A Simulation Game to Promote Systems Thinking and Cross Organizational Collaboration in Government Agencies

William H. Ryder
Retired, Formerly with U.S. Department of Defense
6735 Allview Drive
Columbia, MD 21046
410-730-5541
onebyke2ryders@gmail.com

Abstract: For the past 7 years a management training organization within the U.S. Department of Defense has offered an experiential workshop to train leaders in systems thinking and collaboration skills. Based on a table-top dynamic model resembling a board game, the workshop offers participants opportunities for situational learning and consequent recognition of their own hidden mental models. This paper describes the underlying dynamic model, the system dynamics structures supporting it and shows how the workshop uses these structures to introduce system dynamics concepts and foster collaborative skills. Lessons learned for presenting this and similar workshops are discussed and a measure of acceptance of this workshop is presented.

Keywords: Serious games, Training games, Capability Trap, Collaboration, Mental models, Leadership training

After the September 11, 2001 attacks on the World Trade Center in New York and the Pentagon near Washington D.C., a blue-ribbon panel, popularly called the 9/11 Commission, formed to investigate why the intelligence system of the United States failed to alert the country to the attacks. The panel issued its now famous 9/11 Report in July of 2004 (9/11 Commission, 2004). Among the 9/11 Report's findings, two themes stand out:

1) The cultures of the various agencies within the U. S. Intelligence Community (IC) work to inhibit the sharing of information across agencies. Specifically, "There are no punishments for *not* sharing information. Agencies uphold a "need to know" culture of information protection rather than promoting a "need to share" culture of integration." (9/11 Commission,2004, pp 417)
2) There are six problems, which, taken together, undermine the agility of the overall system. The report identified them as follows: structural barriers, lack of common standards, divided management, weak capacity to set priorities across agencies, too many jobs given to some of the agency heads, and a system that is "too complex and too secret for anyone to understand". (9/11 Commission, 2004, pp 408-410) In other words, the IC struggles with some of the same authority, and process dynamics issues faced by other large, complex organizations.

Since 2004, initiatives in leadership training, both at the IC member agency level and at the national level have worked to address these problems. One such initiative became an experiential workshop designed for training future executives in the IC. Central to the workshop is a table top simulation resembling a board game that loosely models operations in a fictitious community of generic intelligence agencies.

This paper describes the design goals of the workshop, the structure, dynamics, and implementation of the simulation and workshop, and summarizes lessons multiple workshop sessions have taught the designers. The paper intends, by giving an example, to assist others who would develop similar experiential workshops within other story line contexts.

Background

The use of table top simulations to illustrate system dynamics principles has ample precedent. The Beer Distribution Game (Senge 2006, Sterman 2000), The Manufacturing Game® (Sterman 2000, Ledet and Paich 1994); and Friday Night at the ER® (Gardner, 1994) exemplify the genre. Each simulation teaches basic principles of system dynamics, and surfaces the hidden mental models of reality held by the participants, thereby enabling participants to change the way they think about their own work environments. Of these, The Manufacturing Game® (TMG) seemed best suited as a starting point for creating a new simulation game having a government story line. Subsequent discussions with Winston Ledet, co-creator of TMG, suggested that a game-based intervention could overcome the pervasive resistance to change prevalent in complex organizations. Separate discussion with Professor John Sterman of MIT identified the capability trap (Repenning and Sterman, 2001, Repenning, N. and J. Sterman 2002) as an ideal dynamic structure to underpin the new game. It is simple, it deals with the tensions of making allocation decisions, and it explores resistance to long-term investments for improvement. Morrison (2012) has recently analyzed the dynamics of this structure under assumptions of linearity, and Sterman (2014) and Lyneis and Sterman (2014) apply it to the challenges of sustainability.

Design Goals

Because the new workshop would train managers aspiring to the executive level, it had to meet specific design goals.

1. The workshop experience had to underscore the importance of understanding the larger system of which individual decision-makers are a part. Specifically, the workshop had to create an awareness of basic system dynamics concepts as well as the importance of mental models in decisions. The workshop had to render its system dynamics content so accessible that non-technical managers could embrace and pursue it on a personal level when back on the job.
2. The workshop had to surface the dysfunctional behaviors and assumptions that inhibit collaborative behavior and offer participants the opportunity to recognize and to practice successful collaboration.
3. The simulation game within the workshop had to be sufficiently compelling and fun to play to command the attention of habitually distracted managers for extended time periods sufficient to allow the participants to explore alternatives and acknowledge the shortcomings of their own perception. The simulation storyline had to resemble participants' real jobs sufficiently closely that they would identify themselves with it. Video game design analysis (Koster 2005) identified necessary elements to meet this requirement.
4. The simulation would not use computers for calculations. Hidden states or unseen computer operations might offer participants a mental scapegoat to blame for poor performance. The

outcome of the simulation had to be closely and obviously tied to the decisions made by the participants themselves.

5. The workshop could not cost very much on a per-session basis. The game materials would be produced at low cost and facilitators would be mostly in-house volunteers drawn from the pool of past workshop participants.

In using volunteers as facilitators, the workshop would require a deep bench of trained facilitators to ensure a full cast for each session. This led to a companion "train the trainer" workshop whereby recruits would learn to be fully capable facilitators. Much of the material and content for this latter workshop was designed under contract by Ledet Enterprises, Inc.

Simulation Structure

A typical workshop consists of four teams of six participants. Each team operates its own board-game simulation of a notional government intelligence agency. The simulation proceeds by synchronously executed turns. Each of the four independent intelligence agencies attempts to control its own growing terrorist threat by producing intelligence. Left unchecked, the terrorists advance their nefarious plans every turn to reach a threat level threshold, indicating a successful terrorist act. Intelligence, when produced by an agency, decreases that threat level incrementally. If all four agencies can avoid a terrorist act for the duration of the two-day workshop, all workshop participants receive a fabulous prize of chocolate. However, should a terrorist cell reach the threshold in any of the agencies, the simulation terminates for all agencies and nobody receives chocolate. Instead, a facilitated discussion occurs and the game boards are given additional resources to allow participants to try again with, we hope, their newly gained knowledge.

Participant Roles

Each agency has a team of six participants, one for each sector of their agency. Of these, five have key roles directly connected with running the agency. The three chiefs of access, processing, and analysis manage intelligence production operations, while the two chiefs of human resources and infrastructure act to enable those operations by providing personnel and infrastructure. This leaves the chief of liaison to deal with communication with other agencies. The liaison chief meets with other liaison chiefs once per turn to exchange sensitive information needed to produce intelligence. When liaison chiefs are not conferring with other liaisons they perform bookkeeping tasks for their respective agencies.

Interdependent systems

The simulation consists of three systems: 1) the production system, 2) the personnel hiring and training system, and 3) the infrastructure acquisition and maintenance system. Workshop participants manipulate these systems by making personnel and infrastructure allocation decisions. Personnel decisions include assignments and training. Infrastructure decisions include installation allocation and maintenance. In making these decisions, participants continually face the classic "operate now – better before worse" or "invest for later – worse before better" conflict of the capability trap. We shall describe the production system first, then review the generic capability trap structure and show how it generates the dynamics of the personnel and infrastructure systems.

Production System

The production system occupies most of the game board, shown notionally in Figure 1. The production system moves sensitive information, represented by ordinary playing cards, from an "access" sector through a "processing" sector and into a secret "known database". The information card movements are limited by the capacities within "offices" responsible for moving them. Able people with working infrastructure form the capacity in the offices. The cards remain face-down until they reach the database where their faces are exposed for analysis. Offices in the "Analysis" sector match the card face values in the database against individual "intelligence requirements" to form the final intelligence. A set of requirements is a 4 x 4 array of playing card face values needed by exogenous intelligence customers who prosecute terrorists. Shown in Figure 2, a set of requirements resembles a Bingo card. Intelligence becomes actionable (and reduces the terrorist threat) when the analysis offices match enough database information cards to needed information cells in the array to form a complete column, row or diagonal, that is, a "Bingo", and declares it as such by shouting "Bingo!". Declaration of "Bingo!" causes customer use of the intelligence to counteract the terrorists and reduces the threat level by an amount related to the difficulty of the requirement. Unfortunately, the customers' actions in suppressing the terrorists also reveal the secret information in the known database to the terrorists, thus rendering it useless for future intelligence production. Therefore, after a Bingo, all information cards in the known database recycle to the access sector to represent new information. Card matching is limited by the capacities of the analysis offices responsible for matching.

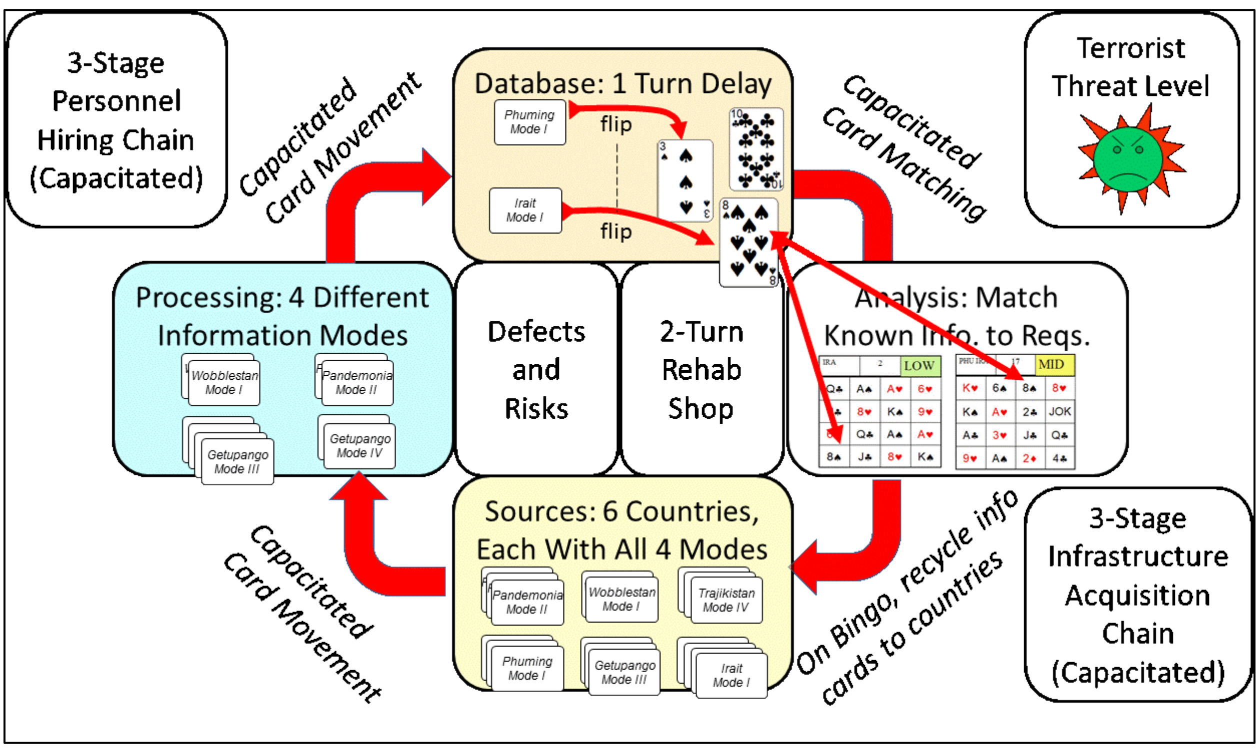


Figure 1: Game Board Layout and Path of Information

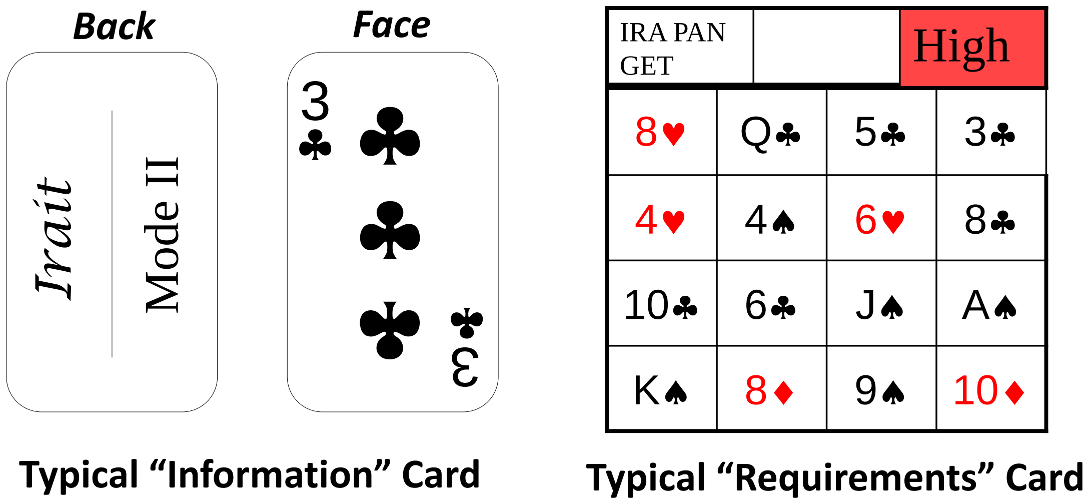


Figure 2: Components of Information

Capability Trap

The personnel and infrastructure systems each contain a capability trap. To illustrate the generic capability trap structure and dynamics, we briefly rephrase the cited works of Repenning, Sterman and Morrison here. We then map the generic structure into the two systems. The open loop structure of the capability trap is shown in Figure 3.

A backlog of work (measured in *tasks*), fed by an exogenous process (*tasks per time*), must be reduced through a capacity-limited process for doing the work (*tasks per time*). The capacity (*tasks per time*) results from a stock of resources performing the work (*resource units*), with each resource unit having an inherent capability (*capacity per resource unit*) to do the work. For example, if the resource is personnel time, the capability could be a function of peoples' alertness, level of expertise, readiness, productivity, or level of health. Capability drains away through a "losing capability" process (*capability per time*) dependent on the current level of capability per resource. Resource units within a limited total pool of resources can occupy either an "in-production" state where they do work or an "in improvement" state where they build capability. A decision-maker adjusts the allocation of resource units between the two states so as to maximize the actual capacity to do work. This allocation of resources and the characteristics of the losing and gaining processes determine open-loop equilibrium values for both the resource stocks and the capability stock.

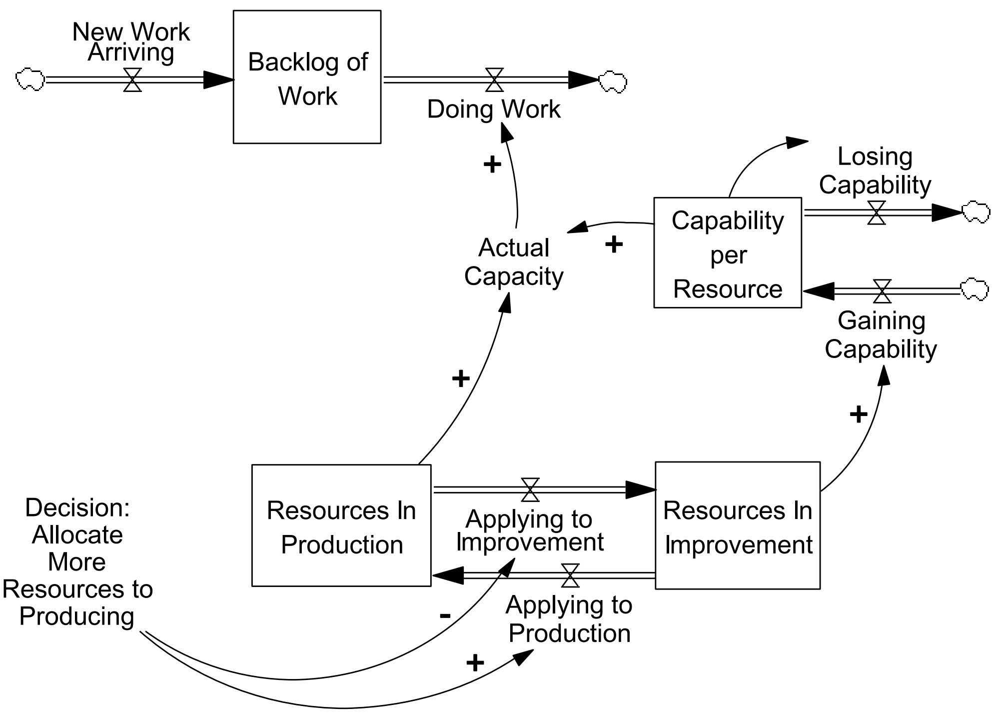


Figure 3: Open Loop Stock and Flow of Capability Trap

If no resources occupy the in-production state, there will be no capacity. Conversely, if all resources are in production, the stock of capability will leak away and drive capacity to zero. Some optimal allocation between these extremes will yield the maximum equilibrium capacity value. Morrison shows that with capacity defined as resources multiplied by average capability, and with first order linear losing and gaining processes, the open-loop equilibrium capacity, as a function of fraction of resources in production, will be unimodal and skewed toward production, as shown notionally in Figure 4. The exact position and value of the optimum will depend on the time constants of the losing and gaining processes. Specifically, letting x represent the Fraction of Resources in Production, t_G the average time it takes for the gaining process, and t_L the time it takes for the losing process, the open-loop equilibrium capacity is given by:

$$Capacity = A\frac{x(1-x)}{b+1-x} \qquad where\ b = \frac{t_G}{t_L}; \quad A\ is\ constant$$

Thus, if it takes much longer to gain capability than it does to lose it, the optimum resource fraction will be slightly higher than 0.5. Conversely, if capability is lost very slowly relative to the gaining rate, the optimum fraction will be closer to 1.

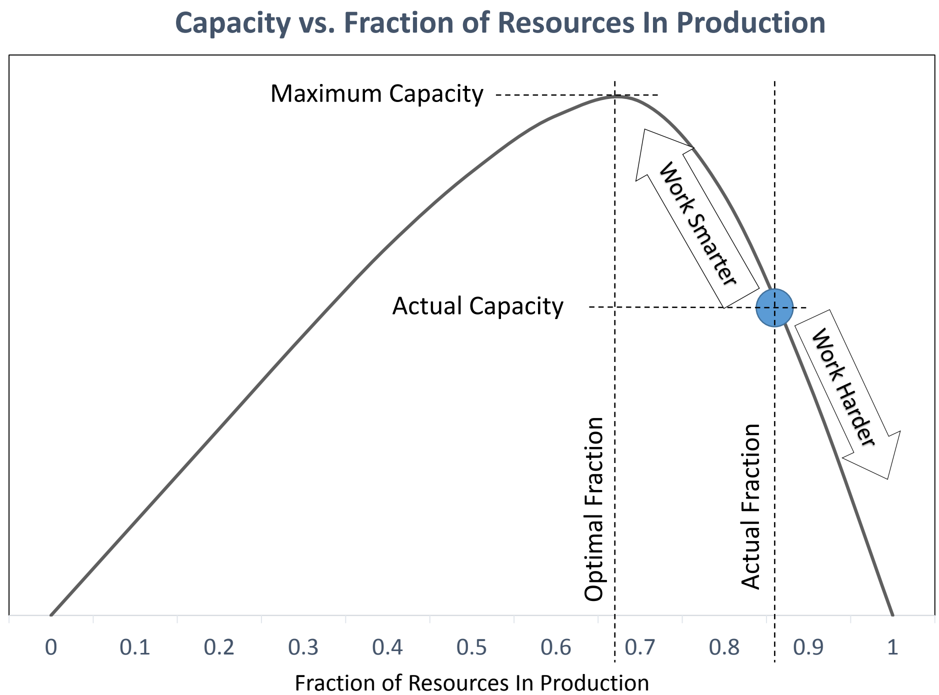


Figure 4: Equilibrium Performance Curve of Open Loop System

The actual capability of resources is frequently difficult or expensive to measure. Moreover, the rates of the losing capability and gaining capability processes may vary with time and have unknown parameters. Therefore, the decision-makers typically do not know precisely how much capability exists at any given time, nor do they know the optimal fraction to attain the maximum capacity. Instead, the decision-makers adjust the fraction of resources in production incrementally by reacting to the size of the backlog. This causal influence creates the action of the capability trap by closing two feedback loops. The sense of the loops depends on which of two decision policies the decision-maker selects at any given time. The two policies, "work harder" and "work smarter", are shown in Figure 5 and Figure 6 respectively, with feedback loops closed. Figures 5 and 6 differ only in the outcome of the decision affecting the fraction of resources in production.

To see the operation of the capability trap, imagine that the system is operating with more resources in production than optimal, as shown in Figure 4. We assume that a large backlog with significant schedule pressure induces the decision-maker to adjust the fraction of resources in production. Clearly, the correct decision would be the work smarter strategy. This would move the fraction in production lower and closer to the optimal point. The eventual increase in equilibrium capability would more than make up for fewer resources in production. But this worse-before-better choice would incur a short term loss of capacity while resources build capability. Instead, the decision-maker, under pressure and not knowing the position of the optimum point, will employ the work harder strategy. This moves the system further away from optimum and decreases capacity as capability approaches the new lower equilibrium. As a result, both the backlog and schedule pressure increase. The cycle repeats until either the decision-maker realizes that the work harder policy makes things worse, or the capacity collapses completely. The now dominant R1 "Seek Production" loop in Figure 5 pulls the organization into a permanent state of low performance.

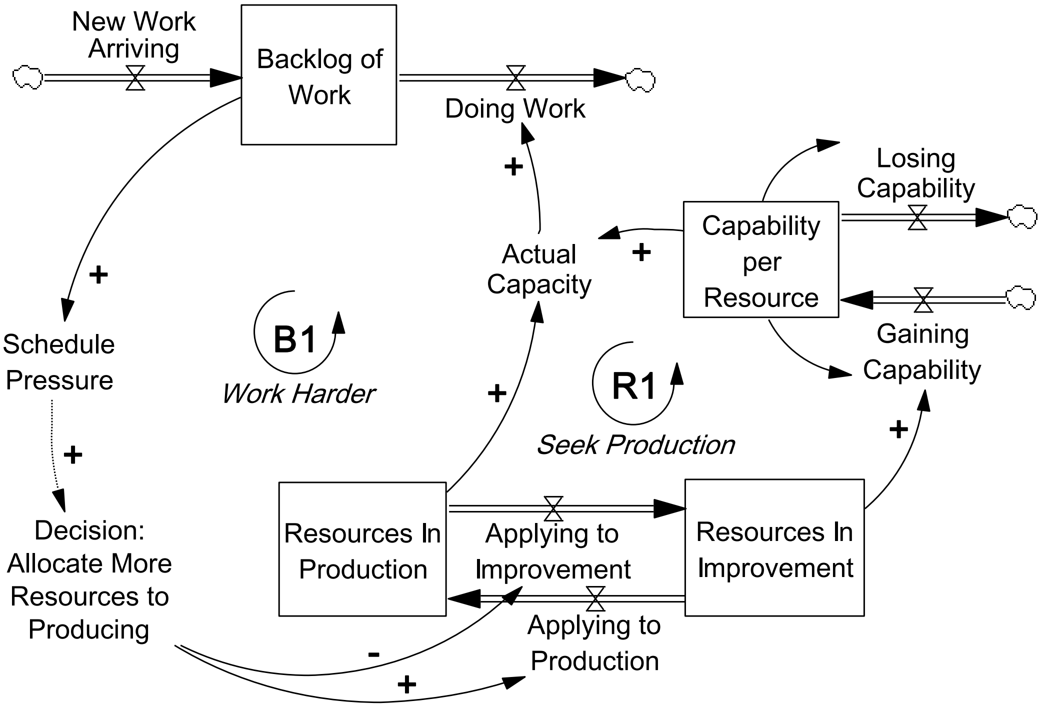


Figure 5: Closed Loop Work Harder Policy

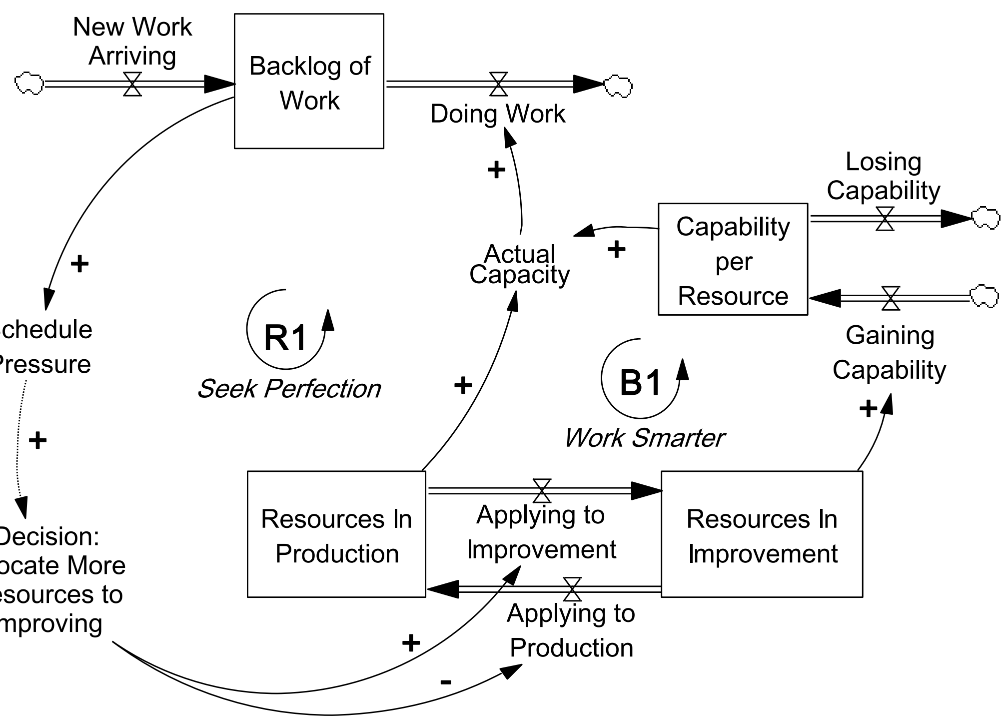


Figure 6: Closed Loop Work Smarter Policy

Alternatively, were the system operating with fewer than optimum resources in production, the work harder strategy would prove the correct choice. It would move the fraction in production higher, closer to optimum, and increase the capacity. This could work down some of the backlog and ease schedule pressure. However, if the backlog were too high to begin with, schedule pressure would remain high and the work harder policy would be repeated. The system would be driven inadvertently past an invisible tipping point and into the trap. Schedule pressure coupled with the belief that working harder solves the problem creates the trap. Conversely, if the decision-maker believed that work smarter was always the best response to schedule pressure, the trap would work in the opposite direction, driven by the “seek perfection” loop in Fig. 4, via repeated application of work smarter decisions. Thus, the optimum point in Fig 2 is unstable. Either strategy, applied persistently, will lead to a state of diminished performance relative to optimum. Readers may recognize this structure as the “fixes that fail” systems thinking archetype (Senge et al.,1994).

Personnel and Infrastructure Playing Pieces

The generic capability trap appears in the simulation game through attributes of personnel and infrastructure playing pieces. The construction of the two types is shown in Figure 7. Personnel pieces consist of standard binder clips with bent handles. The clips represent authorized job positions or billets. Each filled position binder clip holds a laminated paper card representing a personnel cohort. The card edge holds a paper clip over inscribed numbers to show the number of turns before that cohort retires. Yellow strips of sticky paper attached to the top of the card display the ability level for that cohort. Infrastructure pieces consist of stacks of colored plastic poker chips. The one blue chip at the bottom represents the infrastructure itself, red chips in the stack are defects within that infrastructure unit, and a variable number of white chips represent that unit’s remaining expected useful life in turns.

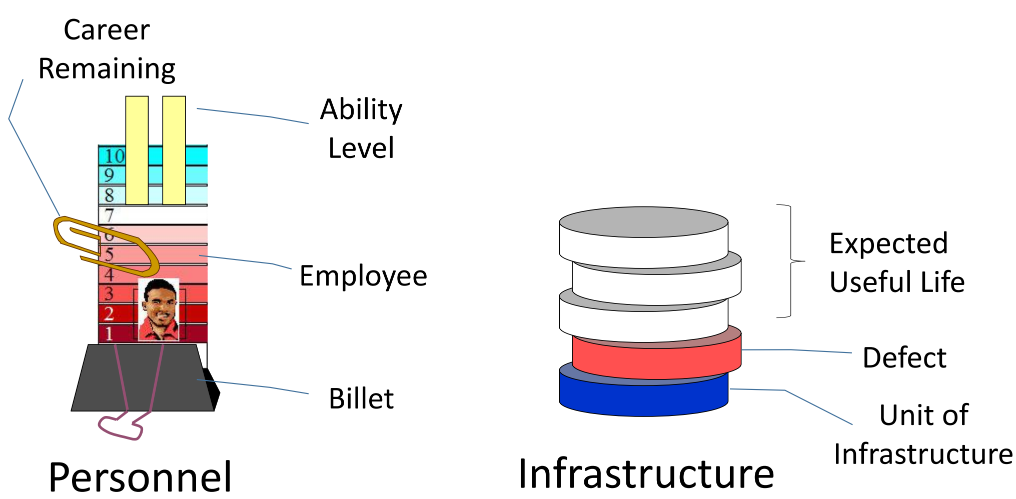


Figure 7: Tokens for Personnel and Resources

Personnel System

Personnel units are assigned by participants to populate offices on the game board. Each “At-Work” personnel unit generates capacity equal to their individual ability level, provided they also each have an infrastructure unit. The personnel in offices are considered to be “At Work” unless participants have put them into an "In-Training" state to increase their ability by placing them on their sides. Personnel units in training do no work regardless of their current ability. Each turn, participants decrement the paper clip turns-until-retirement counters by one. Upon retirement, personnel units leave their offices, lose their ability, and become "Unfilled Positions" in a capacity-limited three-stage hiring chain. Once hired, new personnel units begin with a career length of 9 turns, no ability and are assigned by participants to any of the offices. Participants may also reassign current personnel units to another office or agency, but the transferred units lose one unit of ability to reflect the retraining required for the new job.

The open loop capability trap structure in this system appears in Figure 8. The decisions is shown in bold font. Comparing Figure 8 with Figure 3, we see that the average ability of a personnel unit plays the role of capability per resource. Two losing processes, retiring and reassigning, decrease ability, while training increases it. People in offices are the resources. Those at work are in production and those in training are in improvement. The terrorist threat level represents the backlog, turns until "game over" (not shown) is schedule pressure, and the producing intelligence process equates to doing work. Participants' perceptions (not shown) of the pressure drive decisions balancing hiring versus producing, as well as rates of training, and reassignments. We have assumed that all personnel are enabled by an operating infrastructure unit.

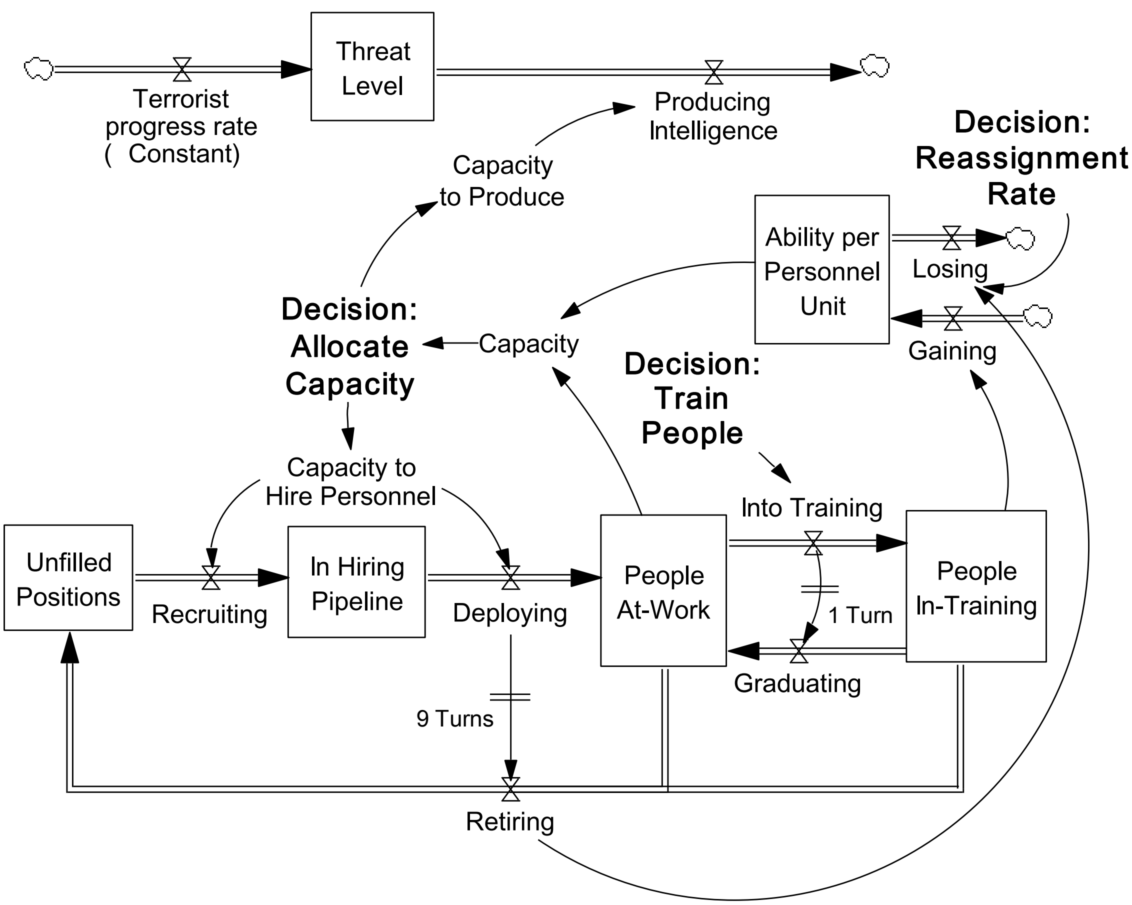


Figure 8: Open Loop Personnel System

Two failure modes can be observed during a typical workshop. First, if participants fail to hire personnel faster than the personnel retire, the overall workforce level will fall toward a low equilibrium level insufficient to defeat the terrorists. It takes 9 turns of constant hiring rate to reach equilibrium, so this peril is not immediately evident. Second, if participants hire personnel at the required rate but do not train enough of them to offset the loss of ability through retirements and reassignments, the average ability level will fall and undermine capacity. Again, the terrorists will win. By allocating enough capacity to hiring, training enough personnel units each turn, and reassigning personnel sparingly, successful participants grow a workforce having enough capacity to control the terrorists. Unfortunately, most participants believe that, when disaster looms, production trumps all other "non-essential" activities such as training and hiring. They apply the "work harder" policy to close the trap loops of Figure 5.

Most participants require help to recognize the stocks in this system, understand their operation, and detect the capability trap unfolding. The workshop provides brief training sessions given by facilitators in the form of pre-scripted skits. They explain the above model, discuss the nature of stocks, and run a computer simulation of the model to show the decision values needed to stave off disaster. Participants can, but do not always, apply these beneficial strategies in subsequent turns.

Infrastructure System

Infrastructure units are acquired via a three stage capacity-limited acquisition process similar to the personnel system hiring process. Newly acquired infrastructure units are installed by participants into offices where they enable personnel to do work. They accumulate "defects" as they operate. The infrastructure units will either break down from the defects, or be taken offline for rehabilitation before they break. If infrastructure units do break down, they leave the game board and must be re-acquired. Alternatively, if participants elect to take them out of service and move them to rehabilitation, they are unavailable to support personnel units for two turns. Thus, the defects that accumulate on the infrastructure units constitute the overall destructive process for infrastructure. Infrastructure breaks down immediately with two defects, but can continue operating until its useful life expires with just one. Infrastructure with no defects, being good as new, receives a fresh round of useful life when its current useful life expires. Participants distribute arriving defects onto working infrastructure and decrement the expected lifetime counters each turn. Like personnel, infrastructure must be brought onto the game board as fast as it leaves. If infrastructure does break down, people and infrastructure units must be allocated to the acquisition pipeline to re-acquire it.

Defects stream onto the game board from seven specific causes or sources. Each produces partial defects at a characteristic rate. The sources represent the common maladies and actions of complex organizations and add realism to the story line. Participants quickly associate the severity of the defect stream with the availability of working infrastructure.

The open loop capability trap structure in the infrastructure system appears in Figure 9. It models the breakdown of infrastructure in a similar fashion to equipment failure in The Manufacturing Game®.

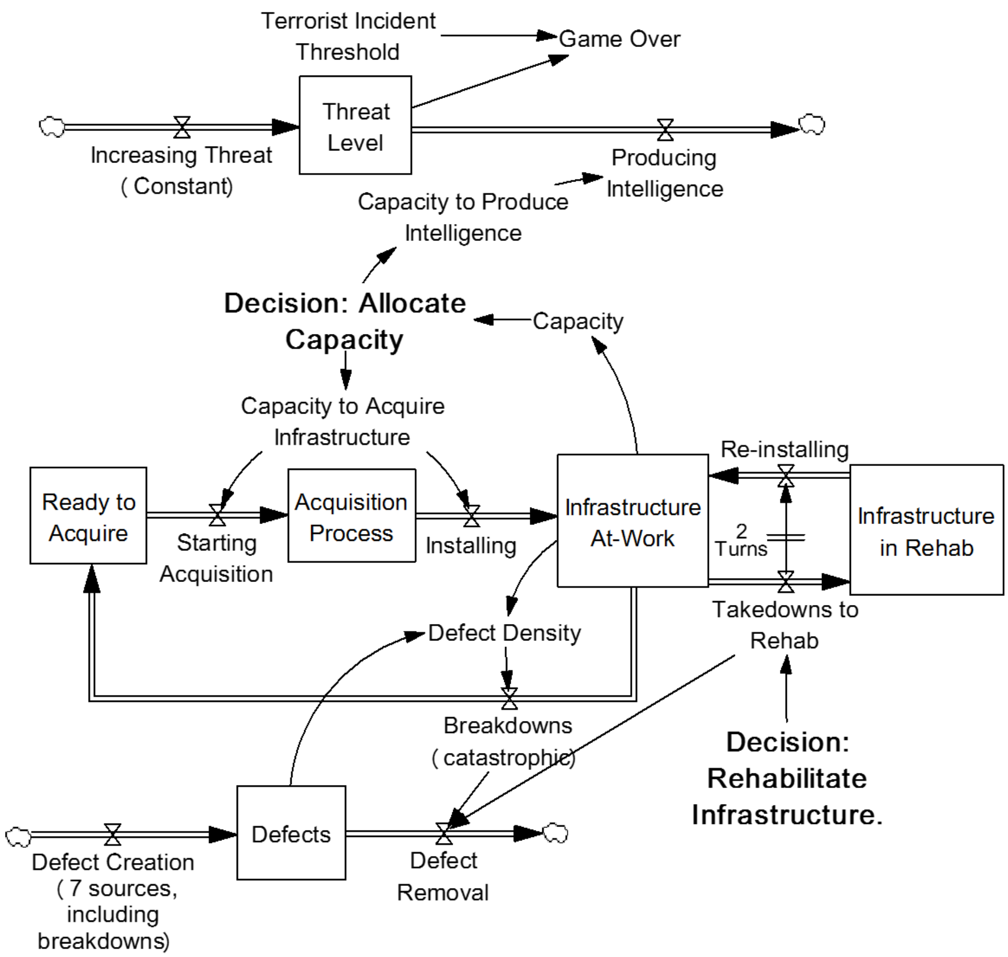


Figure 9: Infrastructure System

Like personnel, infrastructure must be brought onto the game board as fast as it leaves. If infrastructure is leaving through breakdowns, people and infrastructure units must be allocated to the acquisition pipeline to re-acquire it. Here, Capacity is the presence of working infrastructure units. “Capability” is the collective ability of all the working infrastructure to absorb new defects without breaking. Only infrastructure units with no defects can do this. Once all units have one defect, a wave of breakdowns will occur as defects arrive without spaces to absorb them. This sudden, unexpected failure surprises many participants. If, instead, participants place one unit of infrastructure into rehabilitation for each new defect arriving in a turn, the stock of defect-free units will stabilize and no breakdowns will occur. In this system, working harder means to keep damaged infrastructure on line until it breaks. Following terminology in The Manufacturing Game®, we call this the “Reactive Mode”. Working smarter means to put damaged infrastructure into rehabilitation before it breaks. This is the “Planned Mode”. As with personnel, schedule pressure in the form of terrorist advances drives the participants’ allocation and rehabilitation decisions.

Participants may alter the defect creation rates by assigning personnel units to do “improvement projects”. This activity removes two personnel units per project from their offices for two turns. Each successfully completed project will reduce the rate of defect creation for one of the defect sources.

When the simulation starts, all agencies suffer from an almost overwhelming defect rate. However, agencies that do improvement projects consistently will, over the course of about 15 turns, reduce the defect stream sufficiently to operate their agencies efficiently. However, agencies doing too many improvement projects at once will not produce enough intelligence to suppress their terrorists.

Metrics

Once each turn, participants measure key variables in their agencies, such as the size of the workforce, their state of readiness, and the health of infrastructure by counting the positions and conditions of the playing pieces. These metrics, when entered into a spreadsheet, are graphed and displayed to the room. A facilitator leads discussions about the importance of metrics, the concepts of leading and lagging indicators and how models define the meaning of metrics.

Workshop Structure

Collaboration

The simulation is calibrated so that a single agency, once it solves its resource issues, can just stay even with terrorist threat growth using capacity within their agency. However, there are community-wide strategies available within the rules that will soundly defeat the terror threats at all agencies. To recognize and employ these strategies, participants must shift from a focus on their own agencies into a community focus with community-wide decision making. In effect, they must do in the game world what the 911 Commission asks the real Intelligence Community to do in the real world. This transition requires new mental models that embrace a system-wide view and willingness to sacrifice individual advantage to community benefit. Multiple workshops have revealed a common sequence of four development stages for collaborating. We label these stages Competition, Bartering, Cooperation, and Full Collaboration.

In the Competition stage, participants view the simulation as a competition despite the cooperative nature of the chocolate reward system. Participants find themselves competing for scarce resources within their own agency. Arguments over people and infrastructure units quickly develop until tables evolve a decision making process for their agency. The operating mental model values individual effort above effort across the table. Participants in this mode usually recognize their own lack of collaborative effort.

After a few turns, participants begin trading resources or promises of resources with each other. This behavior first emerges within agencies, but frequently becomes the norm across agency boundaries. Its signature is a *quid pro quo* negotiation where both parties give something and get something. This Bartering stage can persist throughout the workshop. If asked if they are collaborating successfully, participants in Bartering stage invariably answer in the affirmative.

When resources become so scarce that a needy party cannot offer anything in return, the Cooperation stage may emerge. Participants will give other participants in other agencies resources without expecting anything in return, providing that the needy party has asked for them. Such person-to-person deals usually grow in frequency to become standard procedure. In Cooperation stage, strong agencies will freely donate their completed, but undeclared, bingo cards to weaker agencies to prevent terrorist attacks at the weaker agencies. Most workshops conclude in this mode. The participants feel they have reached a high level of collaboration.

Occasionally, some individuals will recognize the economies of production the community can obtain by adopting a centrally coordinated Full Collaboration. They persuade the other participants to form a regulating body to determine which agencies will declare Bingo, when they will do so, and what information sources they will work. Such room-wide coordination sometimes reveals an alternative more efficient means to share information across agencies that is within the game rules. These actions combine to multiply the rate of card matching and, hence, intelligence production. This wholesale transformation of mental models is rare and has only happened spontaneously a few times in the history of the workshop.

Workshop Facilitators

A "cast" of four facilitators staffs the typical workshop. A dedicated Game Caller introduces the workshop, explains the rules of the simulation to the participants and maintains the pace of the simulation by announcing the turns. Two Table Facilitators circulate among the agency tables during play to answer questions about the rules and to detect any misinterpretations. A Game Master observes play, adjudicates any rule disputes, directs the other facilitators, and decides what interventions will take place during the simulation. Two of the four facilitators deliver just-in-time training to participants at intervals throughout the workshop. The skits, illustrated in Figure 10, are delivered in character with costumes. We find that this mode contributes to the alternate reality of the simulation and provides comic relief from the stress experienced by participants. The skits deliver system dynamics instruction at the precise time participants may perceive they need something like system dynamics.

Figure 10: Just-In-Time training from Dr. Psycho Metric

The usual facilitator group relies on a larger contingent of trained facilitators to form casts for successive workshop sessions. To become a facilitator, an individual must first play the game in a workshop, then volunteer for a four-day in-depth course of system dynamics and facilitation skill instruction. The facilitators' course might be thought of as the true leverage point of this systemic intervention. The workshop itself identifies and recruits those individuals with the ability and desire to improve their real life environments. By taking the facilitators' course and giving subsequent workshops, they become the exponents of change in the larger community.

Events During a Workshop

When participants arrive, they find the four agency tables configured identically as though they had operated under previous leadership for a long time. Facilitators encourage them to select a role they do not perform in the course of their real life work. Following a welcome speech, Game Caller explains the very complex rules of the simulation and steps participants through the first turn. Play then commences and continues without break except for the just-in-time instruction skits. Various intervention events may occur as determined by Game Master. These events are designed to regulate the pressure participants feel and thus to maintain an uncomfortable but motivating pressure level. By the end of the first day, the dynamics of the game have typically produced a low point of performance and participants go home to think about why things are going badly. Upon returning the next morning, participants usually have many ideas and start experimenting with ways of collaborating more effectively. To reinforce this frame, Game Master conducts a benchmarking exercise where participants circulate among agencies while a stationary representative from each agency explains that agency's strategy. Many insights emerge from this exercise. Play resumes, the pace of turns quickens and players develop a confidence that they can win the chocolate. Just when things are improving, Game Master announces a community wide reorganization. Half of the participants are reassigned to other tables and those that stay at their original agency assume different roles. The disruptive effects of this event generate some of the deepest learnings of the workshop. Local cultures that emerged during the first day become visible and dysfunctional behaviors are revealed. Play then continues until the afternoon of the second day to allow participants time to connect all they have experienced with successful play. A facilitated room-wide de-briefing led by Game Master follows. There, participants describe their feelings during the game, major events that happened to them, and what they perceive they learned during the game. The discussion helps the participants identify the experiences in the simulation with opportunities for improvement in their real jobs. An optional action-planning segment may follow the facilitated discussion. Facilitators contact participants to review progress on the submitted action plans three weeks after the workshop. This sequence of events is shown diagrammatically in Figure 11.

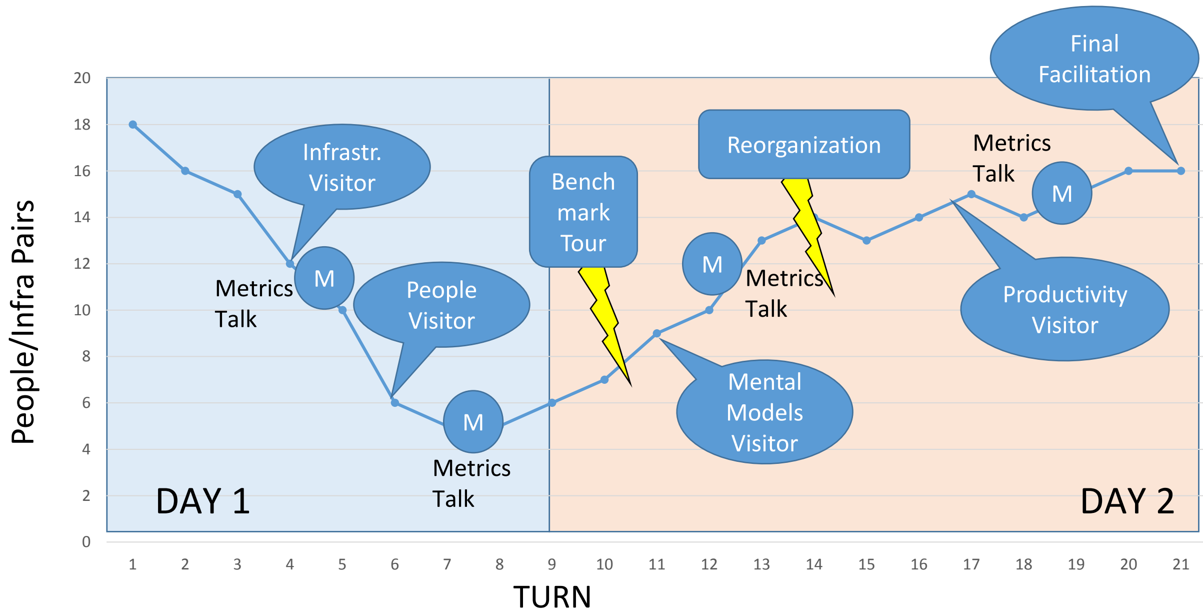


Figure 11: Events Within a Workshop

Learnings and Insights Gained Over Many Workshops

This workshop has grown and changed with each session. Some new and unexpected interpretation of the rules, some key suggestion for improving the game, some better approach to facilitation always seems to emerge. We highlight a few of our own experientially gained learnings here.

Game Insights for Designers

1. The game should challenge participants but begin and end with success. Early success pulls the participants into the game's alternate reality. Ending success brings closure to the experience and enables participants to carry learnings back to the real world.
2. The basic premise of the game must be simple. Participants lose interest with increasing complexity unless they see a strong parallel with their work reality. There is great benefit in choosing a simple underlying system dynamics model to begin the game design.
3. Facilitators must be open to new interpretations of game rules. The participants operate in a reality that is partly the reality created by the game's story and partly that of their perception of their real world jobs. Sometimes participants will uncover important features of reality that make the simulation more believable. The question "What would happen in real life?" answers most participant attempts to interpret the rules creatively.
4. Limit bookkeeping or calculations performed by participants to the simplest acts. If time charts of data are important to the instruction, only ask participants to count things and record the counts on paper. Let facilitators enter the data into a spreadsheet program to plot and display the data.
5. The simulation must admit at least two parallel strategies within its rules. The obvious strategy, supported by the game story line and the participants' normal habits of "correct" operation, should lead to perdition, or at least low performance. The less obvious strategies that force

participants to abandon ingrained mental models should lead to success. This conflict creates extremely uncomfortable tensions for participants, but is needed to surface mental models in action.

Workshop Insights for Facilitators

1. Facilitators must practice their roles before the workshop. Table facilitators must become experts in game rules and in spotting willful violations or misunderstandings. They must practice common interventions so as to avoid conflict with participants. Skit performers must know their parts and rehearse them beforehand. Game Caller must practice the long speeches describing the rules to provide lively presentations and command the attention of the participants. Responsibility to make these things happen lies with the Game Master.
2. Sometimes the Game Master must intervene to deal with anomalous conditions or to promote the intended workshop learnings. Each outside-the-rules intervention, such as adjusting the threat level for all boards due to unseen events or giving a weak agency more resources as a congressionally mandated fiat, weakens the reality experienced by the participants. We try to limit such measures to restarting game play after a terrorist attack and its associated debrief.
3. The overnight break in the middle of the simulation gives participants time to absorb and process the chaotic events of the first day. Only overnight do participants soften the hold of their own mental models. They come on the second day ready to experiment with strategies outside their normal repertoire. Every time we have compressed the workshop into a single day, few participants arrive at the insights the workshop is designed to achieve. If the workshop must be shortened to eight or fewer hours, it is far better to split the time over two days with an overnight break in between.
4. Always treat participants with utmost respect. Their reality is correct as far as they are concerned. Moreover, some of them are confused about the rules and may feel both frustration and anger in the early stages of the workshop. Many will blame the facilitators or the game itself for any shortcomings.
5. Sometimes individual participants will find the simulation so intimidating, frustrating, or incongruent with their own view of reality that they become inactive or physically leave the workshop altogether. If this condition can be detected early, Game Master can work with those individuals offline to deflect the resulting resentment. Frequently, these problems arise when participants have been ordered to do the workshop by their superiors in the face of mounting backlogs back on the job. Always survey the participants before the workshop to find out why they have come and what they expect from the workshop. "My boss told me to be here" is an indicator of future problems.
6. Each workshop demands logistic arrangements. Venue facilities, workshop materials, lunch orders, last-minute participant cancellations and recruiting, and game board setup all have to work smoothly.
7. Few errors are more damaging to a workshop than errors in the initial game board setup. Check each board by three different people to find omissions or other errors in setup.

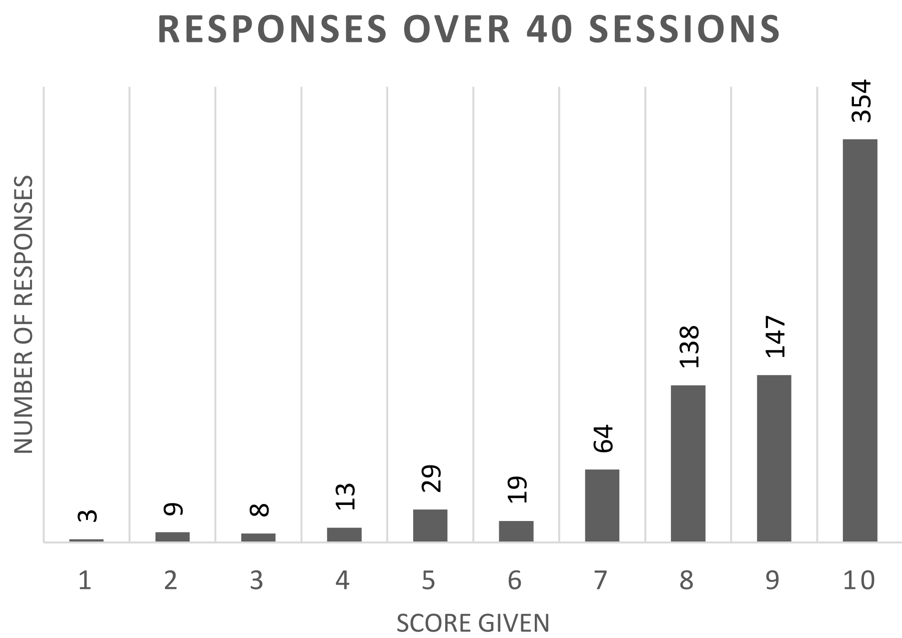
Workshop Performance

Although the workshop was presented only once at the Intelligence Community level, it has been presented more than 40 times at a single agency and has trained more than 750 participants over six years. Unfortunately, we lack a quantitative measure of the impact of the workshop on collaboration and decision-making on the job. Instead, we have measured the ability of the workshop to elicit requests for more workshop sessions in the face of competition from other leadership courses and to propagate to new organizations. As past participants move into leadership positions of increasing responsibility, the influence of the workshop shows up in their adoption of the workshop's vocabulary, their willingness to recommend the workshop to peers and subordinates, and their use of the workshop as a team-building exercise for their own organizations. This increases exposure of the target population to the workshop's lessons and serves as a necessary step on the path to real culture change. Thus, repeated use and spread to new participants has become our proxy measure for the workshop's effectiveness. To this end, we track the responses to the exit survey questions.

At the end of each workshop, participants fill out an evaluation questionnaire to rate the workshop experience. The Net Promoter Score (NPS) has become a widely used standard of performance measurement across a diverse set of industries. (Reichheld and Markey, 2011) The form has two questions: 1) "Using the scale supplied (a 1-10 scale), would you would recommend this workshop to a colleague or friend? 2) Please give a brief explanation for the score you just gave." The scores and comments guide corrections for future workshops and provide both numeric and anecdotal assessments of the perceived effectiveness of the workshop. At the bottom of the form, a third question asks if the participant would like to become a facilitator for the workshop in the future.

Defined as the percentage of promoters minus the percentage of detractors in a pool of respondents, the Net Promoter Score aims to measure the strength of word-of-mouth promotion and to predict repeat sales. A promoter is a respondent who gives a score of 9 or 10. A detractor is a respondent who gives a score less than 7. Respondents giving 7 or 8 are regarded as neutral and not included in the Net Promoter Score. A study of the Net Promoter Scores of 19 industry groups over 175 firms in 2012 found that top performers in airlines, fast food chains, auto dealers, retailers, grocery store chains, and computer makers have Net Promoter Scores in the 45% to 70% range. (Temkin Group, 2012)

Our workshop exit survey results given by 784 respondents in 40 workshop sessions are accumulated as a histogram in Graph 1.



Graph 1 Distribution of Net Promoter Scores

From these numbers, 63.9% of respondents are promoters and will likely speak well of the workshop, while 10.3% are detractors, who will likely speak badly of the experience. Thus, the NPS for this data is 53.6%, and comparable to the high performers in the above industries. We conclude that this workshop enjoys a healthy word-of-mouth promotion effect. Unfortunately, growth of the workshop due to word of mouth promotion is limited by the scarcity of classrooms, as the available classrooms are allocated over an extensive curriculum. We now run 6 to 8 workshops per year.

The rationale players give for the scores they mark does provide insight into initial reactions to the workshop. Here are typical comments offered by participants from the populations of score 10 and scores 2, 4 and 5 in Graph 1:

> (Score 10) "First practical leadership class I've taken here. Puts theory into practice & results speak for themselves. Stresses, and proves, the necessity to balance all aspects of organizations – operations, training, HR, logistics, infrastructure, succession planning, etc."

> (Score 10) "I had difficulty believing that you have to really heavily staff the enablers. That doesn't fit with my usual world-view."

> (Score 10) "Fun, really opened up my eyes to "big picture" ideas.

> (Score 10) "Understanding that every organization impacts all other organizations in the stark way it was brought out in the class is an insight vital to everyone's work. Seeing the impact of a lack of collaboration and support for all (at all levels) in a concrete way puts the importance of holistic awareness front & center. It also emphasized the need to balance strategic and tactical activities in order to sustain a healthy organization."

> (Score 2) "Was looking for something more practical & hands on."

(Score 2) “Too confusing and irrelevant for my REAL job.”

(Score 5) “My position does really not currently utilize this course.”

(Score 4) “Still not sure what I will be taking away and if the time away from my desk was beneficial”.

Detractors seem to have trouble connecting the game world with their own reality. We have not yet identified the most common reasons for this disconnection and recognize the disconnection as an important topic for future research.

Precise data measuring demand for the workshop are not available. However, informal evidence suggests that demand is growing. On average, approximately 15% of workshop participants ask to become future facilitators, and over 100 of them, 30 of whom are currently active, have been trained. The number of internal organizations stating requirements for the workshop has increased over the years, as has the diversity of customer organizations. More customers are asking for special sessions for their intact organizations. Moreover, the workshop has been adopted by one other agency – meaning that agency has developed their own team of facilitators, constructed their own game materials and offers the workshop internally on a repeated basis. This evidence suggests that the reach and impact of the workshop remains in the early stages of adoption, both within the one agency and the larger Intelligence Community.

Conclusion

This paper describes a serious game for management training workshops. It was designed to surface deeply ingrained mental models held by managers in government agencies. It provides a safe, alternate reality for participants in which some of them recognize and confront habits and behaviors that limit their performance in their real jobs. Taking the form of the capability trap of Repenning and Sterman, and Morrison, the game confronts participants with the need to make counterintuitive decisions under pressure. Because it is based on system dynamics analyses and models, it offers the chance for system dynamics instruction just when participants perceive that they need it. It also distinguishes collaboration from other forms of cooperation. Many workshop sessions have shown that this experiential mode of assisted self-instruction is well received by participants, and generates word-of-mouth expansion of the customer base. The adoption of the workshop appears to be now in the early stages of growth. We are optimistic that in time its beneficial effect on the community at large will become apparent.

References

9/11 Commission 2004, *The 9/11 Commission Report*, U.S. Government Printing Office.

Gardner B.,1994, Video obtained through private communication. Game is described on the website www.fridaynightattheer.com. Accessed 12/14/2013.

Koster R. 2005, *A Theory of Fun For Game Design*, Paraglyph Press, Scottsdale, AZ

Ledet W and Paich M 1994, The Manufacturing Game, Text of a presentation at Goal/QPC TQM Conference Boston MA on Nov 16 1994, available online at www.manufacturinggame.com/resources/themanufacturinggame.pdf , Accessed 2/21/2014

Lyneis J, and Sterman JD 2014, *How to Save a Leaky Ship: Capability Traps and the Failure of Win-Win Investments in Sustainability and Social Responsibility,* Draft obtained through private correspondence.

Morrison JB. 2012*, Process Improvement Dynamics Under Constrained Resources: Managing the Work Harder vs. Work Smarter Balance,* System Dynamics Review **28**(4) 329-350.

Reichheld F. and Markey R 2011, *The Ultimate Question 2.0*, Bain and Company.

Repenning NP and Sterman JD 2001, *Nobody Ever Gets Credit For Fixing Problems That Never Happened: Creating and Sustaining Process Improvement,* California Management Review 43(4)*,* 64-88.

Repenning, N. and J. Sterman (2002). *Capability Traps and Self-Confirming Attribution Errors in the Dynamics of Process Improvement.* Administrative Science Quarterly **47**(2): 265-295.

Senge PM. 2006, *The Fifth Discipline*, Currency/Doubleday/Random House, New York.

Senge et al. 1994, *The Fifth Discipline Fieldbook*, Doubleday, New York

Sterman JD. 2000, *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw Hill, Boston

Sterman, JD 2014, *Stumbling Toward Sustainability: Why Organizational Learning and Radical Innovation Are Necessary to Build a Sustainable World – But Not Sufficient*, "Organizational and Strategic Change and the Challenge of Sustainability" Henderson R, Tushman, M, Gulati R.(eds), Oxford University Press (Forthcoming).

Temkin Group, 2012, Net Promoter Score Benchmark Study 2012, www.temkingroup.com/research-reports/net-promoter-score-benchmark-study-2012. Accessed 2/21/2014.