Operator confidence in models:
Comparing two business examples

**Paul Newton, Rod MacDonald, Suzanne Riney, Frank Rasor,
Mary Rita Roux-Zink, Neil Roberts, Rich Bierlein, Eric Garday**

The Boeing Company
1420 South Trenton Street
Room 21A3-6, Building 15-20, Mail Code 2T-44,
Seattle, WA 98108-4848, USA
206-419-4397
paul.c.newton2@boeing.com
(The above are corresponding author Paul Newton's contact information)[1]

Abstract

Forrester (1973, pages 24 & 35) discusses "how operators (people who make decisions to control action) develop confidence in a model's suitability for purpose. …From the viewpoint of the operator…there is a network of contacts between a model and reality. As the multiple contacts are explored without showing serious discrepancy between a model and the real world it represents, confidence in the model increases." We describe two modeling projects begun at essentially the same time with shared sponsorship and a shared modeler, but which used different approaches. One approach we name Comparison of Alternative Policies (CAP). The other we name Comparison of Alternative Mental Models of System Structure (CAMMSS). CAP is classical system dynamics; whereas, the CAMMSS approach seems fairly atypical. However, the CAMMSS-developed model is getting significantly more operator interest and referrals than the CAP-developed model. This leads us to believe that operators have more confidence in the CAMMSS-developed model. We explore how the CAMMSS-developed model may provide operators more non-discrepant contacts with their perception of reality, thus creating Forrester's conditions for more operator confidence. Our origins for CAMMSS lie in Senge's (1990, 2006, pages 47-51 in both editions) description of how to improve performance in the Beer Game.

Keywords:

Safety, productivity, validity, validation, verification, confidence, mental, practice

Introduction

Some might wonder what is meant by "operator" in this paper's title. Forrester (1973, p.24)[2] defines operators as "people who make decisions to control action." He gives several examples:

[1] With the exception of Rod MacDonald, who is a modeler contracted by Boeing for one of the two models described herein, all the other authors are Boeing employees and their roles on the projects described herein are discussed in the paper. If you want the contact information for any of the authors, please contact corresponding author Paul Newton.

business people, politicians, lawyers, medical doctors and engineers. Forrester distinguishes "operators" from "observers." He argues that "Model-building observers take a very different view of models from that of model-building operators. They explain and criticize but do not act. They compare to the operator as the staff advisor compares to a manager" (p. 25).

This paper describes two independent, but simultaneous, modeling projects in the same company, both related to improving safety dynamics, one for hazardous operations in a Service Organization (SO)[3] and the other for an enterprise safety initiative. Both projects were sponsored by the same manager, co-author Suzanne Riney – an "operator" per Forrester. And both projects utilized the same modeler, co-author Paul Newton – an "observer" per Forrester. We will refer to the service organization model as the Hazardous Operations Safety (HOS) model, and the enterprise safety initiative model as the Mental Models Relating Safety and Productivity (MMRSP) model.

Each model offers significant, and probably comparable, value to the company. However the MMRSP model continues to attract decidedly more operator interest, and operator referrals to other operators in the company, than the HOS model. This paper focuses on a novel difference between the two modeling projects that may give rise to this differential operator interest and referrals, and that may be significant for the practice of system dynamics. One aspect of traditional system dynamics practice is that it compares dynamics produced by alternative policies (CAP). The HOS model, the one that has so far received the least operator interest, used this classical CAP approach. The MMRSP model, the one that is receiving more operator interest, instead took the novel approach of comparing dynamics produced by alternative mental models of system structure (CAMMSS).

Forrester (1973) goes on to contrast how operators and observers develop confidence in a model's suitability for purpose. For operators he summarizes, "From the viewpoint of the operator…there is a network of contacts between a model and reality. As the multiple contacts are explored without showing serious discrepancy between a model and the real world it represents, confidence in the model increases," (p. 35).[4] That there is more interest in the MMRSP model than in the HSO model, as expressed by more referrals, is herein assumed as evidence of more operator confidence in the MMRSP model. The purpose of this paper is to explore the ways in which the CAMMSS approach used on the MMRSP model may provide more, or more meaningful, non-discrepant "contacts between a model and reality" for operators than does the CAP approach used on the HOS project. But first, to provide background for the reader, we compare the two modeling projects and models.

[2] While exploring the system dynamics model validity literature in preparation for this paper we came across this statement in Barlas (1990), "…the most complete discussion of the philosophy of system dynamics model validation is given by Forrester (1973) in an unpublished research paper, portions of which were revised and published by Forrester and Senge (1980)."

[3] Service Organization (SO) is a fictitious name for this organization inside the company.

[4] Of course, reality for each of us, including operators, is only our individual perception of reality.

Comparing the two modeling projects and models

The two modeling projects were quite different in their purposes (Table 1), sponsorship and initiation (Table 2), and the activities on their respective timelines (Figure 1). Note in Table 1:

1) Both models had significant shifts in their purposes, and
2) The HOS model began with a very typical purpose for a system dynamics project [to "test proposed mitigations" (policies)], whereas the MMRSP model began with a very general purpose that doesn't even express a definite need for system dynamics (note the "Initiation" cell for MMRSP in Table 2 – Riney's "thinking STim[5] might be useful").

Further, note in Table 2 that system dynamics (STim[5]) was the primary approach from the beginning for the HOS project, but not for the MMRSP project. Yet, as of this writing, the MMRSP model continues to garner more operator interest, referrals, and presumably confidence, as shown in the number of referred presentations in the timeline in Figure 1.

The Hazardous Operations Safety (HOS) model using the CAP approach

Turning our attention to the HOS model, Riney, at Newton's request, contracted co-author Rod MacDonald to focus exclusively on the modeling work. Note in the top row of Figure 1 the nearly one year duration of MacDonald's involvement. Over that year MacDonald, with support from Newton, facilitated the group modeling activity that was a large portion of two two-hour modeling team meetings led each week by co-author Neil Roberts. The subject matter experts on the modeling team primarily included co-authors and operators Rich Bierlein, Eric Garday and Neil Roberts. Roberts provided safety discipline expertise specific to the hazardous operations being modeled, whereas Garday and Bierlein each have decades of experience actually doing and leading the hazardous work being modeled. Additional safety experts[6] followed the work and participated quite often. Every couple of months the team would brief and solicit feedback from Bierlein and Garday's peers on a Senior Technical Team, each member of which also has decades of experience with the hazardous operation being modeled. Likewise the modeling team gave frequent briefings to the management of the hazardous operation, including Riney, co-author Frank Rasor, and Rasor's leadership team. MacDonald, with Newton as a collaborator, did all of the detailed modeling, model testing and calibration to hazardous operation information and data provided by Garday and Bierlein. MacDonald calibrated the HOS model to plausibly replicate project historical data for many hazardous projects in the service organization. In short, this was a classic system dynamics modeling exercise that, as can be seen in Figure 1, lasted almost a full year.

Over a period of years, the service organization had adopted qualitative systems thinking approaches. Causal loop diagrams are prominent on the walls of the service organization's

[5] STim is an acronym for "Systems Thinking aided by Simulation." It has the same meaning as "System Dynamics." We use this acronym at the company because "Systems Thinking" is quite well known. When we introduce "System Dynamics" it causes questions such as, "Right now, are we doing Systems Thinking or System Dynamics?" To avoid such questions, we use the STim acronym.

[6] Other frequent participants included Jennifer Klemisch and Xidong Xu. Klemisch, like Roberts, provided safety discipline expertise specific to the hazardous operations being modeled. Xu is a human factors and safety expert from the company's R&D organization.

conference rooms and cubicles in facilities around the country. However, as is common in the service organization, Garday was skeptical of the correctness of business dynamics inferences from the development and use of such diagrams. However, over the course of this project he came to see that system dynamics provides a way to test the thinking in, and conclusions from, such diagrams. Garday and Bierlein developed a great deal of confidence in the model and its findings – so much so that in the aforementioned meetings with their peers, and with Rasor and his leadership team, Bierlein and Garday were outstanding in describing the work, relating it to real-world examples in the hazardous operations, and answering questions.

	Mental Models Relating Safety & Productivity (MMRSP) model	Hazardous Operations Safety (HOS) model
Initial Purpose	Explore the meaning of a "safety system"	Test proposed mitigations to improve the safety of hazardous work operations in the SO.
Current Purpose	To be the core of an online course to inspire people who make safety-related decisions to discover, challenge and improve their mental models relating safety and productivity, both personal and collective, and short and long term.	To inspire reflection, both in the service organization and its internal customers, on how varying a hazardous operation project's start date can affect both safety and schedule dynamics.

Table 1: Initial and current purposes for each model

	Mental Models Relating Safety & Productivity (MMRSP) model	Hazardous Operations Safety (HOS) model
Sponsorship	Riney initially, with co-author Roux-Zink joining later	Riney initially, with Rasor joining later
Initiation	STim[5] was on the periphery from the beginning of the SMS project. Thinking STim[5] might be useful, Riney engaged modeler Newton to participate on the team. Indeed, the purpose of the team's project did not clearly lend itself to STim.[5] Newton found Cooke's (2006) model, saw relevance, attempted to explain it to the team, but the team never really adopted it.	STim[5] is a primary approach used to address the problem. Riney initially engaged modeler Newton, and later MacDonald at Newton's request, to facilitate the process and build the model. Roberts, the team leader, was interested in using STim[5] from the beginning. Also, the problem statement lent itself to STim.[5]

Table 2: Sponsorship and initiation for each model

All this led to Rasor's seeing business improvement possibilities from the model, in particular in relation to improving collaboration with the internal partner organizations that Rasor's service organization serves. The model's purpose gradually shifted to include not only safety performance, but also hazardous operations schedule performance, especially in relation to when to start the hazardous operations work (see the HOS model's "Current Purpose" in Table 1). Because upstream partner performance is so critical to this choice of start date, we used the model to illustrate how the hazardous operation project's start date can impact both its finish date and its safety over the course of the project. This led to recognition that the boundary of the model should be expanded to include our upstream internal partner organizations' work. As shown in Figure 1, Rasor and Newton presented this idea to the service organization's VP and his leadership team. Although the work was well received with much positive discussion in the meeting, as of yet there has been no subsequent action. However, it is common in the company for system dynamics work to languish and then come back to life at a later date. Hopefully that is what will happen in this case.

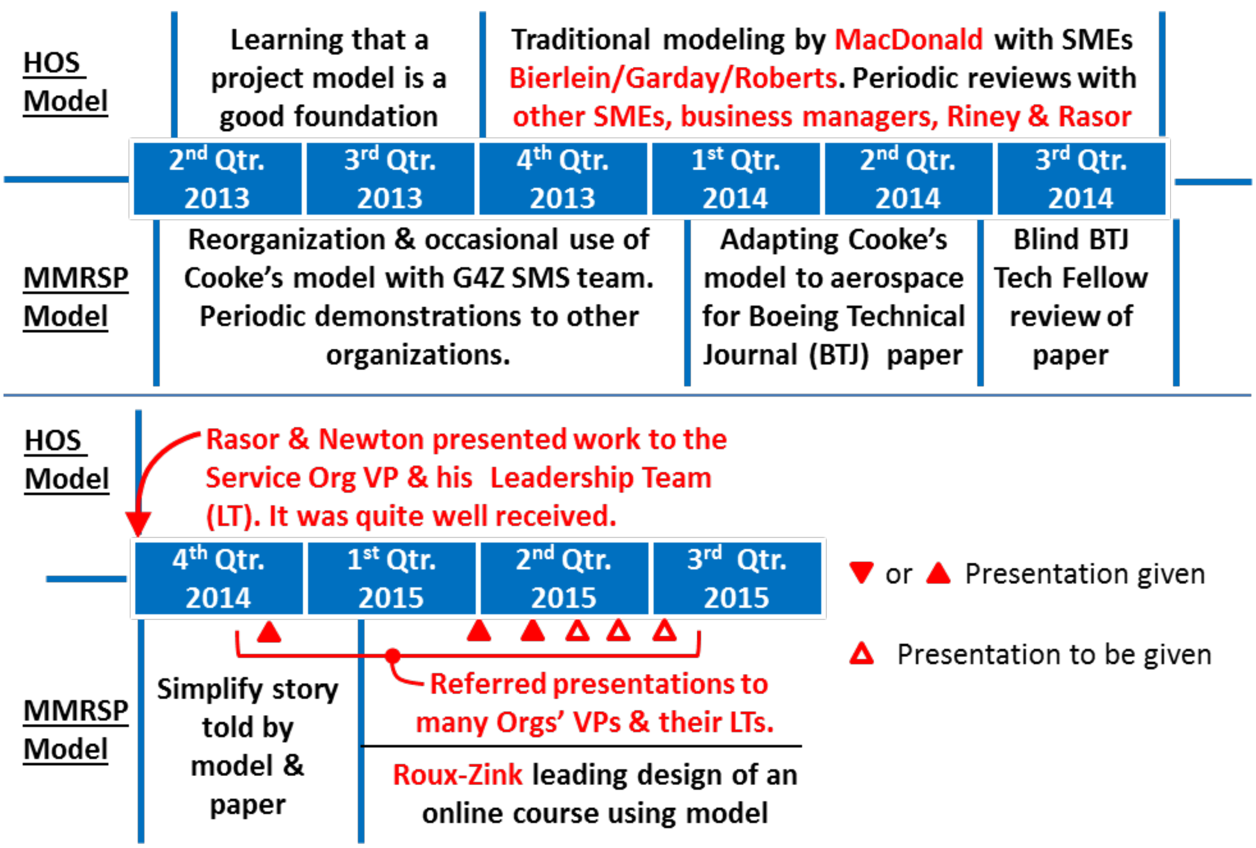

Figure 1: Timeline for each model

The Mental Models Relating Safety and Productivity (MMRSP) model using the CAMMSS approach

Comparing the MMRSP model in the second row of Figure 1 with the HOS model in the first row, note the following differences:

1) Much more elapsed time was spent on reflection in the first phase (approximately 10 months rather than approximately 5 months).
2) There was no explicit formal modeling team. Only Newton, the researchers whose work provided the foundation for his development of the MMRSP model, and his absorption of information from engaging informally with others on this topic as an employee of the company, provided the information used in creating and testing the MMRSP model.
3) The need to explain the work in a peer-reviewed technical paper. The peer review led to both significant paper simplification expanding the number of operators who would find the paper and work of value, and refinement of the CAMMSS approach for this project.

In addition there were other major differences, not shown in Figure 1. As compared to the HOS model, the MMRSP model:

4) Had a much less focused initial purpose (see Table 1)
5) Represented generic system structure, rather than system structure specific to a particular organization. Had the MMRSP model been approached as specific to a particular problem in a particular organization like the HOS model was, it likely would similarly have required experts from that organization as part of the modeling team.
6) Focused on comparative testing of mental models of system structure, rather than comparative testing of policies (called "mitigations" in the HOS model)[7]
7) Used much less calibration to historical time series data. The MMRSP model only:
   a) Produces plausible safety output values for one variable (Lost Workday Case Rate) that are in the range of real-world values, and
   b) Uses plausible worker population and production rate values for a generic typical airplane assembly plant.

Reflecting on these comparisons of the MMRSP model with the HOS model, items 1, 2, 3 and 4 above likely provided the modeler more exploration time, freedom and flexibility, thereby possibly enabling the modeler to respond more confidently, yet reflectively, to a wider range of operator challenges when presenting the work. It seems that such confident reflection by the modeler in response to operator challenges would tend to increase operator confidence in the model. Item 5 would usually be viewed as having the effect of reducing operator confidence in the model. Yet operators seem to have more confidence in this model. Why? This paper explores item 6 as a possible significant cause. Explaining why item 6 is a cause returns us to this paper's purpose: to explore the ways in which the CAMMSS approach used on the MMRSP model may provide more, or more meaningful, non-discrepant "contacts between a model and reality" for operators than does the CAP approach used on the HOS project. Finally, item 7 likewise would usually be viewed as having the effect of reducing operator confidence. However, as discussed later herein, the fuzziness of a CAMMSS model, e.g. the MMRSP model, probably reduces operators' sense of need for calibration to historical time series data (item 7).

[7] Newton's use of the CAMMSS approach originated when he moved from the company's research & technology organization to an organization development group in the Service Organization (SO) in the fall of 2013. That mental models are important in organization development led Newton to the notion of CAMMSS. Senge (1990, 2006, pages 47-51 in both editions) provides precedent for the CAMMSS approach in his description of how to improve performance in the Beer Game.

Gaining confidence and broad interest in implementation

As just described, despite occurring in parallel, the development process for the HOS and MMRSP models were very different from each other. However, in the end, whether or not either modeling project will result in significant and broad change in a large company is a function of the socialization of the model and related recommendations with audiences of operators who were not involved in the model's development. This socialization process is a referral process in which one set of operators that has become aware of the model refers the work to one or more sets of other operators. Such referrals seem to occur only when the potentially referring set of operators has sufficient interest and confidence in the model and related recommendations to feel comfortable making a referral. Per Forrester (1973) such operator confidence requires a non-discrepant "network of contacts between a model and reality." Such contacts include both system structure contacts, and system behavior contacts.

System structure contacts between a model and reality

In system dynamics, system structure is typically presented to an audience of operators in causal loop diagram (CLD) form. This was indeed the case for both the HOS and MMRSP models. The CLDs in Figures 2 and 3 are the CLDs of the HOS and MMRSP models, respectively, as shown to operators when presenting each model. In both cases, these CLDs are not presented in full, but rather are gradually exposed, loop by loop, with accompanying discussion. Note the significant comparative simplicity of the MMRSP CLD in Figure 3.

The reason for this simplicity is that the MMRSP CLD is not intended to represent the full model. Rather, it represents the structure required to describe both a plausible common mental model relating safety and productivity, and a plausibly better (more realistic) mental model (see the Mental Models layer in Figure 4). In general the CAMMSS approach taken on the MMRSP project yields simpler, easier-to-understand, CLDs for operators than does the CAP approach taken on the HOS project, thus more easily providing non-discrepant "contacts between a model and reality" in operators' minds.

Another non-discrepant contact follows from upfront framing of the CLD as a mental model, as shown in the Mental Models of System Structure layer in Figure 4. Reflecting on their own mental models, operators quickly appreciate that mental models are simple relative to the full system relevant to the problem being modeled. Operators easily recognize that our mental models are the foundation of our decision rules, and that if our mental models are significantly different than the real system, then our decision rules will yield poor decision streams, resulting in unintended, and often undesirable, system performance.[8] This, in itself, is a non-discrepant contact between a model and reality, but of a different sort – we could call it a meta-contact. The reality is that we know the way we think about system structure is simpler than the actual system structure, and seeing that the modeling approach recognizes this is a "meta-contact." Seeing plausible dynamic simulation output arising from alternative mental models, as shown in the Patterns Over Time layer in Figure 4, seems to cement this non-discrepant meta-contact "between a model and reality" in operators' minds.

---

[8] In this sense, a mental model is essentially the same thing as a "policy" in classic modeling practice.

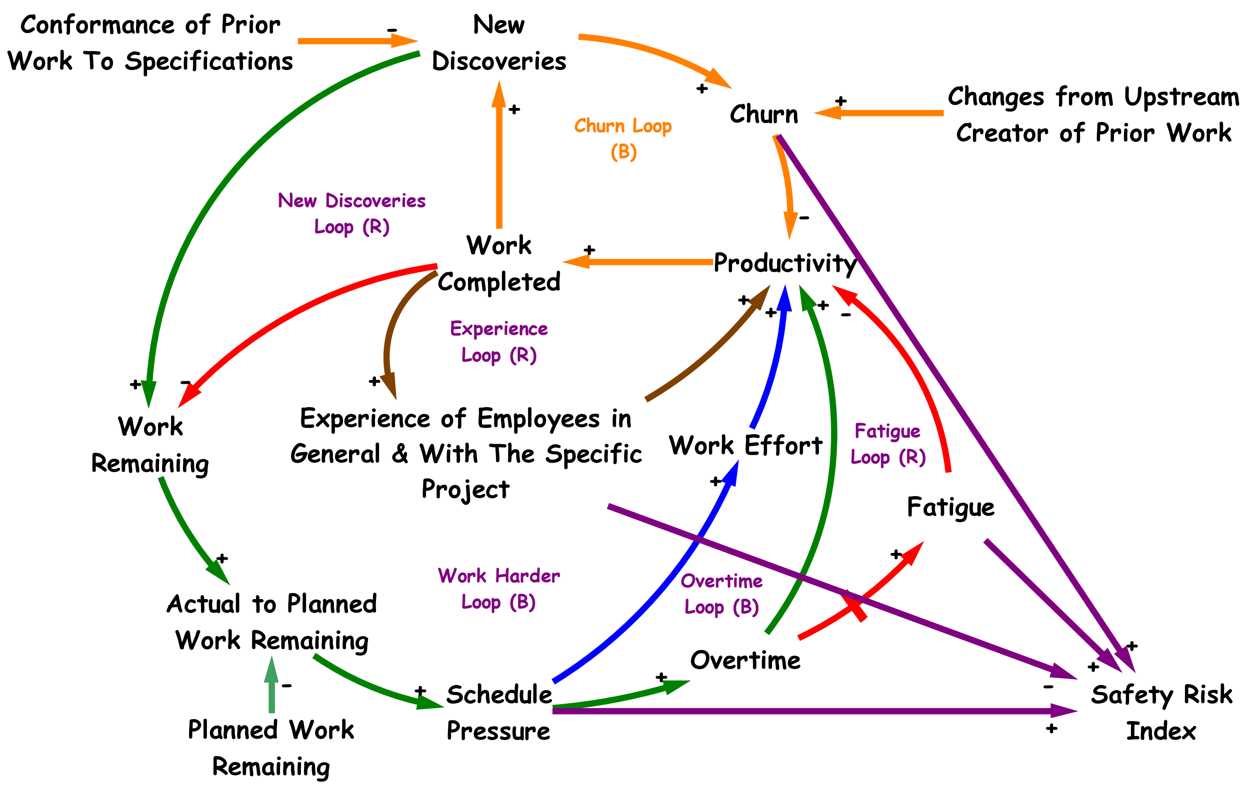

Figure 2: CLD gradually revealed to operators when presenting the structure, and explaining the resulting behavior, of the HOS model.

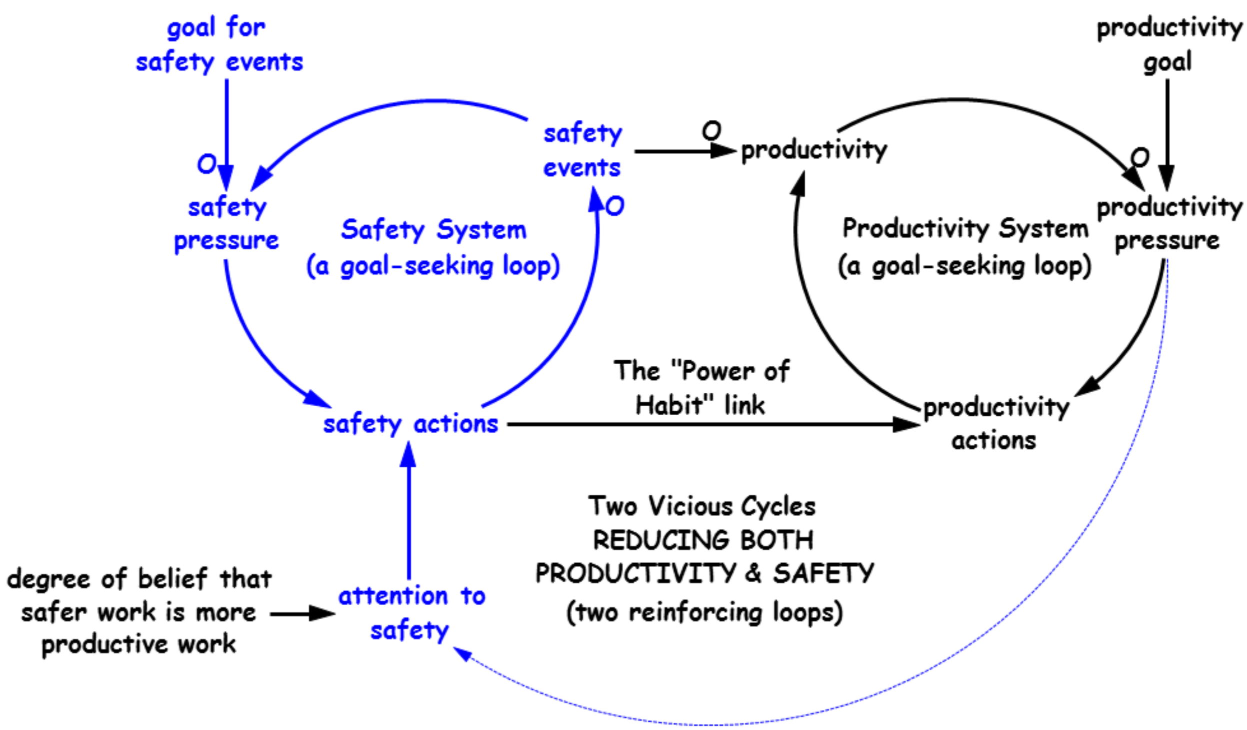

Figure 3: CLD gradually revealed to operators when presenting the mental models of system structure, and explaining the resulting behavior, of the MMRSP model.

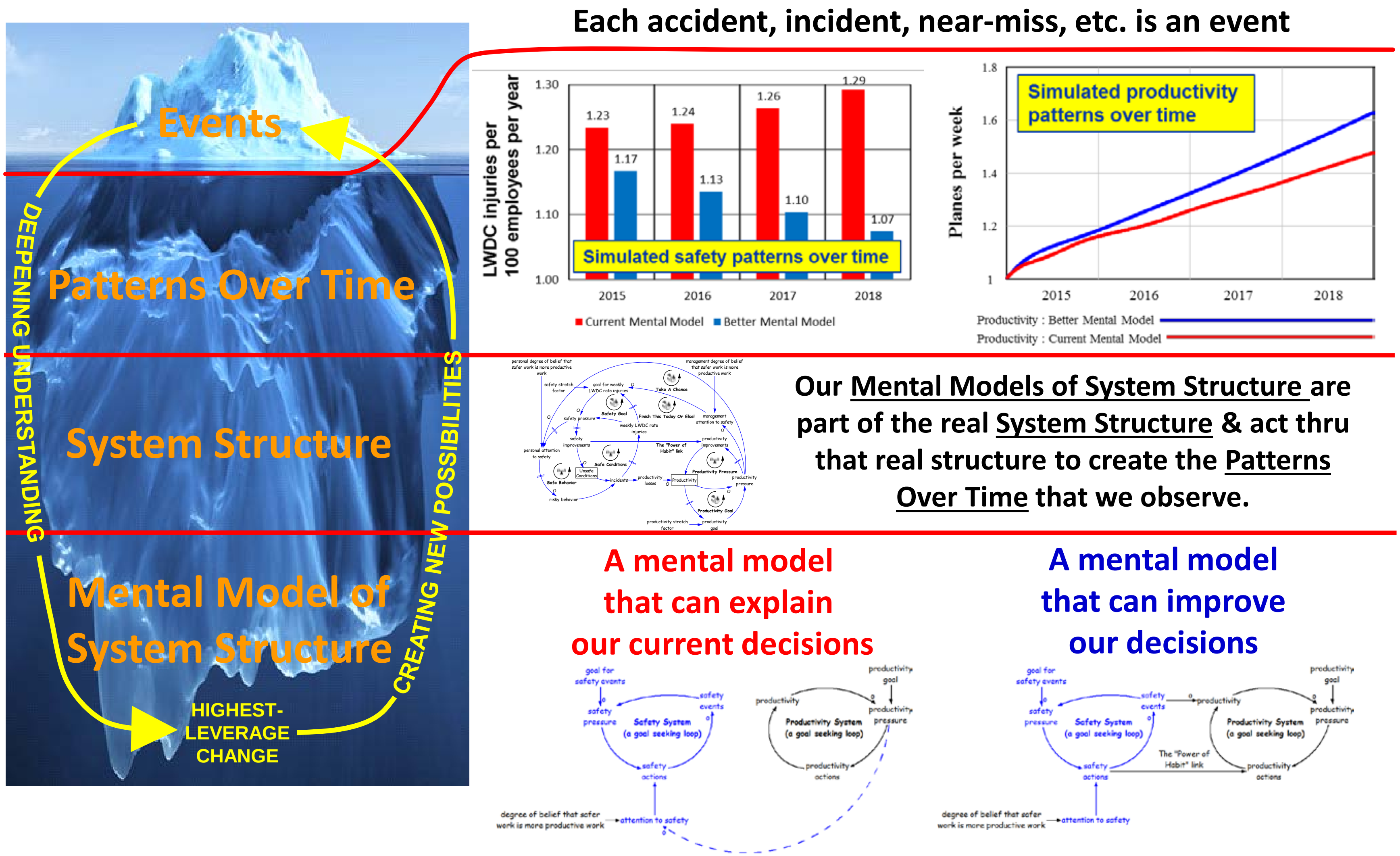

Figure 4: Explicitness in showing that the CLDs explained to the operators in the MMRSP project are not the full system structure relevant to the problem, but rather are only simplistic mental models of that structure, yet very important because they create the differences in the red (worse) and blue (better) patterns over time.

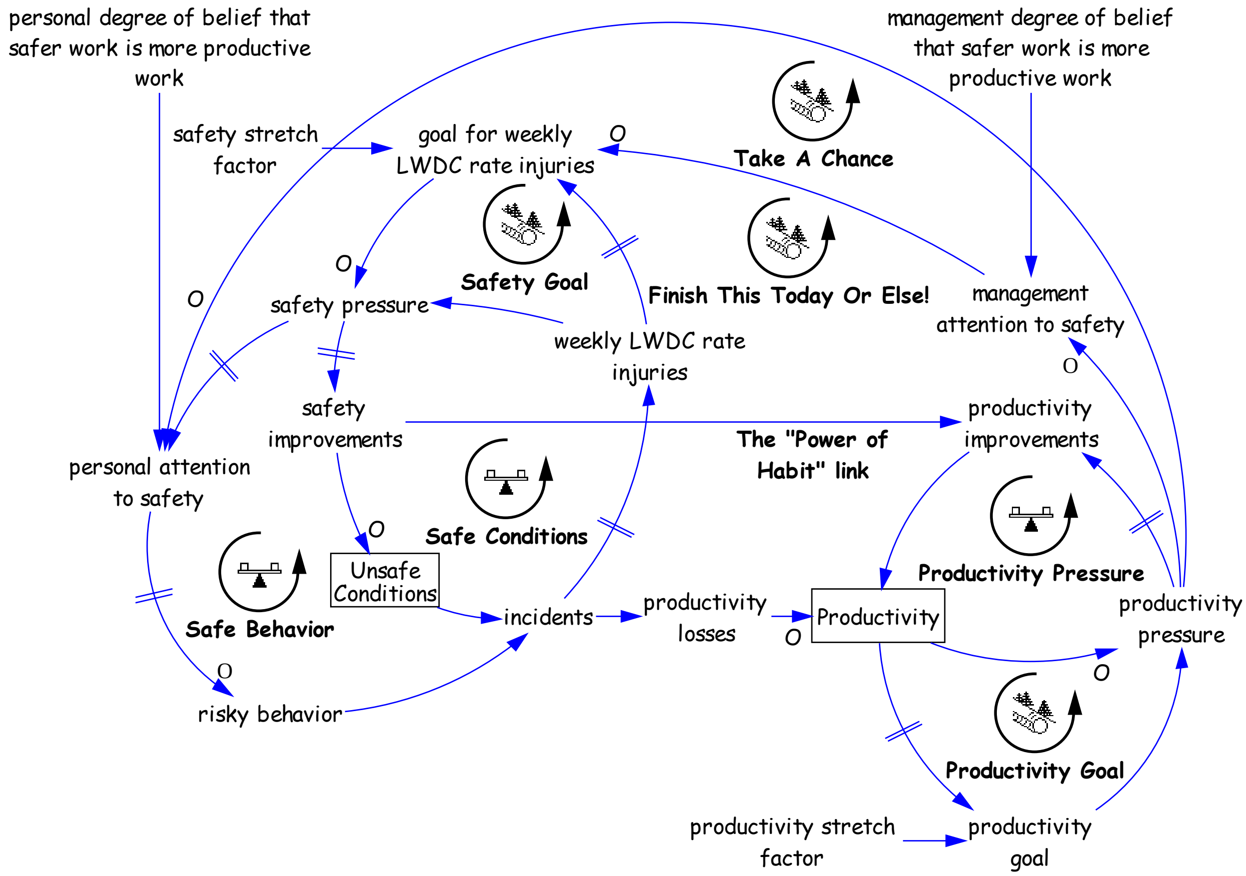

Figure 5: CLD of system modeled as relevant to the MMRSP problem, also shown in the System Structure layer in Figure 4.

Note that a full CLD of the system structure relevant to the MMRSP problem is shown in the System Structure layer of Figure 4 and is enlarged in Figure 5. Unlike Figure 3, Figure 5 is not shown to operators unless requested - except in the unreadable form in the System Structure layer of Figure 4. Comparing Figure 5 to Figure 2, it is easy to see that the MMRSP CLD in Figure 5 is at least as, and probably more, complex than the HOS CLD in Figure 2.[9] Sometimes system dynamicists refer to these full CLD representations of the system relevant to the problem being modeled as mental models. However, the CAMMSS approach frames mental models as even simpler representations of the system, e.g. Figure 3, which framing seems to align with most operators' notions. For example, note that Figure 3, as compared to Figure 5, does not:

1) distinguish between management and personal *degree of belief that safer work is more productive work*,
2) distinguish between safe conditions and safe behavior
3) Include the *Productivity Goal* and *Safety Goal* feedback loops

As mentioned earlier[7] in his description of how to improve performance in the Beer Game, Senge (1990, 2006, pages 47-51 in both editions) provides precedent for this way of viewing mental models, and hence for this CAMMSS approach.

[9] Actually, the HOS model is much more complex than the MMRSP model. However the HOS CLD that we created for the HOS model is somewhat less complex than the one we created for the MMRSP model.

System behavior contacts between a model and reality
Because the HOS project initial purpose was to test proposed mitigations to improve the safety of hazardous work operations in the service organization, it was almost a given to operators that, to have confidence in the mitigation simulation test results, the HOS model would first have to replicate the dynamics of past hazardous work projects. And indeed, the modeling team acquired historical dynamics from four or five such projects, and ensured that the HOS model could plausibly replicate the dynamics of each project. This gave the modeling team much confidence in the model. And since the team was executing such model behavior tests on so many different projects, there were instances in which the model could not replicate some aspect of a specific project's dynamics. These challenges led to improvements in the team's conceptualization and modeling of system structure, such that the model could replicate these dynamics, giving the modeling team even more confidence in the model's structure and behavior.

When presenting the HOS project to operators who weren't involved in the modeling effort, in addition to showing Figure 2 as previously mentioned, the team also chose to show alignment of simulations with the dynamics of past projects as an additional way to gain operator confidence in the model. Obviously, such simulations supply operators with additional non-discrepant "contacts between a model and reality." Operator audiences seem to take it as a given that such replications are helpful.

However, in the case of the MMRSP model, no attempt was made to find historical dynamic data to replicate. We typically only show and discuss with operators the simulation results in the Patterns Over Time layer of Figure 4. The y-axis values on these charts are plausible, but there is no attempt to replicate specific historical dynamics. And operators don't question us on this seeming omission! Why?

Again, we suspect the answer lies in the CAMMSS approach, which, because it is comparing alternative mental models, is fuzzy. Because operators know how fuzzy their own mental models are, they don't expect the simulation to replicate historical dynamics. It is enough for them that the simulation results are "in the ballpark" and plausible. It is enough that a better mental model can improve business performance dynamics, in this case safety and productivity dynamics. They don't demand calibrations to historical data, as so often happens when the CAP approach of comparing alternative policies is employed.

Conclusion: Three Cautions

1) <u>Our models must still be structurally and behaviorally sound.</u>
   No operators have yet requested an opportunity to see, much less understand and question, the full system structure of the MMRSP model. That the CAMMSS approach focuses on mental models probably gives rise to this lack of questioning. However, this lack

of questioning does not mean that CAMMSS models need not be as structurally and behaviorally sound as possible.[10]

2) We shouldn't make a full-scale switch from the CAP to the CAMMSS approach.
The CAMMSS approach will not work for all, and maybe not for many, problems we model. We are simply pointing out that the CAMMSS approach exists, and that it may provide more non-discrepant "contacts between a model and reality" in operators' minds. To the degree that it does, operators may have more confidence in the model, leading to more interest in the model and a stronger desire to share it with others, making it more likely that recommendations arising from the modeling project will be implemented. Our intent is to use the CAMMSS approach when it seems there are possibilities for improving system performance with improved mental models.

3) We should avoid using the CAMMSS approach to give operators the impression that "this is the way you think now, but you should think this way instead."
Rather, use CAMMSS to comparatively test alternative mental models for the purpose of encouraging people to discover, think about, share, challenge and improve their own contextual mental models of systems structure. The wording of the labels for the two mental models on the Mental Models of System Structure layer of Figure 4 is intended to reinforce this intention:
    a) "A mental model that can explain our current decisions," and
    b) "A mental model that can improve our decisions."
These two labels were intentionally chosen to avoid giving operators the impressions that "we know how you think" and "we know how you should think." Rather these two mental models are intended as tools for showing that we can discover, think about, share, challenge and improve our mental models.

References

Barlas, Yaman and Stanley Carpenter (1990) Philosophical roots of model validation: two paradigms. *System Dynamics Review* 6 (No. 2, Summer 1990): 148-166. ISSN 0883-7066.

Cooke, David L. (2006) Learning from incidents: from normal accidents to high reliability. System Dynamics Review Vol.2, No. 3, (Fall 2006): 213-239.

Forrester, Jay W. (1973) Confidence in Models of Social Behavior with Emphasis on System Dynamics Models. System Dynamics Group Working Paper D-1967. Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA. http://www.systemdynamics.org/mit-sdgroup-literature-collection/

[10] We're confident that sooner or later someone in the company will seriously challenge the MMRSP model, and we are prepared. We developed a paper fully documenting the model that we can share with such challengers. Probably there are shortcomings in the model, even in relation to its intended purpose, as there probably are with most models. Certainly we could, and should, do more testing of the model. However, that the CAMMSS approach more readily conveys that the model is for exploration and learning, we feel well prepared to face challenges in this spirit of exploration and learning.

Forrester, Jay W., and Peter M. Senge. (1980) Tests for Building Confidence in System Dynamics Models. In System Dynamics, ed. A. A. Legasto, Jr., J. W. Forrester, and J. M. Lyneis, 201-228. TIMS Studies in the Management Sciences. Vol. 14. New York: North-Holland.

Senge, Peter M. (1990, 2006) The Fifth Discipline: The Art & Practice of the Learning Organization. Currency Doubleday.

