic processes and their application, as well as the effects of specific details of the biological system under consideration. Equally important, they serve as a way to reveal whether students recognize as relevant underlying principles when constructing a plausible explanatory model. A version of this process occurs within the beSocratic exercises used with the CLUE, OCLUE, and biofundamentals course materials (Cooper and Klymkowsky, 2016; Cooper et al., 2019; Klymkowsky et al., 2016). Such activities can reveal how students understand and apply (or not) core ideas in response to how instructors present materials. As an example, such an analysis revealed the persistence of student confusions about the distinction between hydrogen and covalent bonds in conventionally designed courses (Williams et al., 2015) and the factors that influence the behavior of molecular networks (Trujillo et al., 2012) (see below). The process

of "reviewing" a model can reveal such mistakes and omissions. Does the logic of the model produce the behaviors that the model is meant to explain? Typically predictions of model outcomes are qualitative, but increasingly involves quantitative methods (see Ali et al., 2020; Gardner et al., 2000), raising the question of when and at what level mathematical modeling methods should be integrated into biology courses and curricula (see Klymkowsky, 2009; Pevzner and Shamir, 2009). A final test of a model is the extent to which it can predict the system's behavior in response to various perturbations. As the cycle repeats, there are opportunities for revision, the integration of new ideas, new details, and new components. Given the complexity and our often incomplete understanding of biological systems (Lin et al., 2019; Smits et al., 2019; Wilkinson, 2019), working researchers frequently revise models, and such revisions should be expected for students' explanatory models. What is critical in both arenas is the ability to revise models in response to review.

What is likely to be particularly valuable for students, and for course and curriculum designers and instructors, are the insights that building PEMMs can have in terms of identifying the ideas and practices required for their construction. Such an analysis enables us to look backward at what was presented to students (taught), what they are able to recognize as relevant and apply (learned), and to consider where in the curriculum specific ideas are best presented and practiced, so that reasonable models can be constructed and the habits of mind associated with model building can be reinforced. It is worth noting explicitly, and stressing to students, that even the most careful and experienced of model builders benefit from critical "third party" review. Michael Meister's analysis of paramagnetic effects on biological molecules provides a particularly informative example, revealing that certain "claims conflict with basic laws of physics. The discrepancies are large: from 5 to 10 log units. If the reported phenomena do in fact occur, they must have causes entirely different from the ones proposed by the authors" (Meister, 2016).[1]

The process of PEMM construction, analysis, and revision provides students with an authentic introduction to a basic scientific process and training in the habits of mind involved in deciding for themselves whether biology-based arguments make sense. Whether a student's original PEMM turns out to reflect the actual process is less important than that it works and serves as the basis for testing assumptions and responding to critical feedback. To illustrate their value, I describe my journey to appreciating the value of employing a PEMM evaluation model in the context of designing and teaching an upper division developmental biology course, often the last course majors are required to take before graduation.[2]

**Understanding developmental mechanisms:** The study of developmental biology is commonly rooted in examples from various "model" organisms, chosen for historical and practical reasons. Teaching developmental biology poses interesting challenges, because evolutionary adaptations end up producing species specific, and often functionally significant mechanistic variations. As an example, mouse is a common model system for studying early developmental events in mammals. And yet, there are well known and dramatic differences in a number of early (and late) developmental processes (Rossant, 2015), including basic mechanisms involved in cancer and other diseases (Pulendran and Davis, 2020; Seok et al., 2013; Warren et al., 2013). As one recent example, a highly conserved, long, non-coding RNA appears to play different roles in early mouse and human embryos (Sharma and Carninci, 2020), while null mutations can produce different phenotypes in human and mouse (Liao and

[1] A more humorous expansion of this theme can be found here: Magnetofiction – A Reader's Guide

[2] My ruminations on this topic are described here: https://bioliteracy.blog/on-devo/

Zhang, 2008). New genes arise (Zhang et al., 2015) and well-conserved genes are lost (Sharma et al., 2018; Zhu et al., 2007) in various lineages.

Frog and newt embryos have been used as model systems in a number of classic studies, ranging from establishing the feasibility of reprogramming somatic nuclei (Gurdon, 1962) to uncovering a range of inductive interactions and morphogenic processes (Harland and Grainger, 2011; Moriyama and De Robertis, 2018; Shook et al., 2018). Yet, there are substantial differences in developmental processes between amphibian species. For example, the maternal mRNA VegT (involved in the regulation of Nodal signaling) is localized to the oocyte's vegetal cortex in the clawed frog *Xenopus laevis*, a common model system for studies of vertebrate development, but found in the animal region of the marsupial frog *Eleutherodactylus coqui* oocytes (Elinson and del Pino, 2012). Of note, orthologs of VegT are absent from mammals. Similarly, the asymmetric distribution of bicoid protein in the fruit fly *Drosophila melanogaster* egg is often used to introduce how asymmetries in gene expression are established in the early embryo, yet the bicoid gene is "unique to higher dipterans" – it is absent from other insects (Lynch and Desplan, 2003). There are dramatic differences in the functional organization of HOX genes, involved in anterior-posterior (and other) embryonic asymmetries, between species (Darbellay et al., 2019; Duboule, 2007). So, if we are to follow the advice of Bill Wood (2008) to teach concepts and not (often species-specific) facts, on which concepts should we focus? Here the individual instructor is often provided little guidance and may come to rely on increasingly encyclopedic textbooks.

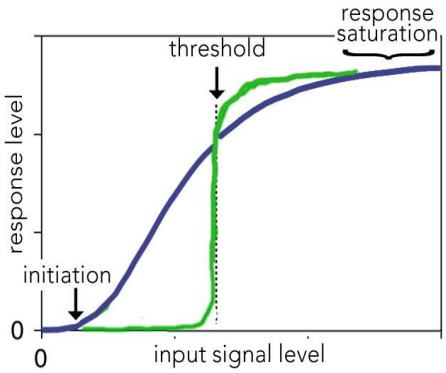


FIG.1: Examples of a standard sigmoidal dose-response curve (blue) indicating the point of response initiation and saturation, together with a threshold response (green) in which signal inputs for initiation and saturation levels are close to one another

**Threshold responses – a repeated theme in developing systems:** A key feature of developing systems, whether uni- or multicellular, is that they change over time in response to various signals.[3] A system's responses to these signals are generally not linear but display a distinct sigmoidal, and in the extreme case, a sharp "threshold" shape (Chow et al., 2011)(**FIG. 1**→). In the case of the unicellular slime mold *Dictyostelium discoideum,* the cellular response involves two distinct threshold signals, a "quorum sensing" system that monitors the number of cells per volume and a second system that reflects the cells' nutritional state (Loomis, 2014). Below a distinct signal molecule concentration, there is little or no cellular response. At a slightly higher signal concentration, that is above a "threshold," the signal response increases sharply and quickly saturates.

Student responses to questions related to signaling systems, delivered through the web-based beSocratic system (Bryfczynski et al., 2015), displayed some evident confusions associated with response onset, saturation, and threshold effects (**FIG. 2** ↓). Moreover, when asked to "provide a plausible molecular mechanism to produce that behavior (a threshold effect)," most students invoke active mechanisms associated with feedback loops, protein structure, nuclear import, or DNA modification – few explicitly recognized the need to overcome homeostatic processes (unpub. obs.). In fact, we had

[3] https://bioliteracy.blog/2018/12/15/on-teaching-developmental-biology-in-the-21st-century/

previously described similar confusions displayed by late stage molecular biology majors when consid-

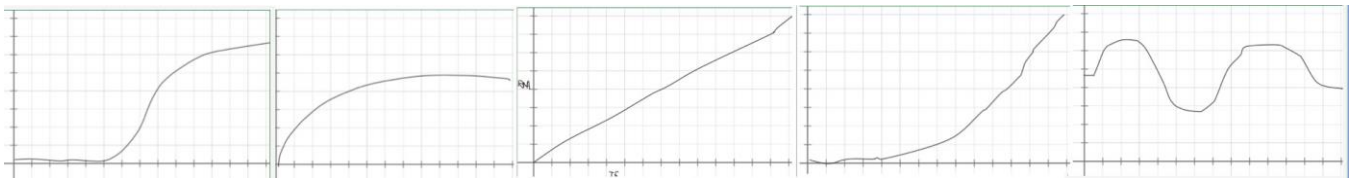


*FIG. 2 A representative set of student responses to the question, "In terms of increasing concentration of transcription factor (TF)(x-axis), draw your prediction of the expression level (RNA) of a particular target gene)."*

ering molecular networks (Trujillo et al., 2012). It appears that many students come to late stage "capstone" courses with a fragile understanding of the molecular processes involved in common signal-response behaviors. It follows that if students are to make sense of developmental processes, we must re-design instructional approaches so that they address these persistent difficulties. In the past, a common approach has been to identify students' problematic ideas and then design instruction to “overcome” or replace them with more canonical ideas. Unfortunately, there is little evidence that this approach works, beyond replacing misremembered facts. When considering the interplay of complex ideas, it seems entirely inadequate. We must address how to help students construct, connect, and contextualize their knowledge so that it becomes useful. I suggest that a strategy based on the construction, analysis, and revision of PEMMs is one such approach.

**What do we need to consider when building a signal response PEMM?** When getting specific about building PEMMs to explain sigmoidal responses and threshold behaviors, we first need to define exactly what we mean by the response. There are a number of possibilities for the student to consider – is it the immediate effect that follows the binding of the signaling molecule to its receptor, or is it the end behavior. In the case of slime mold cells, it is an observable behavior - the cells' migration toward each other to form a slug that goes on to differentiate. It can be the appearance of a particular pattern, such as the distinctive segments of a *Drosophila* embryo, or the patterns of gene expression that occur along the dorsal-ventral axes of the larvae or the vertebrate neural tube. Next our model needs to explain (i) why there is little or no response below a signal concentration, and (ii) why the response rises and then saturates as the signal concentration increases above its threshold concentration.

When originally faced with teaching these behaviors, I looked to the literature for established and concrete mechanisms to scaffold my presentation, rather than working through the possible (plausible) generic processes that might be involved. I found this search for accessible (teachable) mechanisms frustrating – often the mechanism(s) underlying "delayed" signal initiation was not clearly elucidated, if described or considered at all (see the text box above – Harland, 2011). While not described explicitly, mechanisms of saturation seemed inherently simpler (but not always obvious in student explanations); typically they involve limited numbers of regulatory targets – for example, there are generally only two copies of a particular gene per cell. Based on this view, although the mechanism(s) that lead to a higher initiation concentration may be complex, a threshold effect is a "simple" variant of a sigmoidal response, a response associated with a small difference between response onset and saturation.

So, what factors can influence the response initiation concentration? There are many, ranging from the concentration of receptors, signal-receptor binding affinity and "dwell" time (the half-life of the bound state), the effect of binding on receptor behavior, which can include receptor interactions with

other molecules and allosteric effects on enzymatic activity leading to the post-translational modification of targets that alter interactions, cellular localization, and rates of degradation. Such "downstream" and feedback effects can alter the numbers of available receptors, their response to signals, and to changes in gene expression, which in turn can influence the system's responsiveness through the expression of a range of agonists, antagonists, and modifiers. These changes do not occur in isolation; rather, they occur in the context of various homeostatic processes that act to return the system to its state before exposure to the signal. If the signal response alters gene expression, the state after a signaling response may well be different and so respond differently to the same signal molecule. The result is that the system is a product of its history, together with its energetic state, as these responses all involve coupled chemical reactions. Signal response systems are information processing molecular machines, in the sense described by Alberts (1998).

**Modeling response initiation**: To build a simple generic model of a sigmoidal response, a student will have to consider the various factors described above and the opposing homeostatic processes that act to reverse the effects of signal molecule-receptor "activation." At low signal concentrations, when few receptors are bound, homeostatic mechanisms will oppose the resulting signal-induced activity, e.g., targeted phosphorylation will be opposed by dephosphorylation. The response to a particular signal concentration therefore reflects both signaling effects and their "reversal rates." As an example, consider the case of a signal associated with the opening of an ion channel. The effect of opening a small number of channels will be offset by the on-going restorative, energy-dependent pumping mechanisms that act to maintain the cell's resting state. Only with increasing signal, which activates and opens more receptors, will the maintenance system be overwhelmed, and a response generated. The situation is further complicated when the initial steps in generating a response involve the assembly of a multiple component complex. The frequency and stability (lifetime) of the steps involved in this process will impact the probability that a functional and stable complex will be formed. That said, from a modeling perspective, the steps involved in such processes can often be "collapsed or telescoped" into a single step (*see* Chow et al., 2011). Together, these factors will determine the signal concentration at which a discernible response will occur.

While students may be expected to generate generic response models, they will require more detailed interaction diagrams when asked to consider "real" systems (as illustrated below). For example, the presence of antagonists can lead to an effective reduction in signal or receptor concentrations over time, while accessory factors can significantly increase dwell times compared to simpler systems, as suggested by observations on transcription factor binding (Gurdon et al., 2020). In this light, it appears that students' general understanding of the energetics of molecular interactions within biological systems is often weak (Cooper and Klymkowsky, 2013; Kohn et al., 2018). When considering response saturation, we need to recognize that the numbers of targets that can be activated, whether receptors, molecular machines, or genes, are limited. Once all targets are activated, the response will necessarily plateau. For students to be able to apply this idea within their models, including models of gene expression, teachers must lay the necessary groundwork in earlier courses.

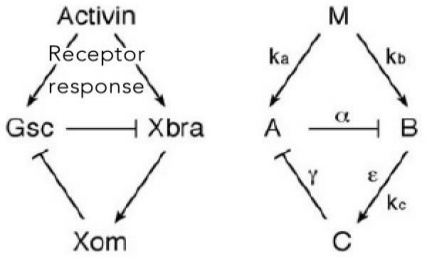



**FIG.3**: A schematic of regulatory circuit described by Saka & Smith (2007)

How can we present these complexities to students so that they are not overwhelmed? After some searching I was drawn to a signaling response model presented by Saka and Smith (2007). They aimed to explain how different levels of an extracellular signaling molecule (activin, a member of the TGFβ family) could differentially regulate the expression of one or another target gene, a system active during *Xenopus* development (**FIG. 3** →). What is particularly noteworthy about this system is its apparent simplicity and that it can produce two opposing outcomes depending on the assumptions made. The result is an accessible scenario to introduce the various considerations involved and how they impact signal-response outcomes and illustrate the range of behaviors that can be generated by "simple" systems. Similar considerations apply to a wide range of signaling systems and can be extended to considering immediate, steady state, adaptive, and cascading (evolving) responses (Lemmon et al., 2016; Li and Elowitz, 2019) as well as more complex fold-change sensing systems (Adler and Alon, 2018; Goentoro and Kirschner, 2009).

In the Saka and Smith model, secreted activin protein and the responding receptor system trigger an "upstream" process. The response (modified from Chaikuad and Bullock, 2016)(**FIG. 4** →) involves binding, changes in receptor structure, and interactions with other regulatory factors that lead to receptor kinase activation and the modification of cytoplasmic receptor-regulated SMAD proteins (R-SMADs). These processes are all reversible through various mechanisms. Phosphorylated R-SMADs dimerize in the cytoplasm and associate with the "co-SMAD," SMAD4. The cytoplasmic SMAD4:R-SMAD complex is then imported into the nucleus where it binds to specific DNA sequences and interacts with various accessory proteins leading to altered gene expression.

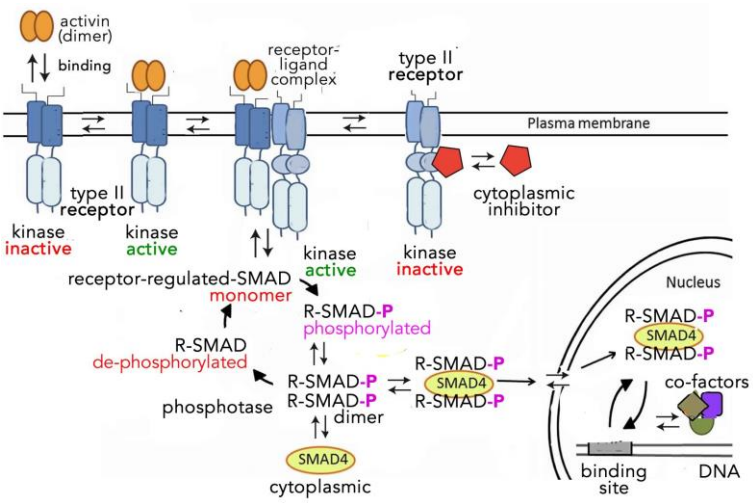



**FIG.4**: A schematic of the activin (TGFb) signaling pathway adopted with modifications from Chaikuad & Bullock,

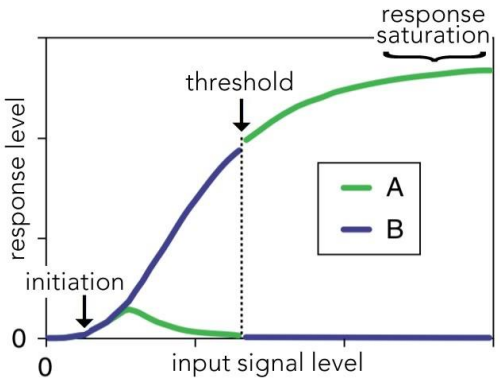



**FIG.5**: Outcome of the relationship between input signal (activin) and gene expression based on one set of parameters - modified from Saka & Smith (2007)

In the Saka and Smith model (←**FIG 5**

described in more detail in Appendix 1) exposure to activin leads to regulation of the expression of two genes, *Gsc* and *Xbra.*[4] Both encode sequence-specific DNA binding proteins and act as regulators of transcription. In their scenario, activin-activated R-SMAD/SMAD4 complexes directly regulate both *Gsc* and *Xbra*; there are no intervening genes whose transcription and translation are necessary for activin-regulated *Xbra* and *Gsc* gene expression. The gene products that are needed (and there are many) are already present within the cells. The network does, however, involve indirect effects: Gsc acts to inhibit *Xbra* expression and Xbra acts to induce *Xom* expression, with the Xom protein acting to repress *Gsc* expression – all other regulatory targets of Gsc, Xbra, and Xom are ignored. The model predicts dramatically different behaviors in terms of gene expression, based on assumptions about rates of target protein accumulation, a function of synthesis and degradation rates, together with target gene binding affinities.

Reproducing the Saka & Smith model requires a level of mathematical sophistication that most undergraduate biology students are unlikely to possess even though many biology degree programs have (or had) a calculus course requirement. That said, after an introduction to the system, students can build on general assumptions and develop qualitative models that produce clear mechanistic predictions. We can reasonably ask students to justify predictions as to how variations in various parameters (e.g. differences in protein stability, binding affinities, cytoplasmic localization, as well as various forms of feedback interactions) will influence network behaviors over time and space. As an example, we can expect students to be able to predict the effects of local signal sources, resulting in signaling (morphogen) gradients, as well as the effects of changes in gene and transcript size, as described by Harima et al (2013). An example question is supplied in Appendix 2.

**How do we prepare for the effects of a PEMM-centric approach?** The point of the PEMM approach is to focus instruction and to prepare students so that they can analyze processes they encounter in various biological contexts, ranging from the molecular to the ecological. As pointed out by McClymer & Knowles (1992), students' preparation needs to include information that every practitioner knows. We must therefore think hard about what that information consists of – is it details of specific systems or general principles such as how molecules interact and how those interactions influence their various activities? I would argue that general information should be defined not, for example, as the details of the Krebs cycle but instead as what principles are involved in coupling chemical reactions. Instructors can then introduce details as needed so students can consider specific processes, including whether they are reversible or not, and if not (e.g. proteolytic processing) how the system resets over time.

The question is whether the information we ask students to remember is useful to them in understanding and explaining a range of processes. In the case of sigmoidal response curves and threshold behaviors discussed here, we are working to develop a general understanding of a ubiquitous feature of a wide range of biological processes – from quorum sensing in microbial communities, the patterning of embryonic development, as well as (arguably) a range of physiological and social processes. The desired result is that students will be able to generate plausible explanatory models for how various regulators (extracellular and intracellular antagonists) and perturbations (mutations, toxic molecules, and environmental stressors) influence a system.

[4] Gene names are italicized, protein names are not.

So how does incorporating a PEMM-centered approach to presenting materials impact course and curricular design? The impact is likely to be dramatic, depending upon the extent of its implementation. In part, the effect reflects the need to introduce students' to systems thinking, model building, and the process of evaluating the implications of their assumptions. Through such a process, students gain direct experience with authentic scientific practices without adding an excessive rote memorization load. To prepare students, they need to be reminded of, or in some cases introduced to for the first time, the various cellular and molecular processes involved and how they interact in specific situations. This includes calling out the relevance of underlying (and universal) processes and necessary details. Practice in model building, presentation, analysis, and revision takes class time. In my own situation, materials are introduced in one class period; these are applied in the context of beSocratic activities that students' complete on their own. At the start of the next class period we review these activities and students are asked to present their solutions.[5]

Adoption of a PEMM approach encourages us to critically evaluate what has been "covered" in past courses and to focus on materials that students will need to use to construct and evaluate models of developmental processes. It means that we must consider what topics can be omitted or de-emphasized in order to make room for such explanatory model building, feedback, and revision. Given that instructors may worry that others will criticize course changes as an inappropriate over-simplification of course content (a criticism leveled at the CLUE chemistry curricula), it is important to have established benchmarks to evaluate student learning outcomes.[6] [line deleted at the request of the reviewer]

Model building invites the incorporation of versions of peer review, revision, and resubmission, and helps us move from a one-off evaluative system to a more developmental process focused on applying and mastering underlying concepts and their application. We change the emphasis of assessment from the binary of right-wrong to the competency of learning how to perform a task. The idea that biological systems all share common features becomes the theme that unites them into a coherent and comprehensible whole – it provides a context within which to explicitly recognize and incorporate the details of specific systems as needed. Such an altered emphasis does, of necessity, demand a change of "coverage" and resource allocation, particularly in terms of instructor feedback and how we evaluate students learning outcomes. I would suggest moving from the use of timed high stakes exams, something no scientist would willingly accept (imagine if manuscripts or grant proposals had to be developed in a timed context), to one more like that of the preparation, peer review, revision, and eventual "publication" of a course dossier that documents a student's mastery of the materials and skills presented. Centering instruction and assessment, formative and summative on PEMMs and their presentation and revision would change course emphasis, hopefully making courses more engaging, inclusive, and effective in terms of learning outcomes and in fostering an inquisitive mindset. The PEMM approach explicitly values reflective thinking, something too often in short supply in a range of educational and broader social settings, .

[5] The course website can be visited here: http://virtuallaboratory.colorado.edu/DEVO@CU/index.html and will be archived.

[6] In the case of CLUE nationally normed exams generated by the American Chemical Society's Exams Institute to demonstrate no decrease in scores on these very traditional exams.

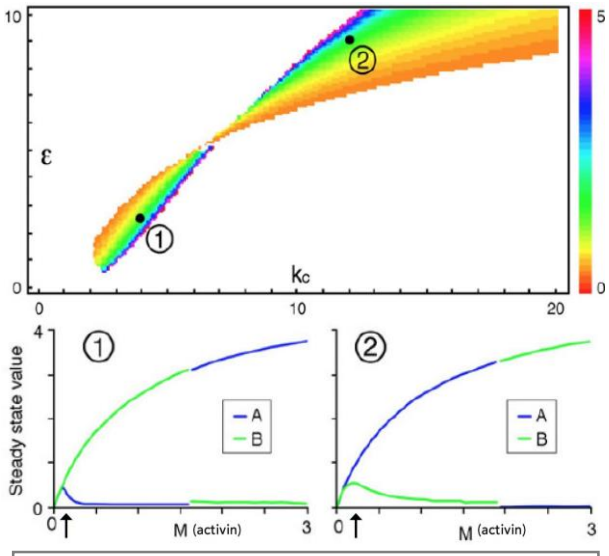
**Acknowledgements**: Aspects of this work have been supported by NSF and ASSETT@CU. Ongoing discussions with Melanie M. Cooper have deeply influenced my thoughts and their evolution. I have learned much from students' questions and beSocratic responses, as well as discussions with Eric Stade and Justin Brumbaugh. I appreciate the reference suggestion from Richard Harland and helpful comments and edits by R.P. Klymkowsky, Esq, although final errors are my own doing.

**Supplemental Information: Modeling a response:** In the Saka and Smith model we abstract and generalize the system, replacing protein and gene names with symbols. In such a model, many of the molecular mechanisms involved are telescoped into more general variables and used to generate systems of (solvable) differential equations, such as the activin-induced activation of the R-SMAD:SMAD4 transcriptional regulator (→). These equations enable us to model the effects of changes in various system parameters on outcomes. Here we characterize the relationship between the concentration of the signaling molecule [M], and the effects on the accumulation of the proteins encoded by the direct (A and B) and indirectly (C) regulated genes. These variables can take on a range of values and can be regulated through post-translational modifications, molecular interactions, and cellular localization together with their "reversal" rates. In addition, while the same signaling system directly regulates the expression of both A and B genes, the rates of A and B protein accumulation and steady state levels may be different – transcript and coding region lengths, RNA stability and codon usage, the rates of folding, assembly rates (for multimeric proteins), and interactions with competing targets and partners all influence them.

$$\frac{dA}{dt} = \frac{k_a}{1+C^{\gamma}} \cdot \frac{M^{\mu}}{1+M^{\mu}} - kd_a \cdot A$$

$$\frac{dB}{dt} = \frac{k_b}{1+A^{\alpha}} \cdot \frac{M^{\mu}}{1+M^{\mu}} - kd_b \cdot B$$

$$\frac{dC}{dt} = k_c \cdot \frac{B^{\varepsilon}}{1+B^{\varepsilon}} - kd_c \cdot C$$




*$k_a$, $k_b$ and $k_c$ are the synthesis rates of A, B and C. α and γ reflect the cooperativities of repression by A and C, ε and μ are the cooperativities of induction by B and M. $kd_a$, $kd_b$ and $kd_c$ are degradation rates of A,B, and C proteins. Graph plots the steady state expression levels of two genes (A and B) in response to increasing levels of a signaling molecule. As described, this system displays a sharp threshold where expression of one gene drops and the other increases dramatically. Our goal is to consider what resources students need to access so that they can produce plausible models of response initiation, saturation, and threshold behaviors. Figure modified from Saka & Smith, 2007*

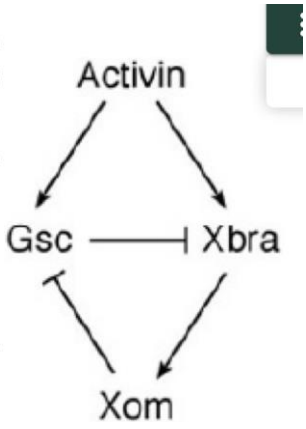
How the system behaves depends on the set of parameters that are applied, values that may or may not be easily determined experimentally, and may also vary between cells. Saka and Smith modeled the system's behavior at two parameter positions (marked 1 and 2 in the top graph (←). In both, behavior is similar at low concentrations of signaling [M] molecule (bottom graphs), a domain in which both genes A and B are expressed at similar low levels, the upper arrows in the lower panels. Expression behavior changes dramatically as [M] increases. In the two domains, expression of one or the other of the target genes increases, while the other drops to near zero. Expression of the active gene continues to increase until [M] crosses a threshold, at which point expression flips, the expression of the previously expressed gene drops to near zero while the expression of the previously unexpressed gene jumps to a high (plateauing) level. If we were to think of a plane of cells, in which there is a localized source of M that decreases with diffusion from the source, resulting in an [M] gradient, we might predict that we would see a domain of cells expressing gene A surrounded by a domain of cells expressing gene B or vice versa. A sharp boundary would separate the two domains. We would then expect the expression of A or B to lead to different "downstream" effects in terms of differences in gene expression and cellular behaviors.

**Q8 (6 points)**: In this system, activin activates the expression of *Gsc* and *Xbra* (→). Saka & Smith modeled this gene network based on estimates of the various parameters involved, including synthesis rates, half-lives, and the affinities of the Gsc, Xbra, and Xom proteins for their DNA target sites.

Assume that activin activates *Gsc* and *Xbra* expression equally, but that the movement of Gsc into the nucleus is faster than that of Xbra & Xom,

**A:** Predict (↓) the level of Gsc protein as a function of time (activin is added at time 10 and remains in the system)

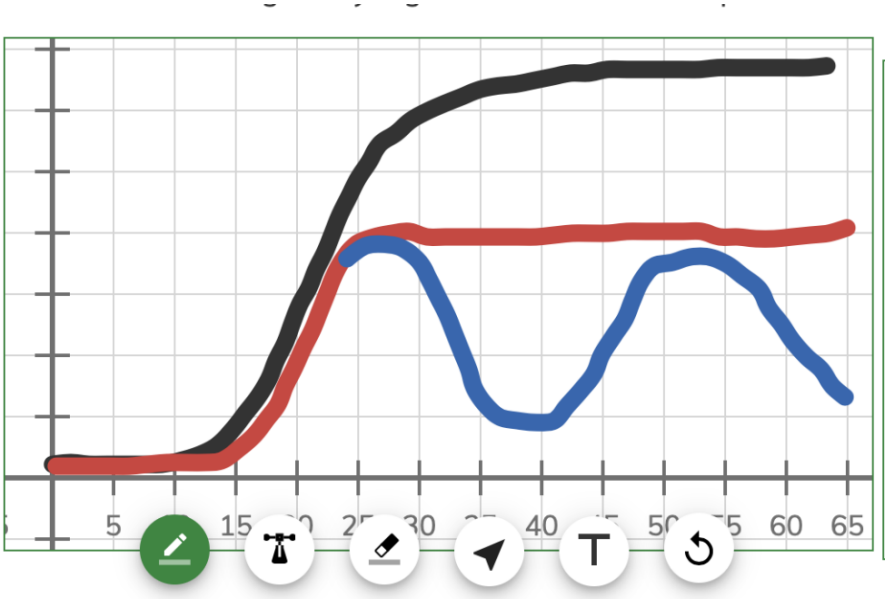
**B:** Predict **(in a different color)** How will the response change if Gsc also negatively regulates the activin receptor? describe your assumptions (↓)




**Appendix 2**: As an example of an exam question, we asked

**Target answer:** Assuming that the level of activin is sufficient to activate the system (1), we would expect that Gsc would accumulate in nuclei before Xbra. It would therefore act to repress *Xbra* expression (2), resulting in the long term inhibition of *Xbra* (and *Xom*) expression.

If Gsc also negatively regulates expression of the Activin Receptor, we might predict that the maximum response to a particular level of Activin would decrease (3 red line) - and perhaps even fall into the "no response" range, leading to a decrease and perhaps even the disappearance of Gsc expression (4 blue line).
Assuming that Activin remains present, once Gsc levels drop, Activin Receptor activity would increase and the cycle would repeat (5 blue line). It is worth noting that other, more complex responses are also possible



## Literature cited:


Adler, M. and Alon, U. (2018). Fold-change detection in biological systems. *Current Opinion in Systems Biology* 8, 81-89..

Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* 92, 291-294.

Ali, M. Z., Parisutham, V., Choubey, S. and Brewster, R. C. (2020). Inherent regulatory asymmetry emanating from network architecture in a prevalent autoregulatory motif. *eLife* 9, e56517.

Beatty, J. (1995). The evolutionary contingency thesis. Concepts, theories, and rationality in the biological sciences 45, 81.

Bryfczynski, S., Pargas, R. P., Cooper, M. M., Klymkowsky, M., Hester, J. and Grove, N. P. (2015). Classroom Uses for BeSocratic. In *The Impact of Pen and Touch Technology on Education* (ed. T. Hammond, S. Valentine, A. Adler & M. Payton): Springer.

Chaikuad, A. and Bullock, A. N. (2016). Structural basis of intracellular TGF-β signaling: Receptors and Smads. *Cold Spring Harbor perspectives in biology* 8, a022111.

Chow, C. C., Ong, K. M., Dougherty, E. J. and Simons Jr, S. S. (2011). Inferring mechanisms from dose–response curves. In *Methods in enzymology*, pp. 465-483: Elsevier.

Cooper, M. M. and Klymkowsky, M. W. (2013). The trouble with chemical energy: why understanding bond energies requires an interdisciplinary systems approach. *CBE Life Sci Educ* 12, 306-312.

Cooper, M. M. and Klymkowsky, M. W. (2016). Chemistry, Life, the Universe and Everything: An Evidence-Based, Transformed General Chemistry Curriculum. In *AAAS NSF EnFuse*. Washington, D.C.: AAAS and NSF.

Cooper, M. M., Stowe, R., Crandell, O. and Klymkowsky, M. W. (2019). Organic Chemistry, Life, the Universe and Everything (OCLUE): A Transformed Organic Chemistry Curriculum. *J. Chem. Educ.* 96, 1858-1872.

Darbellay, F., Bochaton, C., Lopez-Delisle, L., Mascrez, B., Tschopp, P., Delpretti, S., Zakany, J. and Duboule, D. (2019). The constrained architecture of mammalian Hox gene clusters. *Proceedings of the National Academy of Sciences* 116, 13424-13433.

Duboule, D. (2007). The rise and fall of Hox gene clusters. *Development* 134, 2549-2560..

Gardner, T. S., Cantor, C. R. and Collins, J. J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature* 403, 339-342.

Goentoro, L. and Kirschner, M. W. (2009). Evidence that fold-change, and not absolute level, of β-catenin dictates Wnt signaling. *Molecular cell* 36, 872-884.

Gurdon, J., Javed, K., Vodnala, M. and Garrett, N. (2020). Long-term association of a transcription factor with its chromatin binding site can stabilize gene expression and cell fate commitment. *Proceedings of the National Academy of Sciences*.

Harima, Y., Takashima, Y., Ueda, Y., Ohtsuka, T. and Kageyama, R. (2013). Accelerating the tempo of the segmentation clock by reducing the number of introns in the Hes7 gene. *Cell Reports* 3, 1-7.

Harland, R. (2011). We hold these principles to be self-evident. Oxford University Press for The Company of Biologists Limited.

Harland, R. M. and Grainger, R. M. (2011). Xenopus research: metamorphosed by genetics and genomics. *Trends in Genetics* 27, 507-515.

Klymkowsky, M. W. (2009). Make room for computing. *Science* 326, 227.

Klymkowsky, M. W. (2010). Thinking about the conceptual foundations of the biological sciences. *CBE Life Science Education* 9, 405-407.

Klymkowsky, M. W., Rentsch, J. D., Begovic, E. and Cooper, M. M. (2016). The design and transformation of Biofundamentals: a non-survey introductory evolutionary and molecular biology course. *LSE Cell Biol Edu.* pii: ar70.

Klymkowsky, M. W., Underwood, S. M. and Garvin-Doxas, K. (2010). Biological Concepts Instrument (BCI): A diagnostic tool for revealing student thinking. In *arXiv*: Cornell University Library.

Kohn, K. P., Underwood, S. and Cooper, M. M. (2018). Energy Connections and Misconnections across Chemistry and Biology. *CBE—Life Sciences Education* 17.

Lemmon, M. A., Freed, D. M., Schlessinger, J. and Kiyatkin, A. (2016). The dark side of cell signaling: positive roles for negative regulators. *Cell* 164, 1172-1184.

Lewin, R. (1982). Biology is not postage stamp collecting. *Science* 216, 718-720.

Li, P. and Elowitz, M. B. (2019). Communication codes in developmental signaling pathways. *Development* 146, dev170977.

Liao, B.-Y. and Zhang, J. (2008). Null mutations in human and mouse orthologs frequently result in different phenotypes. *Proceedings of the National Academy of Sciences*.

Lin, A., Giuliano, C. J., Palladino, A., John, K. M., Abramowicz, C., Yuan, M. L., Sausville, E. L., Lukow, D. A., Liu, L. and Chait, A. R. (2019). Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Science translational medicine* 11, eaaw8412.
Loomis, W. F. (2014). Cell signaling during development of Dictyostelium. *Developmental biology* 391, 1-16.
Mayr, E. (1961). Cause and effect in biology. *Science* 134, 1501-1506.
Mayr, E. (1985). The Growth of Biological Thought: Diversity, Evolution, and Inheritance. Cambridge, MA: Belknap Press of Harvard University Press.
McClymer, J. F. and Knowles, L. Z. (1992). Ersatz learning, inauthentic testing. *J. Excell. Coll. Teaching* 3, 33-50.
Meister, M. (2016). Physical limits to magnetogenetics. *Elife* 5, e17210.
Moriyama, Y. and De Robertis, E. M. (2018). Embryonic regeneration by relocalization of the Spemann organizer during twinning in Xenopus. *Proceedings of the National Academy of Sciences* 115, E4815-E4822.
Nehm, R. H. (2019). Biology education research: building integrative frameworks for teaching and learning about living systems. *Disciplinary and Interdisciplinary Science Education Research* 1, 15.
Nehm, R. H., Poole, T. M., Lyford, M. E., Hoskins, S. G., Carruth, L., Ewers, B. E. and Colberg, P. J. (2009). Does the segregation of evolution in biology textbooks and introductory courses reinforce students' faulty mental models of biology and evolution? *Evolution: Education and Outreach* 2, 527-532.
NRC (2012). A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Washington, DC:: The National Academies Press,.
Pevzner, P. and Shamir, R. (2009). Computing has changed biology—biology education must catch up. *Science* 325, 541-542.
Pulendran, B. and Davis, M. M. (2020). The science and medicine of human immunology. *Science* 369.
Saka, Y. and Smith, J. C. (2007). A mechanism for the sharp transition of morphogen gradient interpretation in Xenopus. *BMC Dev Biol* 7, 47.
Seok, J., Warren, H. S., Cuenca, A. G., Mindrinos, M. N., Baker, H. V., Xu, W., Richards, D. R., McDonald-Smith, G. P., Gao, H. and Hennessy, L. (2013). Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences* 110, 3507-3512.
Sharma, V., Hecker, N., Roscito, J. G., Foerster, L., Langer, B. E. and Hiller, M. (2018). A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. *Nature communications* 9, 1-9.
Shook, D. R., Kasprowicz, E. M., Davidson, L. A. and Keller, R. (2018). Large, long range tensile forces drive convergence during Xenopus blastopore closure and body axis elongation. *Elife* 7, e26944.
Smits, A. H., Ziebell, F., Joberty, G., Zinn, N., Mueller, W. F., Clauder-Münster, S., Eberhard, D., Savitski, M. F., Grandi, P. and Jakob, P. (2019). Biological plasticity rescues target activity in CRISPR knock outs. *Nature Methods* 16, 1087-1093.
Trujillo, C., Cooper, M. M. and Klymkowsky, M. W. (2012). Using graph-based assessments within socratic tutorials to reveal and refine students' analytical thinking about molecular networks. *Biochem Mol Biol Educ* 40, 100-107.
Warren, H. S., Tompkins, R. G., Mindrinos, M. N., Xiao, W. and Davis, R. W. (2013). Reply to Cauwels et al.: Of men, not mice, and inflammation. *Proceedings of the National Academy of Sciences* 110, E3151-E3151.
Wilkinson, M. F. (2019). Genetic paradox explained by nonsense. Nature Publishing Group.
Williams, L. C., Underwood, S. M., Klymkowsky, M. W. and Cooper, M. M. (2015). Are Noncovalent Interactions an Achilles Heel in Chemistry Education? A Comparison of Instructional Approaches. *Journal of Chemical Education* 92, 1979–1987.
Zhang, W., Landback, P., Gschwend, A. R., Shen, B. and Long, M. (2015). New genes drive the evolution of gene interaction networks in the human and mouse genomes. *Genome biology* 16, 202.
Zhu, J., Sanborn, J. Z., Diekhans, M., Lowe, C. B., Pringle, T. H. and Haussler, D. (2007). Comparative genomics search for losses of long-established genes on the human lineage. *PLoS computational biology* 3, e247.

# Making mechanistic sense: are we teaching students what they need to know?

**Michael W. Klymkowsky**
Molecular, Cellular & Developmental Biology,
University of Colorado Boulder, Boulder, CO. 80309

**Summary**: Evaluating learning outcomes depends upon objective and actionable measures of what students know – that is, what can they do with what they have learned. In the context of a developmental biology course, a capstone of many molecular biology degree programs, I asked students to predict the behaviors of temporal and spatial signaling gradients. Their responses led me to consider an alternative to conventional assessments, namely a process in which students are asked to build and apply plausible explanatory mechanistic models ("PEMMs"). A salient point is not whether students' models are correct, but whether they "work" in a manner consistent with underlying scientific principles. Analyzing such models can reveal the extent to which students recognize and accurately apply relevant ideas. An emphasis on model building, analysis and revision, an authentic scientific practice, can be expected to have transformative effects on course and curricular design as well as on student engagement and learning outcomes.

## Introduction

> *If the book is to remain manageable in size, it is inevitable that some favorite topics of the reader might be glossed over. However, despite the admirable emphasis on principles and concepts, I occasionally felt shortchanged. With these authors, I might have expected a deeper treatment of what surely must be one of the most important principles: the existence of threshold responses to morphogens – molecules that diffuse from a source and set up a graded concentration. Instead, when it gets to the nitty gritty of boundaries, the activating and repressing activities of one gradient (such as hunchback or dorsal) are mentioned, but how one protein both activates and represses is not explained.*
> – Richard Harland

Biology education has long struggled with the impression that it is more about memorization than the thoughtful consideration and application of widely relevant ideas (Lewin, 1982). Many teachers and academics often see and present areas of biology as distinct disciplines (Nehm, 2019; Nehm et al., 2009). The result is that some introductory biology courses read as surveys of implicitly unrelated topics, rather than a consideration of evolutionary processes acting on systems of interacting entities, whether molecules, cells, or organisms (Klymkowsky, 2010). To determine whether teaching is effective, that is, whether learning has occurred, there is often an emphasis on whether students know or can recognize the correct answer, rather than whether they can explain, in mechanistic terms, biological processes. Failing to emphasize the need to think about biological systems in a critical and mechanistic perspective contributes to a Dunning-Kruger effect (unwarranted confidence in one's understanding) in students, citizens, and politicians called on to make decisions related to a range of biomedical subjects (e.g., vaccine safety and the efficacy of homeopathic, naturopathic, and ineffective (and ecologically destructive) "folk" remedies). The ability to recognize correct answers in the context of a multiple-choice test is quite different from the ability to construct a relevant, plausible, and verifiable explanation, to justify the assumptions that support it, and to recognize the predictions that it implies. Various reform efforts in science education have focused on the importance of supporting learning within a coherent narrative that engages students, in part through a greater emphasis on research processes and common, cross-disciplinary principles (NRC, 2012).

As discussed by Beatty (1995), biology is distinctly different from chemistry and physics. From both a practical and a theoretical perspective, there are no "laws" in biology. While constrained by the laws and general principles of chemistry and physics, biological systems and processes are contingent upon their evolutionary history. As I will discuss, even closely related organisms can vary from one another in significant mechanistic detail. Typically, explanations of biological systems take two complementary forms: the mechanistic (how) and the evolutionary (why) (Mayr, 1961; 1985). At the cellular and molecular levels, mechanistic explanations address how a process occurs and involve the behaviors of molecular machines, a point made explicitly by Bruce Alberts (1998). The properties of molecules, the thermodynamics governing their interactions with one another, and the various chemical reactions in play, particularly the coupling of favorable reactions to drive unfavorable processes, contribute to and constrain such behaviors. "Why" explanations involve evolutionary adaptations, population behaviors (bottlenecks, founder effects, and genetic drift), and the ecological and environmental considerations driving speciation and species-specific behaviors. Both how and why mechanisms involve emergent behaviors, explicable only in terms of systems of interacting processes. While there are no biologic laws analogous to Newton's laws of motion or the laws of thermodynamics, the outcomes of biological processes reflect various, sometimes contradictory tendencies (see Beatty, 1995). While the complexities of biological systems complicate "how" explanations, the unobservable nature of evolutionary events that occurred in the distant past inherently constrain "why" explanations. When approaching the teaching and learning of biology, there are therefore two related questions: what is important to teach so that students can make sense of and appreciate the limits of our current knowledge, and how do we determine whether students are building an effective understanding, can they use their knowledge critically and constructively? How these questions are to be answered is rarely explicitly discussed within the (developmental) biology education community.

One approach to determining what one, and one's students know can be "borrowed" from working scientists and reflects the Socratic tradition. It is to have students engage in the process of building and defending plausible explanatory mechanistic models (“PEMMs”). Building a PEMM involves asking ourselves what combination of processes can produce the behaviors observed and then responding to questions raised by ourselves and others (fellow students, instructors, or peer reviewers) in order to determine:

(i) whether the model's underlying assumptions are consistent with established chemical and physical principles;
(ii) whether the model produces the expected behavior(s); and
(iii) how well the model predicts the response of the system to various perturbations.

Such model building and testing exercises can reveal misunderstandings about basic processes and their application, as well as the effects of specific details of the biological system under consideration. Equally important, they serve as a way to reveal whether students recognize as relevant underlying principles when constructing a plausible explanatory model. A version of this process occurs within the beSocratic exercises used with the CLUE, OCLUE, and biofundamentals course materials (Cooper and Klymkowsky, 2016; Cooper et al., 2019; Klymkowsky et al., 2016). Such activities can reveal how students understand and apply (or not) core ideas in response to how instructors present materials. As an example, such an analysis revealed the persistence of student confusions about the distinction between hydrogen and covalent bonds in conventionally designed courses (Williams et al., 2015) and the factors that influence the behavior of molecular networks (Trujillo et al., 2012) (see below). The process

of "reviewing" a model can reveal such mistakes and omissions. Does the logic of the model produce the behaviors that the model is meant to explain? Typically predictions of model outcomes are qualitative, but increasingly involves quantitative methods (see Ali et al., 2020; Gardner et al., 2000), raising the question of when and at what level mathematical modeling methods should be integrated into biology courses and curricula (see Klymkowsky, 2009; Pevzner and Shamir, 2009). A final test of a model is the extent to which it can predict the system's behavior in response to various perturbations. As the cycle repeats, there are opportunities for revision, the integration of new ideas, new details, and new components. Given the complexity and our often incomplete understanding of biological systems (Lin et al., 2019; Smits et al., 2019; Wilkinson, 2019), working researchers frequently revise models, and such revisions should be expected for students' explanatory models. What is critical in both arenas is the ability to revise models in response to review.

What is likely to be particularly valuable for students, and for course and curriculum designers and instructors, are the insights that building PEMMs can have in terms of identifying the ideas and practices required for their construction. Such an analysis enables us to look backward at what was presented to students (taught), what they are able to recognize as relevant and apply (learned), and to consider where in the curriculum specific ideas are best presented and practiced, so that reasonable models can be constructed and the habits of mind associated with model building can be reinforced. It is worth noting explicitly, and stressing to students, that even the most careful and experienced of model builders benefit from critical "third party" review. Michael Meister's analysis of paramagnetic effects on biological molecules provides a particularly informative example, revealing that certain "claims conflict with basic laws of physics. The discrepancies are large: from 5 to 10 log units. If the reported phenomena do in fact occur, they must have causes entirely different from the ones proposed by the authors" (Meister, 2016).[1]

The process of PEMM construction, analysis, and revision provides students with an authentic introduction to a basic scientific process and training in the habits of mind involved in deciding for themselves whether biology-based arguments make sense. Whether a student's original PEMM turns out to reflect the actual process is less important than that it works and serves as the basis for testing assumptions and responding to critical feedback. To illustrate their value, I describe my journey to appreciating the value of employing a PEMM evaluation model in the context of designing and teaching an upper division developmental biology course, often the last course majors are required to take before graduation.[2]

**Understanding developmental mechanisms:** The study of developmental biology is commonly rooted in examples from various "model" organisms, chosen for historical and practical reasons. Teaching developmental biology poses interesting challenges, because evolutionary adaptations end up producing species specific, and often functionally significant mechanistic variations. As an example, mouse is a common model system for studying early developmental events in mammals. And yet, there are well known and dramatic differences in a number of early (and late) developmental processes (Rossant, 2015), including basic mechanisms involved in cancer and other diseases (Pulendran and Davis, 2020; Seok et al., 2013; Warren et al., 2013). As one recent example, a highly conserved, long, non-coding RNA appears to play different roles in early mouse and human embryos (Sharma and Carninci, 2020), while null mutations can produce different phenotypes in human and mouse (Liao and

[1] A more humorous expansion of this theme can be found here: Magnetofiction – A Reader's Guide

[2] My ruminations on this topic are described here: https://bioliteracy.blog/on-devo/

Zhang, 2008). New genes arise (Zhang et al., 2015) and well-conserved genes are lost (Sharma et al., 2018; Zhu et al., 2007) in various lineages.

Frog and newt embryos have been used as model systems in a number of classic studies, ranging from establishing the feasibility of reprogramming somatic nuclei (Gurdon, 1962) to uncovering a range of inductive interactions and morphogenic processes (Harland and Grainger, 2011; Moriyama and De Robertis, 2018; Shook et al., 2018). Yet, there are substantial differences in developmental processes between amphibian species. For example, the maternal mRNA VegT (involved in the regulation of Nodal signaling) is localized to the oocyte's vegetal cortex in the clawed frog *Xenopus laevis*, a common model system for studies of vertebrate development, but found in the animal region of the marsupial frog *Eleutherodactylus coqui* oocytes (Elinson and del Pino, 2012). Of note, orthologs of VegT are absent from mammals. Similarly, the asymmetric distribution of bicoid protein in the fruit fly *Drosophila melanogaster* egg is often used to introduce how asymmetries in gene expression are established in the early embryo, yet the bicoid gene is "unique to higher dipterans" – it is absent from other insects (Lynch and Desplan, 2003). There are dramatic differences in the functional organization of HOX genes, involved in anterior-posterior (and other) embryonic asymmetries, between species (Darbellay et al., 2019; Duboule, 2007). So, if we are to follow the advice of Bill Wood (2008) to teach concepts and not (often species-specific) facts, on which concepts should we focus? Here the individual instructor is often provided little guidance and may come to rely on increasingly encyclopedic textbooks.

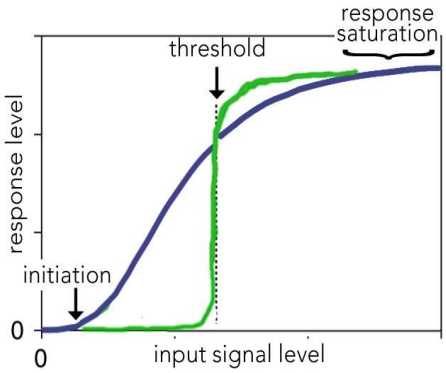



FIG.1: Examples of a standard sigmoidal dose-response curve (blue) indicating the point of response initiation and saturation, together with a threshold response (green) in which signal inputs for initiation and saturation levels are close to one another

**Threshold responses – a repeated theme in developing systems:** A key feature of developing systems, whether uni- or multicellular, is that they change over time in response to various signals.[3] A system's responses to these signals are generally not linear but display a distinct sigmoidal, and in the extreme case, a sharp "threshold" shape (Chow et al., 2011)(**FIG. 1**→). In the case of the unicellular slime mold *Dictyostelium discoideum,* the cellular response involves two distinct threshold signals, a "quorum sensing" system that monitors the number of cells per volume and a second system that reflects the cells' nutritional state (Loomis, 2014). Below a distinct signal molecule concentration, there is little or no cellular response. At a slightly higher signal concentration, that is above a "threshold," the signal response increases sharply and quickly saturates.

Student responses to questions related to signaling systems, delivered through the web-based beSocratic system (Bryfczynski et al., 2015), displayed some evident confusions associated with response onset, saturation, and threshold effects (**FIG. 2** ↓). Moreover, when asked to "provide a plausible molecular mechanism to produce that behavior (a threshold effect)," most students invoke active mechanisms associated with feedback loops, protein structure, nuclear import, or DNA modification – few explicitly recognized the need to overcome homeostatic processes (unpub. obs.). In fact, we had

[3] https://bioliteracy.blog/2018/12/15/on-teaching-developmental-biology-in-the-21st-century/

previously described similar confusions displayed by late stage molecular biology majors when consid-

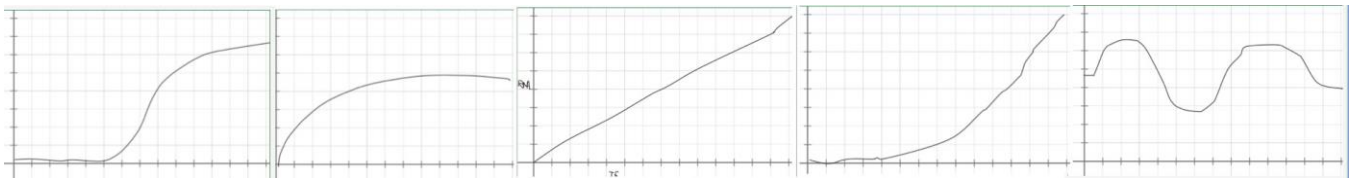


*FIG. 2 A representative set of student responses to the question, "In terms of increasing concentration of transcription factor (TF)(x-axis), draw your prediction of the expression level (RNA) of a particular target gene)."*

ering molecular networks (Trujillo et al., 2012). It appears that many students come to late stage "capstone" courses with a fragile understanding of the molecular processes involved in common signal-response behaviors. It follows that if students are to make sense of developmental processes, we must re-design instructional approaches so that they address these persistent difficulties. In the past, a common approach has been to identify students' problematic ideas and then design instruction to “overcome” or replace them with more canonical ideas. Unfortunately, there is little evidence that this approach works, beyond replacing misremembered facts. When considering the interplay of complex ideas, it seems entirely inadequate. We must address how to help students construct, connect, and contextualize their knowledge so that it becomes useful. I suggest that a strategy based on the construction, analysis, and revision of PEMMs is one such approach.

**What do we need to consider when building a signal response PEMM?** When getting specific about building PEMMs to explain sigmoidal responses and threshold behaviors, we first need to define exactly what we mean by the response. There are a number of possibilities for the student to consider – is it the immediate effect that follows the binding of the signaling molecule to its receptor, or is it the end behavior. In the case of slime mold cells, it is an observable behavior - the cells' migration toward each other to form a slug that goes on to differentiate. It can be the appearance of a particular pattern, such as the distinctive segments of a *Drosophila* embryo, or the patterns of gene expression that occur along the dorsal-ventral axes of the larvae or the vertebrate neural tube. Next our model needs to explain (i) why there is little or no response below a signal concentration, and (ii) why the response rises and then saturates as the signal concentration increases above its threshold concentration.

When originally faced with teaching these behaviors, I looked to the literature for established and concrete mechanisms to scaffold my presentation, rather than working through the possible (plausible) generic processes that might be involved. I found this search for accessible (teachable) mechanisms frustrating – often the mechanism(s) underlying "delayed" signal initiation was not clearly elucidated, if described or considered at all (see the text box above – Harland, 2011). While not described explicitly, mechanisms of saturation seemed inherently simpler (but not always obvious in student explanations); typically they involve limited numbers of regulatory targets – for example, there are generally only two copies of a particular gene per cell. Based on this view, although the mechanism(s) that lead to a higher initiation concentration may be complex, a threshold effect is a "simple" variant of a sigmoidal response, a response associated with a small difference between response onset and saturation.

So, what factors can influence the response initiation concentration? There are many, ranging from the concentration of receptors, signal-receptor binding affinity and "dwell" time (the half-life of the bound state), the effect of binding on receptor behavior, which can include receptor interactions with

other molecules and allosteric effects on enzymatic activity leading to the post-translational modification of targets that alter interactions, cellular localization, and rates of degradation. Such "downstream" and feedback effects can alter the numbers of available receptors, their response to signals, and to changes in gene expression, which in turn can influence the system's responsiveness through the expression of a range of agonists, antagonists, and modifiers. These changes do not occur in isolation; rather, they occur in the context of various homeostatic processes that act to return the system to its state before exposure to the signal. If the signal response alters gene expression, the state after a signaling response may well be different and so respond differently to the same signal molecule. The result is that the system is a product of its history, together with its energetic state, as these responses all involve coupled chemical reactions. Signal response systems are information processing molecular machines, in the sense described by Alberts (1998).

**Modeling response initiation**: To build a simple generic model of a sigmoidal response, a student will have to consider the various factors described above and the opposing homeostatic processes that act to reverse the effects of signal molecule-receptor "activation." At low signal concentrations, when few receptors are bound, homeostatic mechanisms will oppose the resulting signal-induced activity, e.g., targeted phosphorylation will be opposed by dephosphorylation. The response to a particular signal concentration therefore reflects both signaling effects and their "reversal rates." As an example, consider the case of a signal associated with the opening of an ion channel. The effect of opening a small number of channels will be offset by the on-going restorative, energy-dependent pumping mechanisms that act to maintain the cell's resting state. Only with increasing signal, which activates and opens more receptors, will the maintenance system be overwhelmed, and a response generated. The situation is further complicated when the initial steps in generating a response involve the assembly of a multiple component complex. The frequency and stability (lifetime) of the steps involved in this process will impact the probability that a functional and stable complex will be formed. That said, from a modeling perspective, the steps involved in such processes can often be "collapsed or telescoped" into a single step (*see* Chow et al., 2011). Together, these factors will determine the signal concentration at which a discernible response will occur.

While students may be expected to generate generic response models, they will require more detailed interaction diagrams when asked to consider "real" systems (as illustrated below). For example, the presence of antagonists can lead to an effective reduction in signal or receptor concentrations over time, while accessory factors can significantly increase dwell times compared to simpler systems, as suggested by observations on transcription factor binding (Gurdon et al., 2020). In this light, it appears that students' general understanding of the energetics of molecular interactions within biological systems is often weak (Cooper and Klymkowsky, 2013; Kohn et al., 2018). When considering response saturation, we need to recognize that the numbers of targets that can be activated, whether receptors, molecular machines, or genes, are limited. Once all targets are activated, the response will necessarily plateau. For students to be able to apply this idea within their models, including models of gene expression, teachers must lay the necessary groundwork in earlier courses.

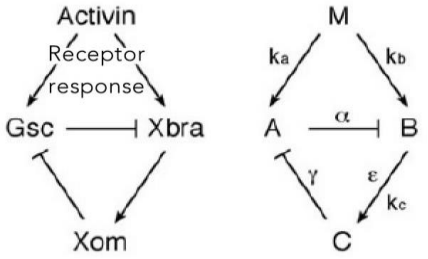



**FIG.3**: A schematic of regulatory circuit described by Saka & Smith (2007)

How can we present these complexities to students so that they are not overwhelmed? After some searching I was drawn to a signaling response model presented by Saka and Smith (2007). They aimed to explain how different levels of an extracellular signaling molecule (activin, a member of the TGFβ family) could differentially regulate the expression of one or another target gene, a system active during *Xenopus* development (**FIG. 3** →). What is particularly noteworthy about this system is its apparent simplicity and that it can produce two opposing outcomes depending on the assumptions made. The result is an accessible scenario to introduce the various considerations involved and how they impact signal-response outcomes and illustrate the range of behaviors that can be generated by "simple" systems. Similar considerations apply to a wide range of signaling systems and can be extended to considering immediate, steady state, adaptive, and cascading (evolving) responses (Lemmon et al., 2016; Li and Elowitz, 2019) as well as more complex fold-change sensing systems (Adler and Alon, 2018; Goentoro and Kirschner, 2009).

In the Saka and Smith model, secreted activin protein and the responding receptor system trigger an "upstream" process. The response (modified from Chaikuad and Bullock, 2016)(**FIG. 4** →) involves binding, changes in receptor structure, and interactions with other regulatory factors that lead to receptor kinase activation and the modification of cytoplasmic receptor-regulated SMAD proteins (R-SMADs). These processes are all reversible through various mechanisms. Phosphorylated R-SMADs dimerize in the cytoplasm and associate with the "co-SMAD," SMAD4. The cytoplasmic SMAD4:R-SMAD complex is then imported into the nucleus where it binds to specific DNA sequences and interacts with various accessory proteins leading to altered gene expression.

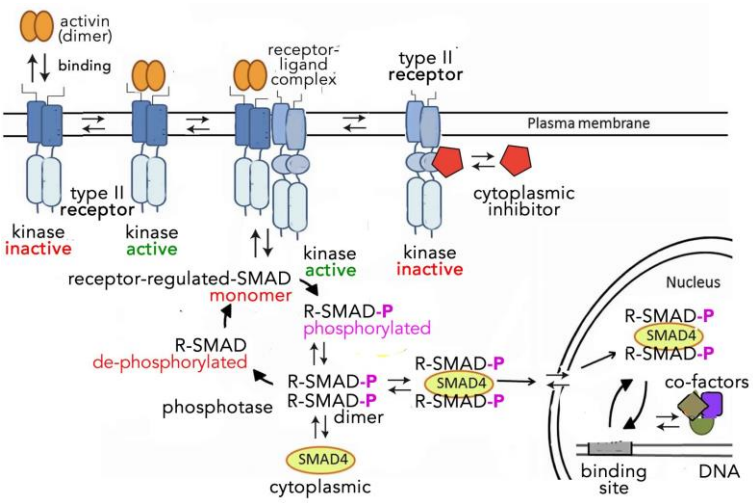



**FIG.4**: A schematic of the activin (TGFb) signaling pathway adopted with modifications from Chaikuad & Bullock,

In the Saka and Smith model (←**FIG 5**

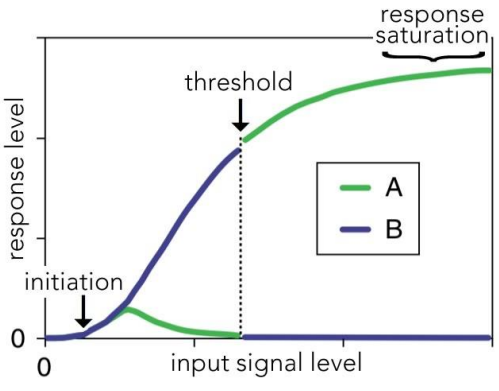



**FIG.5**: Outcome of the relationship between input signal (activin) and gene expression based on one set of parameters - modified from Saka & Smith (2007)

described in more detail in Appendix 1) exposure to activin leads to regulation of the expression of two genes, *Gsc* and *Xbra.*[4] Both encode sequence-specific DNA binding proteins and act as regulators of transcription. In their scenario, activin-activated R-SMAD/SMAD4 complexes directly regulate both *Gsc* and *Xbra*; there are no intervening genes whose transcription and translation are necessary for activin-regulated *Xbra* and *Gsc* gene expression. The gene products that are needed (and there are many) are already present within the cells. The network does, however, involve indirect effects: Gsc acts to inhibit *Xbra* expression and Xbra acts to induce *Xom* expression, with the Xom protein acting to repress *Gsc* expression – all other regulatory targets of Gsc, Xbra, and Xom are ignored. The model predicts dramatically different behaviors in terms of gene expression, based on assumptions about rates of target protein accumulation, a function of synthesis and degradation rates, together with target gene binding affinities.

Reproducing the Saka & Smith model requires a level of mathematical sophistication that most undergraduate biology students are unlikely to possess even though many biology degree programs have (or had) a calculus course requirement. That said, after an introduction to the system, students can build on general assumptions and develop qualitative models that produce clear mechanistic predictions. We can reasonably ask students to justify predictions as to how variations in various parameters (e.g. differences in protein stability, binding affinities, cytoplasmic localization, as well as various forms of feedback interactions) will influence network behaviors over time and space. As an example, we can expect students to be able to predict the effects of local signal sources, resulting in signaling (morphogen) gradients, as well as the effects of changes in gene and transcript size, as described by Harima et al (2013). An example question is supplied in Appendix 2.

**How do we prepare for the effects of a PEMM-centric approach?** The point of the PEMM approach is to focus instruction and to prepare students so that they can analyze processes they encounter in various biological contexts, ranging from the molecular to the ecological. As pointed out by McClymer & Knowles (1992), students' preparation needs to include information that every practitioner knows. We must therefore think hard about what that information consists of – is it details of specific systems or general principles such as how molecules interact and how those interactions influence their various activities? I would argue that general information should be defined not, for example, as the details of the Krebs cycle but instead as what principles are involved in coupling chemical reactions. Instructors can then introduce details as needed so students can consider specific processes, including whether they are reversible or not, and if not (e.g. proteolytic processing) how the system resets over time.

The question is whether the information we ask students to remember is useful to them in understanding and explaining a range of processes. In the case of sigmoidal response curves and threshold behaviors discussed here, we are working to develop a general understanding of a ubiquitous feature of a wide range of biological processes – from quorum sensing in microbial communities, the patterning of embryonic development, as well as (arguably) a range of physiological and social processes. The desired result is that students will be able to generate plausible explanatory models for how various regulators (extracellular and intracellular antagonists) and perturbations (mutations, toxic molecules, and environmental stressors) influence a system.

[4] Gene names are italicized, protein names are not.

So how does incorporating a PEMM-centered approach to presenting materials impact course and curricular design? The impact is likely to be dramatic, depending upon the extent of its implementation. In part, the effect reflects the need to introduce students' to systems thinking, model building, and the process of evaluating the implications of their assumptions. Through such a process, students gain direct experience with authentic scientific practices without adding an excessive rote memorization load. To prepare students, they need to be reminded of, or in some cases introduced to for the first time, the various cellular and molecular processes involved and how they interact in specific situations. This includes calling out the relevance of underlying (and universal) processes and necessary details. Practice in model building, presentation, analysis, and revision takes class time. In my own situation, materials are introduced in one class period; these are applied in the context of beSocratic activities that students' complete on their own. At the start of the next class period we review these activities and students are asked to present their solutions.[5]

Adoption of a PEMM approach encourages us to critically evaluate what has been "covered" in past courses and to focus on materials that students will need to use to construct and evaluate models of developmental processes. It means that we must consider what topics can be omitted or de-emphasized in order to make room for such explanatory model building, feedback, and revision. Given that instructors may worry that others will criticize course changes as an inappropriate over-simplification of course content (a criticism leveled at the CLUE chemistry curricula), it is important to have established benchmarks to evaluate student learning outcomes.[6]

Model building invites the incorporation of versions of peer review, revision, and resubmission, and helps us move from a one-off evaluative system to a more developmental process focused on applying and mastering underlying concepts and their application. We change the emphasis of assessment from the binary of right-wrong to the competency of learning how to perform a task. The idea that biological systems all share common features becomes the theme that unites them into a coherent and comprehensible whole – it provides a context within which to explicitly recognize and incorporate the details of specific systems as needed. Such an altered emphasis does, of necessity, demand a change of "coverage" and resource allocation, particularly in terms of instructor feedback and how we evaluate students learning outcomes. I would suggest moving from the use of timed high stakes exams, something no scientist would willingly accept (imagine if manuscripts or grant proposals had to be developed in a timed context), to one more like that of the preparation, peer review, revision, and eventual "publication" of a course dossier that documents a student's mastery of the materials and skills presented. Centering instruction and assessment, formative and summative on PEMMs and their presentation and revision would change course emphasis, hopefully making courses more engaging, inclusive, and effective in terms of learning outcomes and in fostering an inquisitive mindset. The PEMM approach explicitly values reflective thinking, something too often in short supply in a range of educational and broader social settings, .

[5] The course website can be visited here: http://virtuallaboratory.colorado.edu/DEVO@CU/index.html and will be archived.

[6] In the case of CLUE nationally normed exams generated by the American Chemical Society's Exams Institute to demonstrate no decrease in scores on these very traditional exams.

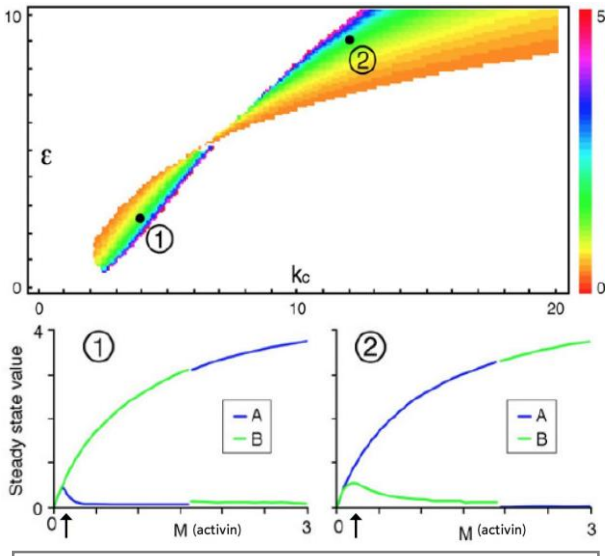
**Acknowledgements**: Aspects of this work have been supported by NSF and ASSETT@CU. Ongoing discussions with Melanie M. Cooper have deeply influenced my thoughts and their evolution. I have learned much from students' questions and beSocratic responses, as well as discussions with Eric Stade and Justin Brumbaugh. I appreciate the reference suggestion from Richard Harland and helpful comments and edits by R.P. Klymkowsky, Esq, although final errors are my own doing.

**Supplemental Information: Modeling a response:** In the Saka and Smith model we abstract and generalize the system, replacing protein and gene names with symbols. In such a model, many of the molecular mechanisms involved are telescoped into more general variables and used to generate systems of (solvable) differential equations, such as the activin-induced activation of the R-SMAD:SMAD4 transcriptional regulator (→). These equations enable us to model the effects of changes in various system parameters on outcomes. Here we characterize the relationship between the concentration of the signaling molecule [M], and the effects on the accumulation of the proteins encoded by the direct (A and B) and indirectly (C) regulated genes. These variables can take on a range of values and can be regulated through post-translational modifications, molecular interactions, and cellular localization together with their "reversal" rates. In addition, while the same signaling system directly regulates the expression of both A and B genes, the rates of A and B protein accumulation and steady state levels may be different – transcript and coding region lengths, RNA stability and codon usage, the rates of folding, assembly rates (for multimeric proteins), and interactions with competing targets and partners all influence them.

$$\frac{dA}{dt} = \frac{k_a}{1+C^{\gamma}} \cdot \frac{M^{\mu}}{1+M^{\mu}} - kd_a \cdot A$$

$$\frac{dB}{dt} = \frac{k_b}{1+A^{\alpha}} \cdot \frac{M^{\mu}}{1+M^{\mu}} - kd_b \cdot B$$

$$\frac{dC}{dt} = k_c \cdot \frac{B^{\varepsilon}}{1+B^{\varepsilon}} - kd_c \cdot C$$




*$k_a$, $k_b$ and $k_c$ are the synthesis rates of A, B and C. α and γ reflect the cooperativities of repression by A and C, ε and μ are the cooperativities of induction by B and M. $kd_a$, $kd_b$ and $kd_c$ are degradation rates of A,B, and C proteins. Graph plots the steady state expression levels of two genes (A and B) in response to increasing levels of a signaling molecule. As described, this system displays a sharp threshold where expression of one gene drops and the other increases dramatically. Our goal is to consider what resources students need to access so that they can produce plausible models of response initiation, saturation, and threshold behaviors. Figure modified from Saka & Smith, 2007*

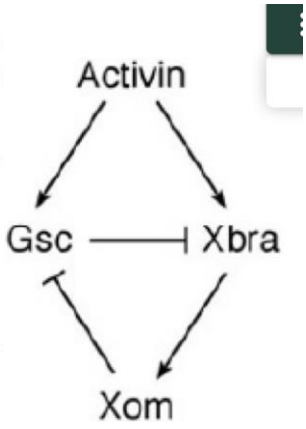
How the system behaves depends on the set of parameters that are applied, values that may or may not be easily determined experimentally, and may also vary between cells. Saka and Smith modeled the system's behavior at two parameter positions (marked 1 and 2 in the top graph (←). In both, behavior is similar at low concentrations of signaling [M] molecule (bottom graphs), a domain in which both genes A and B are expressed at similar low levels, the upper arrows in the lower panels. Expression behavior changes dramatically as [M] increases. In the two domains, expression of one or the other of the target genes increases, while the other drops to near zero. Expression of the active gene continues to increase until [M] crosses a threshold, at which point expression flips, the expression of the previously expressed gene drops to near zero while the expression of the previously unexpressed gene jumps to a high (plateauing) level. If we were to think of a plane of cells, in which there is a localized source of M that decreases with diffusion from the source, resulting in an [M] gradient, we might predict that we would see a domain of cells expressing gene A surrounded by a domain of cells expressing gene B or vice versa. A sharp boundary would separate the two domains. We would then expect the expression of A or B to lead to different "downstream" effects in terms of differences in gene expression and cellular behaviors.

**Q8 (6 points)**: In this system, activin activates the expression of *Gsc* and *Xbra* (→). Saka & Smith modeled this gene network based on estimates of the various parameters involved, including synthesis rates, half-lives, and the affinities of the Gsc, Xbra, and Xom proteins for their DNA target sites.

Assume that activin activates *Gsc* and *Xbra* expression equally, but that the movement of Gsc into the nucleus is faster than that of Xbra & Xom,

**A:** Predict (↓) the level of Gsc protein as a function of time (activin is added at time 10 and remains in the system)

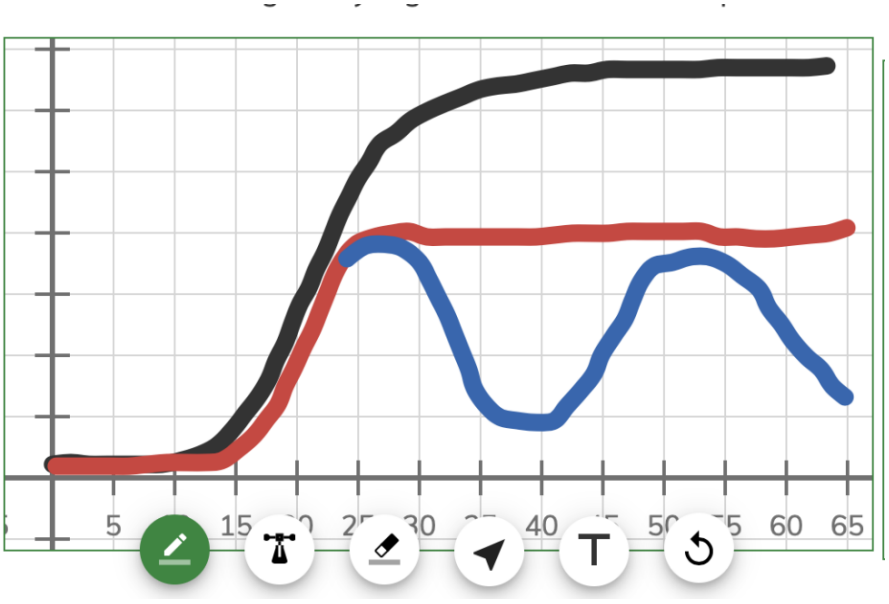
**B:** Predict **(in a different color)** How will the response change if Gsc also negatively regulates the activin receptor? describe your assumptions (↓)




**Appendix 2**: As an example of an exam question, we asked

**Target answer:** Assuming that the level of activin is sufficient to activate the system (1), we would expect that Gsc would accumulate in nuclei before Xbra. It would therefore act to repress *Xbra* expression (2), resulting in the long term inhibition of *Xbra* (and *Xom*) expression.

If Gsc also negatively regulates expression of the Activin Receptor, we might predict that the maximum response to a particular level of Activin would decrease (3 red line) - and perhaps even fall into the "no response" range, leading to a decrease and perhaps even the disappearance of Gsc expression (4 blue line).
Assuming that Activin remains present, once Gsc levels drop, Activin Receptor activity would increase and the cycle would repeat (5 blue line). It is worth noting that other, more complex responses are also possible



## Literature cited:


Adler, M. and Alon, U. (2018). Fold-change detection in biological systems. *Current Opinion in Systems Biology* 8, 81-89..

Alberts, B. (1998). The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* 92, 291-294.

Ali, M. Z., Parisutham, V., Choubey, S. and Brewster, R. C. (2020). Inherent regulatory asymmetry emanating from network architecture in a prevalent autoregulatory motif. *eLife* 9, e56517.

Beatty, J. (1995). The evolutionary contingency thesis. Concepts, theories, and rationality in the biological sciences 45, 81.

Bryfczynski, S., Pargas, R. P., Cooper, M. M., Klymkowsky, M., Hester, J. and Grove, N. P. (2015). Classroom Uses for BeSocratic. In *The Impact of Pen and Touch Technology on Education* (ed. T. Hammond, S. Valentine, A. Adler & M. Payton): Springer.

Chaikuad, A. and Bullock, A. N. (2016). Structural basis of intracellular TGF-β signaling: Receptors and Smads. *Cold Spring Harbor perspectives in biology* 8, a022111.

Chow, C. C., Ong, K. M., Dougherty, E. J. and Simons Jr, S. S. (2011). Inferring mechanisms from dose–response curves. In *Methods in enzymology*, pp. 465-483: Elsevier.

Cooper, M. M. and Klymkowsky, M. W. (2013). The trouble with chemical energy: why understanding bond energies requires an interdisciplinary systems approach. *CBE Life Sci Educ* 12, 306-312.

Cooper, M. M. and Klymkowsky, M. W. (2016). Chemistry, Life, the Universe and Everything: An Evidence-Based, Transformed General Chemistry Curriculum. In *AAAS NSF EnFuse*. Washington, D.C.: AAAS and NSF.

Cooper, M. M., Stowe, R., Crandell, O. and Klymkowsky, M. W. (2019). Organic Chemistry, Life, the Universe and Everything (OCLUE): A Transformed Organic Chemistry Curriculum. *J. Chem. Educ.* 96, 1858-1872.

Darbellay, F., Bochaton, C., Lopez-Delisle, L., Mascrez, B., Tschopp, P., Delpretti, S., Zakany, J. and Duboule, D. (2019). The constrained architecture of mammalian Hox gene clusters. *Proceedings of the National Academy of Sciences* 116, 13424-13433.

Duboule, D. (2007). The rise and fall of Hox gene clusters. *Development* 134, 2549-2560..

Gardner, T. S., Cantor, C. R. and Collins, J. J. (2000). Construction of a genetic toggle switch in Escherichia coli. *Nature* 403, 339-342.

Goentoro, L. and Kirschner, M. W. (2009). Evidence that fold-change, and not absolute level, of β-catenin dictates Wnt signaling. *Molecular cell* 36, 872-884.

Gurdon, J., Javed, K., Vodnala, M. and Garrett, N. (2020). Long-term association of a transcription factor with its chromatin binding site can stabilize gene expression and cell fate commitment. *Proceedings of the National Academy of Sciences*.

Harima, Y., Takashima, Y., Ueda, Y., Ohtsuka, T. and Kageyama, R. (2013). Accelerating the tempo of the segmentation clock by reducing the number of introns in the Hes7 gene. *Cell Reports* 3, 1-7.

Harland, R. (2011). We hold these principles to be self-evident. Oxford University Press for The Company of Biologists Limited.

Harland, R. M. and Grainger, R. M. (2011). Xenopus research: metamorphosed by genetics and genomics. *Trends in Genetics* 27, 507-515.

Klymkowsky, M. W. (2009). Make room for computing. *Science* 326, 227.

Klymkowsky, M. W. (2010). Thinking about the conceptual foundations of the biological sciences. *CBE Life Science Education* 9, 405-407.

Klymkowsky, M. W., Rentsch, J. D., Begovic, E. and Cooper, M. M. (2016). The design and transformation of Biofundamentals: a non-survey introductory evolutionary and molecular biology course. *LSE Cell Biol Edu.* pii: ar70.

Klymkowsky, M. W., Underwood, S. M. and Garvin-Doxas, K. (2010). Biological Concepts Instrument (BCI): A diagnostic tool for revealing student thinking. In *arXiv*: Cornell University Library.

Kohn, K. P., Underwood, S. and Cooper, M. M. (2018). Energy Connections and Misconnections across Chemistry and Biology. *CBE—Life Sciences Education* 17.

Lemmon, M. A., Freed, D. M., Schlessinger, J. and Kiyatkin, A. (2016). The dark side of cell signaling: positive roles for negative regulators. *Cell* 164, 1172-1184.

Lewin, R. (1982). Biology is not postage stamp collecting. *Science* 216, 718-720.

Li, P. and Elowitz, M. B. (2019). Communication codes in developmental signaling pathways. *Development* 146, dev170977.

Liao, B.-Y. and Zhang, J. (2008). Null mutations in human and mouse orthologs frequently result in different phenotypes. *Proceedings of the National Academy of Sciences*.

Lin, A., Giuliano, C. J., Palladino, A., John, K. M., Abramowicz, C., Yuan, M. L., Sausville, E. L., Lukow, D. A., Liu, L. and Chait, A. R. (2019). Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Science translational medicine* 11, eaaw8412.
Loomis, W. F. (2014). Cell signaling during development of Dictyostelium. *Developmental biology* 391, 1-16.
Mayr, E. (1961). Cause and effect in biology. *Science* 134, 1501-1506.
Mayr, E. (1985). The Growth of Biological Thought: Diversity, Evolution, and Inheritance. Cambridge, MA: Belknap Press of Harvard University Press.
McClymer, J. F. and Knowles, L. Z. (1992). Ersatz learning, inauthentic testing. *J. Excell. Coll. Teaching* 3, 33-50.
Meister, M. (2016). Physical limits to magnetogenetics. *Elife* 5, e17210.
Moriyama, Y. and De Robertis, E. M. (2018). Embryonic regeneration by relocalization of the Spemann organizer during twinning in Xenopus. *Proceedings of the National Academy of Sciences* 115, E4815-E4822.
Nehm, R. H. (2019). Biology education research: building integrative frameworks for teaching and learning about living systems. *Disciplinary and Interdisciplinary Science Education Research* 1, 15.
Nehm, R. H., Poole, T. M., Lyford, M. E., Hoskins, S. G., Carruth, L., Ewers, B. E. and Colberg, P. J. (2009). Does the segregation of evolution in biology textbooks and introductory courses reinforce students' faulty mental models of biology and evolution? *Evolution: Education and Outreach* 2, 527-532.
NRC (2012). A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Washington, DC:: The National Academies Press,.
Pevzner, P. and Shamir, R. (2009). Computing has changed biology—biology education must catch up. *Science* 325, 541-542.
Pulendran, B. and Davis, M. M. (2020). The science and medicine of human immunology. *Science* 369.
Saka, Y. and Smith, J. C. (2007). A mechanism for the sharp transition of morphogen gradient interpretation in Xenopus. *BMC Dev Biol* 7, 47.
Seok, J., Warren, H. S., Cuenca, A. G., Mindrinos, M. N., Baker, H. V., Xu, W., Richards, D. R., McDonald-Smith, G. P., Gao, H. and Hennessy, L. (2013). Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences* 110, 3507-3512.
Sharma, V., Hecker, N., Roscito, J. G., Foerster, L., Langer, B. E. and Hiller, M. (2018). A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. *Nature communications* 9, 1-9.
Shook, D. R., Kasprowicz, E. M., Davidson, L. A. and Keller, R. (2018). Large, long range tensile forces drive convergence during Xenopus blastopore closure and body axis elongation. *Elife* 7, e26944.
Smits, A. H., Ziebell, F., Joberty, G., Zinn, N., Mueller, W. F., Clauder-Münster, S., Eberhard, D., Savitski, M. F., Grandi, P. and Jakob, P. (2019). Biological plasticity rescues target activity in CRISPR knock outs. *Nature Methods* 16, 1087-1093.
Trujillo, C., Cooper, M. M. and Klymkowsky, M. W. (2012). Using graph-based assessments within socratic tutorials to reveal and refine students' analytical thinking about molecular networks. *Biochem Mol Biol Educ* 40, 100-107.
Warren, H. S., Tompkins, R. G., Mindrinos, M. N., Xiao, W. and Davis, R. W. (2013). Reply to Cauwels et al.: Of men, not mice, and inflammation. *Proceedings of the National Academy of Sciences* 110, E3151-E3151.
Wilkinson, M. F. (2019). Genetic paradox explained by nonsense. Nature Publishing Group.
Williams, L. C., Underwood, S. M., Klymkowsky, M. W. and Cooper, M. M. (2015). Are Noncovalent Interactions an Achilles Heel in Chemistry Education? A Comparison of Instructional Approaches. *Journal of Chemical Education* 92, 1979–1987.
Zhang, W., Landback, P., Gschwend, A. R., Shen, B. and Long, M. (2015). New genes drive the evolution of gene interaction networks in the human and mouse genomes. *Genome biology* 16, 202.
Zhu, J., Sanborn, J. Z., Diekhans, M., Lowe, C. B., Pringle, T. H. and Haussler, D. (2007). Comparative genomics search for losses of long-established genes on the human lineage. *PLoS computational biology* 3, e247.