What is the Use of Basic Dynamic Tasks?

Eva Jensen[1]
Department of War Studies
Swedish National Defence College
P.O. Box 27805
SE-115 93 Stockholm, Sweden
Telephone: +46 8 55342610
Fax: +46 8 55342600
E-mail: eva.jensen@fhs.se

Agata Sawicka
Faculty of Science and Engineering
Agder University College
NO-4876 Grimstad, Norway
Telephone: +47 37 25 34 25
Fax: +47 37 25 30 01
E-mail: agata.sawicka@hia.no

The various purposes for which a dynamic tasks might be constructed, such as to test for knowledge, teach, or to assist professionals or the lay public in understanding the systems they are dealing with (or part of), are discussed. The idea analysis method is suggested as a means to fit a task to its purpose. Idea analysis entails analysing the task in terms of what basic ideas need to be familiar if one is to be able solve the task. It is just as important to know what knowledge a task does not require as to know what it does require, and if the requirements corresponds to the goal(s) motivating the construction of the task. To provide an example, the Computer Security Incident Response Team (CSIRT) task, a close analogue to the one-stock reindeer management task by Moxnes, is analysed, and several issues of general importance are revealed.

KEY WORDS: BASIC DYNAMIC CONCEPTS, TASK ANALYSIS, TASK REQUIREMENTS, KNOWLEDGE TESTING, TEACHING, LAY PUBLIC COMMUNICATION

Introduction

A vast number of studies demonstrating people's misperceptions of basic dynamics, have accumulated over the years (see e.g., B. T. Bakken & Vamraak, 2003; Booth Sweeney & Sterman, 2000; Brehmer, 1995; Jensen & Brehmer, 2003; Kainz & Ossimitz, 2002; Moxnes, 2000; Moxnes & Saysel, 2004; Sterman, 1989a, , 1989b), and efforts have been made with systems archetypes (Senge, 1990; Wolstenholme, 2003) and "flight simulators" (B. E. Bakken, Gould, & Kim, 1994; Sterman, 1992) to help people overcome these misperceptions and gain some understanding of the basics of dynamic systems. This paper discusses the various purposes for which "simple" dynamic tasks are constructed, and suggest what analyses need to be made to assess how well a task fits its purpose, or better, what to consider prior to constructing a dynamic task, for any purpose.

[1] The authors are indepted to Dr. Berndt Brehmer for his comments on earlier drafts of this paper.

Testing for knowledge and understanding

Tasks used to test for knowledge or understanding need to be analyzed in terms of what one has to know to be able to

a) solve the task, or give a correct answer

b) understand the system

These are not necessarily identical requirements, which will be demonstrated later in this paper. A person who understands the system is likely to be able to solve the task and correctly answer questions, while a person who solves a given task does not necessarily have a good understanding of the system. A task may be correctly performed on the basis of a flawed model of the system, a model that just happens to be consistent with the presently required actions and replies (Moray, 1987, 1999). For tasks that are used as knowledge tests, it is crucial to know what knowledge is demonstrated by good performance, and what knowledge is not.

Booth Sweeney and Sterman (2000) commenced the commendable effort of developing an inventory of tasks testing for systems thinking skills, in particular basic systems concepts, many of which are taught as part of most high school curricula. One result of this effort was the, by now famous, bathtub tasks. Booth Sweeney and Sterman (2000) dissected the tasks into features required for a correct solution. These features were then used as performance criteria. We suggest taking one step further and investigate the ideas entailed in these features, in accordance with the idea analysis method proposed by Lakoff and Nuñes (2000).

Teaching and learning dynamics

When people learn about mathematical concepts, they start with a few innate ideas. When they extend their mathematical knowledge by adopting new ideas, the endorsement of certain "old" ideas are a prerequisite for learning these new ideas. What people learn are not facts about the world, but ways of thinking about the world that help them predict the behaviour of the reality they encounter. One could say that we learn to construct abstract worlds to use as thinking tools. Within the field of mathematics a range of such abstract, sometimes conflicting, worlds are used (Lakoff & Nuñes, 2000). When learning about the number 0, for example, a child has to mentally construct and accept a world in which nothing is a concrete something, namely the number 0.

System dynamics is one way of constructing worlds that are abstract simplified representations of phenomena in the "real world". Constructing and using system dynamics models requires the endorsement of a range and hierarchy of fundamental ideas (basic systems concepts). It is important to identify these basic ideas, and how they build upon each other, to identify idea hierarchies.

If we could analyze dynamic tasks in terms of basic ideas, it would be possible to infer what ideas need to be mastered to solve what tasks. It would also be possible to devise test items that test for the specific ideas. This would inform educators about what ideas need to be taught and what ideas their students do already master.

It is important that when we try to extract the basic ideas needed to understand dynamic systems, we avoid a one-eyed focus on the mastering of system dynamics concepts. System dynamics is but *one* way to represent dynamic systems (Richardson, 1991). People may be familiar with an idea but unfamiliar with how it is represented, and therefore fail to understand. It is important for educators to be able to assess prior knowledge among their pupils to know what there is to build upon. If the idea is already mastered and only the representation new, the pedagogical challenge is to facilitate analogical transfer (Chen, 2002; Gentner & Markman, 1997; Holyoak & Thagard, 1997) from one representational practice to another, rather than the formation of a new concept (Chi, Slotta, & de Leeuw, 1994; diSessa & Sherin, 1998; Slotta, Chi, & Joram, 1995).

Assisting professionals in understanding the systems they are dealing with (or are part of)

The task, to teach the public or professionals in other fields about the behavior of dynamic systems they are part of, influencing, or influenced by, is a complex one. Here the idea analysis becomes even more important, because we have to find shortcuts to understanding. We do not intend to make these people into full-fledged system dynamicists, understanding the systems in all their minute details. What we want, and they probably do as well, is that they acquire some understanding at a more conceptual level. The question, then, is how conceptual knowledge is built in an efficient way that allow for the omission of technical details. Knowledge that is too superficial is seldom particularly useful. You need to reach a certain, unspecified, level of understanding to be able to make use of your acquired knowledge.

This raises another important question affecting our interaction with people who are not system dynamicists: Do we want them to learn about and understand dynamic systems or do we want to convince them that system dynamics is an important issue they need to seriously consider? The answer is probably that we want to achieve a bit of both, but maybe different means are best suited for the different purposes.

Either way, it is important to know the people we are addressing. What knowledge can we expect them to have that we might build upon? What is it, exactly, and in detail, that we want them to understand? And, are there any ways by which we can ascertain that the message gets through as we intend it to?

THE PRESENT STUDY

In the present paper, one particular task, the Computer Security Incident Response Team (CSIRT) task (Sawicka, Gonzales, & Qian, 2005) is analysed in terms of what kind of basic system concepts it is based on, and the ideas inherent in these concepts. The CSIRT task is a close analogue to the one-stock reindeer management task used by Moxnes (2004). The CSIRT system and the one-stock reindeer-lichen system by Moxnes (2004) are isomorphs, and the task demands are identical. It is only the context that differs, as in the case of the bathtub and cash flow tasks by Booth-Sweeney and Sterman (2000).

The CSIRT task is interesting because it is used as a knowledge test, but suggestions are made that the task might be useful in the classroom to illustrate challenges faces by CSIRT managers, and in this way contribute to the computer security field (Sawicka, Gonzales, & Qian, 2005). This means it is intended to tap into all the abovementioned fields of use.

The CSIRT task is also interesting because of the attempt to take one existing system and task structure and apply it to another domain. This is akin to how systems archetypes are supposed to work. How well does it work? What difficulties are run into and what are the benefits of this approach?

So, the major benefit of choosing the CSIRT task as the object of study is that it can be used to address several important issues, issues that are vital to any dynamic task constructed.

In the following, the CSIRT task is described and the correct solution, together with a simpler one, is explicated in detail. The CSIRT task is analyzed in terms of what basic concepts are required for producing a correct solution, and what concepts are required for a full understanding of the system but not necessary for producing the correct solution. This analysis reveals a range of concepts of general importance to the understanding of dynamic systems. Finally, the possible uses of the CSIRT task, as knowledge test, teaching instrument or as a means to communicate with CSIRT managers, are discussed in the light of the preceding analysis. The issues raised are of general relevance and not limited to the CSIRT task.

The CSIRT Task

In the reindeer management task (Moxnes, 2004), the player can, once a year, decide on the size of a reindeer herd. Lichen is the main source of winter fodder for the reindeer. Lichen is a perennial, which means that the yearly growth of lichen depends on the amount of lichen left when spring arrives. The task is constructed as a simulator game and the goal given to the participants is to keep as many reindeer as possible without exhausting the lichen resource, in other words to achieve and maintain the largest possible long-term sustainable reindeer herd, and to do this as fast as possible.

A Computer Security Incident Response Team (CSIRT) is a service organization that assists business or governmental organizations with handling computer security incidents, both by dealing with upcoming threats and by taking proactive measures. In the CSIRT task, providing services to customers corresponds to the grazing of lichen by reindeer in the reindeer management task. What is consumed is the work capacity available in the CSIRT. Work capacity grows by revitalizing and competence building activities, corresponding to lichen growth in the reindeer management task. The task is then to provide as many services as possible while maintaining work capacity in the CSIRT, and to achieve this situation as fast as possible (Sawicka, Gonzales, & Qian, 2005).

The instructions to the CSIRT can be found in the Appendix to this paper, and screenshots of the interface are presented in Figure 1.

Information on CSIRTs

Participants performing the CSIRT task receive an initial description of CSIRTs, in addition to the task instructions parallel to the reindeer management task instructions (see Appendix pp. A1-A3). This text describes what a Computer Security Incident Response Team is, how capacity is measured in average person-hours, how capacity at the individual level varies with experience, how capacity is reduced and how capacity is increased.

This adds to the demands put on the participants, compared to the requirements of the reindeer management task, in two ways:

First, there is more material for the participants to study. Reading tires the participants, and probably increases the frequency of skimming.

Second, understanding the concept of an average person-hour probably requires some effort, and the processes involved in the growth and consumption of CSIRT capacity are not as easily visualized as growing lichen and munching reindeer.

This ought to make the CSIRT task a somewhat more difficult task than the reindeer management task.

Instructions

The task instructions to the CSIRT task (see Appendix pp. A4-A5) are an as close as possible parallel to the instructions to the one-stock reindeer management task (Moxnes, 2004, Appendix 1). The mayor points are:

- The *goal*: Aim for the highest possible available service level, and reach it as soon as possible.

- You will make your decisions quarterly for the next four years.

- There are no competitors.

- The growth in capacity is an inverted U-shaped function of the present capacity. (This information is implicit in the text. It is not explicitly stated; see Appendix)

- In one quarter, providing 100 services reduces the capacity level by 4 average person-hours.

- There are no restrictions on how many services to provide. It is allowed to not provide any services at all. (There must, of course, be an upper limit corresponding to the capacity available.)

- The historical data given is perfect. There is no noise, no distortions, and no random variations.

Last in the instructions is the following section: “Before you take over your CSIRT, you need to know that the previous manager has increased steadily the number of services provided from 115 to 185 over the past 4 years. As a consequence, the CSIRT capacity [average person-hours] has dropped from the initial 50 average person-hours to 24.4 average person-hours at present. This development is shown in the diagrams and table below.” The diagrams and the table are the same as those shown on the second screen shot of the interface in Figure 1 (see also Appendix, p. 5).

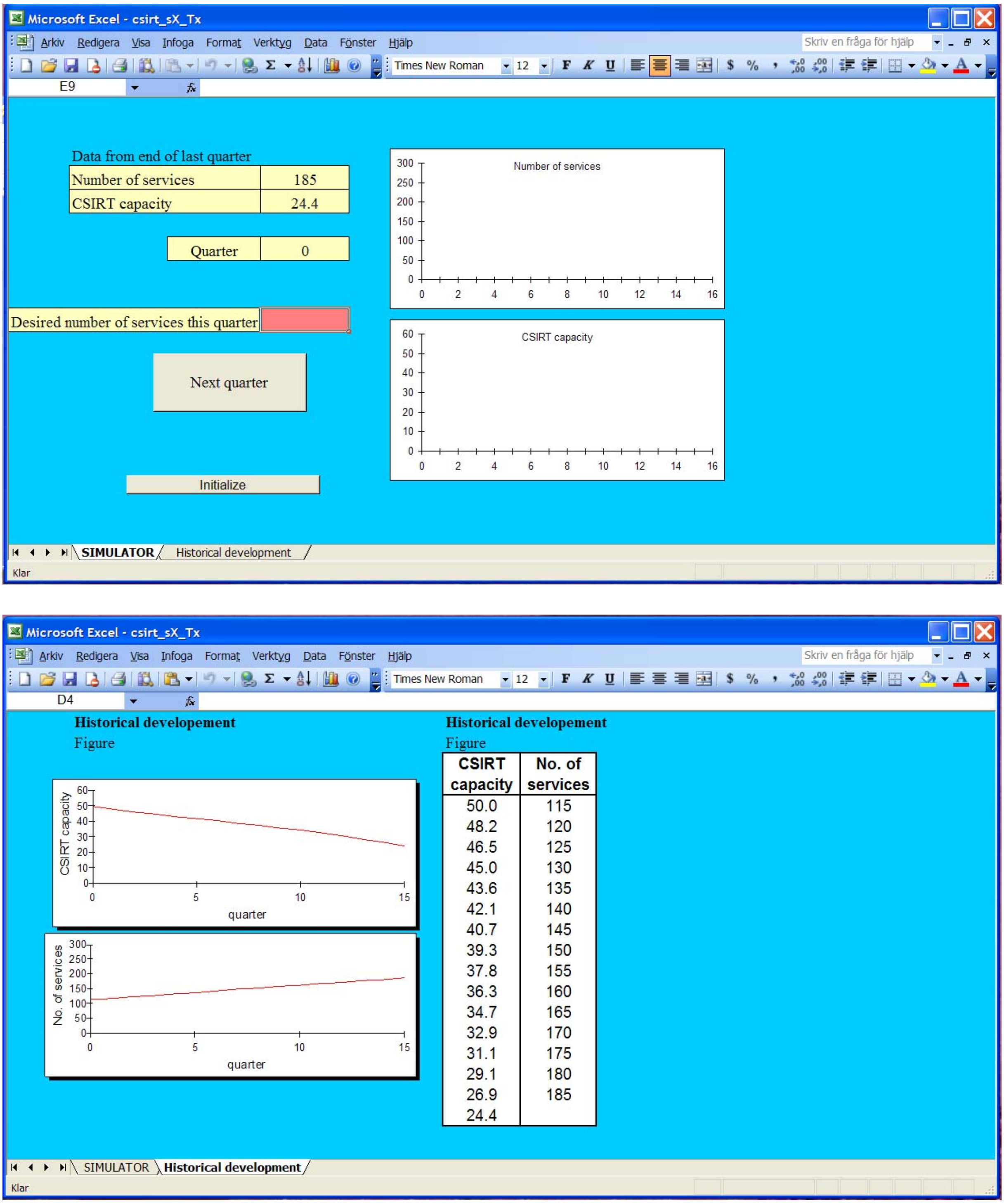


Figure 1. The simulator interface for managing CSIRT services. The interface is analogous to the interface used by Moxnes (2004).

Analysis of the Solution(s)

A CORRECT (OPTIMAL) SOLUTION

To solve this task optimally, or at all, one has to understand:

1. that to maintain a certain capacity level, the number of services provided must correspond to the capacity growth at the present capacity level.

2. that to achieve the highest sustainable provision of services, the capacity level must be such that it produces maximum capacity growth. Then the number services that produce a capacity depletion that equals the maximum capacity growth, can be provided. This would be the highest possible sustainable level of provided services. If this other consideration is not taken into account one could settle at a capacity of e.g. 50 providing 70 services. (Missing this other consideration easily leads to a Simplified Solution, see below.)

3. how to use the provided historical information to calculate the maximum capacity growth. This entails constructing the growth curve of the work capacity.

Construction and interpretation of the capacity growth curve

In the following, the solution suggested by Moxnes (2004) and Sawicka, Gonzalez and Qian (2005) is analyzed. It represents the most straightforward way to produce a growth curve for the CSIRT capacity.

To construct the capacity growth curve, one needs to understand that the resulting capacity after one time period [C_{t+1}] is the capacity at the beginning of this time period, that is the capacity resulting from the previous time period [C_t], plus the growth in capacity $G_{[t, t+1]}$[2] from capacity-increasing activities during the time interval, minus the capacity consumed by the provision of services $R_{[t, t+1]}$ during the interval (see Moxnes, 2004).

(1) $$C_{t+1} = C_t + G_{[t, t+1]} - R_{[t, t+1]}$$

Note that $G_{[t, t+1]}$ and $R_{[t, t+1]}$ are the accumulated growth and reduction of capacity during the time interval. These are stocks, and not the growth and reduction rates.[3]

By rearranging the terms in equation (1) into

(2) $$G_{[t, t+1]} = C_{t+1} - C_t + R_{[t, t+1]}, \text{ where } R_{[t, t+1]} = \text{Services provided} \cdot 4 / 100,$$

the capacity growth for previous 15 quarters can be calculated. The capacity growth in the first time quarter:

(3) $$G_{[0, 1]} = 48.2 - 50.0 + 115 \cdot 0.04 = 2.8$$

The results for all 15 quarters are presented in Table 1.

Table 1. The capacity growth in the 15 previous quarters. (The information provided with the instructions is marked in grey.)

CSIRT capacity	No. of services	Capacity growth

[2] $G_{[t, t+1]} = \int_t^{t+1}(g_t)dt$.

[3] Moxnes (2004) makes the lapse of mixing stock and flow variables. He writes the equation $C_{t+1} = C_t + g_t - r_t$ (although dealing with lichen and with another expression for r_t). This cannot possibly be. A correct expression would be $C_{t+1} = C_t + \int_t^{t+1}(g_t - r_t)dt$.

50.0	115	2.8
48.2	120	3.1
46.5	125	3.5
45.0	130	3.8
43.6	135	3.9
42.1	140	4.2
40.7	145	4.4
39.3	150	4.5
37.8	155	4.7
36.3	160	4.8
34.7	165	4.8
32.9	170	5.0
31.1	175	5.0
29.1	180	5.0
26.9	185	4.9
24.4		

Plotting the growth in capacity in the next quarter for each of the capacity levels of the recent 15 quarters produces Figure 2.

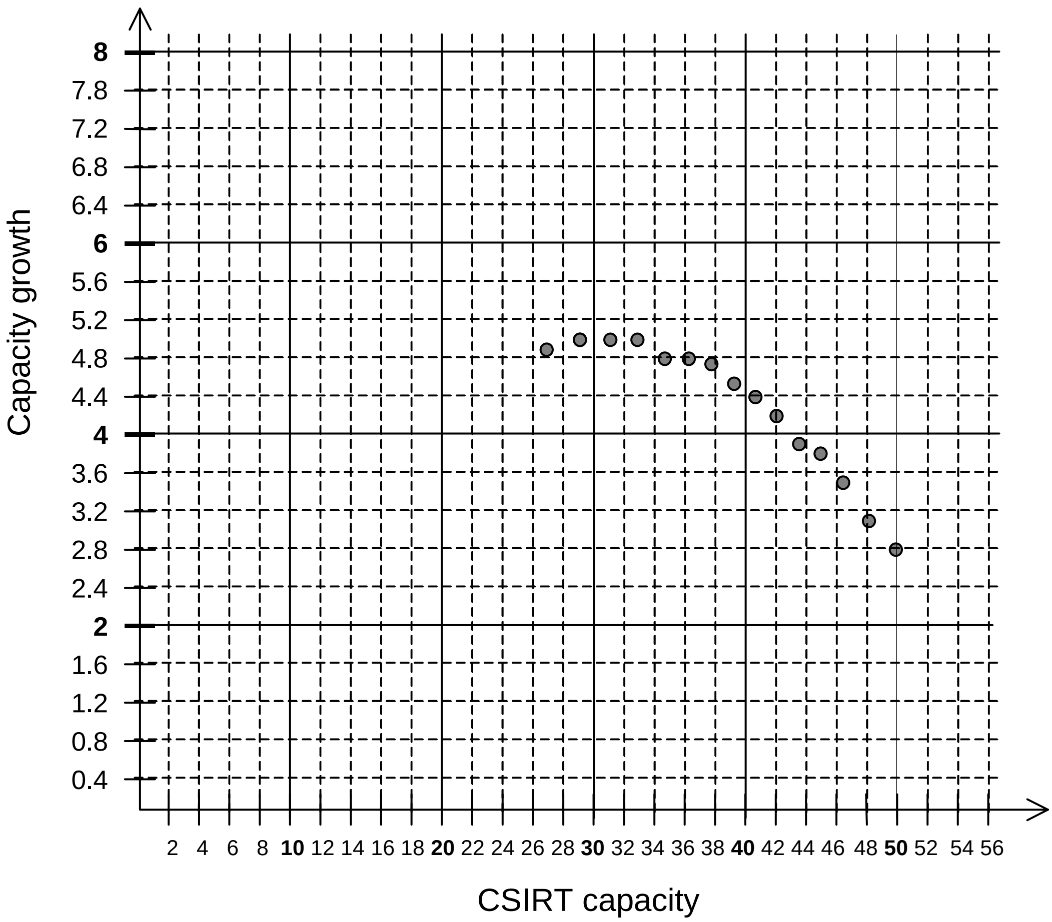


Figure 2. Growth of capacity, in average person-hours, in the following quarter as a function of present capacity.

The maximum growth of capacity, 5.0 average person-hours in a quarter, occurs at a capacity level of approximately 30 average person-hours. This corresponds to 5.0 · 100 / 4 = 125 services.

How to reach this goal?

The closest to optimal way to find this out would be by performing the following calculations (see Sawicka, Gonzales, & Qian, 2005):

The present level of capacity is 24.4 average person-hours. The CSIRT capacity growth curve (Figure 2) does not indicate directly what the equilibrium level of services at this capacity level is. If one assumes, however, that the growth curve is symmetric and that its maximum is at the capacity level of 30 average person-hours, then one knows that the capacity needs to grow by 5.6 average person-hours to reach its maximum level:
30 – 24.4 = 5.6

This is more than it is possible to gain in one quarter, so by not providing any services at all the first quarter at the capacity level of 24.4 average person-hours (corresponding to the 35.6 average person-hour level of capacity due to the symmetry), a growth in capacity of 4.8 average person-hours is achieved. Now only 0.8 average person-hours remain to reach the desired capacity level.

The capacity growth in the next quarter at a capacity level of 29.2 average person-hours is 5 average person-hours according to the table (or the graph). If the reduction of capacity due to provision of services is kept at 4.2 average person-hours in the second quarter the goal of a capacity level of 30 average person-hours should be reached. This means providing 4.2 · 25 = 105 services the second quarter. So, the decisions for the second quarter should be to provide 105 services, and for the remaining quarters to provide 5 · 25 = 125 services/quarter.

A SIMPLIFIED SOLUTION

Performing a correct solution, finding the maximum sustainable level of services, is not that straightforward, as has been demonstrated above. Someone, who is not able to figure out how to obtain the actual optimal number of services, has to settle for a goal to provide as many services as possible without reducing net capacity.

One can see that all the previous four years the capacity level has been constantly decreasing, while, at the same time, the number of services provided has been constantly increasing. Hence, the number of services provided has all the time been larger than what it should be if one wanted to sustain the present capacity level. This suggests the following course of action:

1. Reduce the number of services until the capacity starts to increase when the growth of capacity exceeds the consumption. There is no way to know how large reductions to make. A certain amount of trial-and-error is necessary.

2. Provide few services until the capacity approaches a “good” level. It is up to the task solver to decide on what is a desirable capacity level to maintain

3. The next step is to find the number of services to provide that will keep the capacity at its present (desired) level. This also entails some trial-and-error. This would result in a sustainable solution, but how close or far from the optimal there is no way to know.

4. With repeated trials it is possible to systematically test higher "good" levels, and get closer to the optimal. In case a participant grasps that there exist *one* optimal capacity level, he or she could strive to encircle it.

Aspects of Dynamics Systems Relevant to the CSIRT Task and to Dynamic Tasks in General

We now turn to the issue of what understanding and knowledge is required, both to produce a correct solution to the task, and to have a thorough understanding of the system. The ability to produce a correct solution is not necessarily perfectly related to the latter, as will soon be demonstrated.

We begin by looking at what is needed to understand the system structure, what models are and how they can be used. This is of some importance to the CSIRT task (and the reindeer management task), and of great importance to the use of models of dynamic systems in general.

We then go on to discuss the understanding of the effect of flows on stocks, the accumulation process. A fairly superficial grasp would suffice to be able to produce the growth curve required to perform the correct solution. The requirements of a full system understanding are treated as well, since they are generally important for a correct and thorough understanding of dynamic systems.

There is also the question of the effect of stocks on flows, as in the reproduction of people, plants or animals, and factors putting a limit to boundless growth. These aspects are vital to an understanding of the inverted-U-shaped growth curve for CSIRT capacity as well as lichen, and of the existence of only *one* optimal situation.

Finding the functional relationship between the growth of work capacity and present work capacity, the growth curve, as required for producing the correct solution to the task, is a demand that is more specific to this task and perhaps not as closely related as the other aspects to the understanding of dynamic tasks in general.

The issue of time steps, on the other hand, in truly important to a correct treatment of dynamic systems, while not necessary for producing the correct solution to the CSIRT task (or the reindeer management task). For a full and deep understanding of these systems, however, such knowledge is essential.

The aspects and their relevance to the correct solution to the CSIRT task are summarized in Table 2, by the end of this section.

SYSTEM STRUCTURE

Model thinking

Thinking about a dynamic system entails identifying its structure, its constituent parts and the relations among these parts. Imperative to the use of models is "the consciousness that we deal with models of our reality and not with the reality itself" (Ossimitz, 2000b, p. 10). The ability to build models comes with fully developed model thinking (Ossimitz, 2000a, pp. 58-60, 2000b).

It is possible to think of at least three different levels of model thinking:

1. understanding the model concept
2. model visualization
3. model construction

The third level, model construction, is the ability to extract the central variables important to a question at hand and construct a simplified model of a certain part of reality. This includes the ability to distinguish between important and unimportant factors, and finding an adequate level of detail for the present purpose. It is part of all kinds of modeling efforts. This level is, however, not relevant to the CSIRT task.

The second level, model visualization, is the ability to visualize a model described to you. In the CSIRT task, the system is described in writing, and it is up to the reader to mentally construct a representation (mental model) of the system's structure. The written instructions inform that the capacity growth depends on present capacity, and that providing services reduces capacity. This information has to be sifted out and put together to visualize the basic structure of the system.

The first level, model understanding, is the insight into what a model is. One has to understand and accept the *idea* that a model is a simplified abstract representation. It is not like a cutout, made with a cookie cutter, of the real world. It is a sort of world in itself, a world that only, and strictly, obeys the rules describing it.

Thinking with models is a kind of make-believe play. What happens in the model world may or may not resemble what would happen in the real world. That depends on how closely the model world resembles the real world in the important aspects one wishes to study.

This first-level understanding of the model concept is, as noted by Ossimitz (2000a, pp. 58-60, 2000b), a prerequisite for model thinking of any kind. That happenings in a model world are only influenced by what exists in the model seems to easily escape people unfamiliar with using models. In a study by Jensen and Brehmer (2003), undergraduate psychology students were presented with the rabbits-and-foxes task, where the system is described by the following two sentences: "Every rabbit produces two offspring a year, and every fox eats 4 % of the rabbits a year. For every 180 rabbits consumed a new fox is born, and 20 % of the fox population dies each year." When informed that the foxes were fat (and they could see the rabbits were abundant) and asked about other consequences for the foxes due to the abundance of food (a higher birth rate of fox puppies was the expected answer), various "non-model" answers were received: that the foxes would die from obesity, the foxes would become slower hunters because they were fat and stuffed, or the foxes would get lazy and not eat their share of rabbits (the system description interpreted as a recommendation). Starving foxes received a corresponding set of non-model explanations and expectations (Jensen &

Brehmer, 2006). It seemed as if the participants failed to understand that the system description is the complete and exhaustive description of the simulated world. They did not understand the rules of "the modeling game". If we want people to make use of and reason with models, we have better ensure that they understand what a model is. This understanding cannot, apparently, be taken for granted.

Identify and classify central system variables

Whether constructing or visualizing a model, one has to be able to identify the important parts of the system. Depending on the level of modeling, it might be, and frequently is, important to correctly categorize the variables in the model. Within system dynamics variables are categorized into stock or flow variables (or possibly levels and rates, state and transition variables, or integrals and derivatives). The *idea* is to distinguish between changes taking place in the system (flows) and quantifiable entities (stocks).

Identify the relations among the system variables

When constructing or visualizing a model, one also has to be able to identify the relations connecting the system variables. One way to study relations among system variables is in terms of causal relations, as in done with, for example, causal loop diagrams.

The basic concept of cause and effect is present even in rodents (Schwartz & Robbins, 1995), so it seems reasonable to assume its presence in most of the human population as well. Simple cause-effect relations can be expressed in everyday language, while keeping track of more complex causal structures requires some means of representations, such as causal loop diagrams, and is not as easily handled even by grown-up humans (see e.g., Dörner, 1996; Dörner, Kreuzig, Reither, & Stäudel, 1983).

The causal structure of the CSIRT system is fairly simple, however, leaving modeling tools unnecessary.

Express the nature of the connections.

If one wants a more precise model of a situation, it requires knowledge of the functional relations between the variables. These are generally described by mathematical expressions. The understanding of the system therefore ought to depend on the grasp of the mathematical concepts describing the relations.

In addition to the functional form of the relations, one needs to know the values of the parameters (constants and initial values) describing the system. There is the question of finding them and the question of understanding them.

The CSIRT task is a bit special in this respect, because the functional relation between the available capacity and capacity growth is not described by a mathematical expression, but by a growth curve constructed from the historical information. (This issue is further treated below in the FUCTIONAL RELATIONSHIP section.)

EFFECTS OF FLOWS ON STOCKS

To understand a dynamic system, a system of stocks and flows, one has to understand how flows affect stocks, that things added to a stock make it larger, and that things removed from a stock make it smaller. One also has to understand that when more is added than removed, the stock will become larger, when less is added than removed the stock will become smaller, and when additions are equally large as removals the stock will remain the same (a dynamic equilibrium is maintained).

The average preschooler understands these basic principles (Bryant & Nuñes, 2002; Flavell, Miller, & Miller, 1993). School complicates matters by introducing numbers and strange signs (such as +, -, and =), but the fundamental principles are the same. The problem with these principles is to recognize them when they appear in an unfamiliar context.

In the task where people are entering or leaving a hotel or parking lot (Kainz & Ossimitz, 2002; Ossimitz, 2002), or a department store (Sterman, 2002), more people are entering than leaving up to a certain point in the depicted time period, and for the remaining time more people are leaving than entering. People can easily tell when the most people were entering, or leaving, but find it more difficult to figure out when there were the most people (or cars) in the hotel, parking lot or department store (Kainz & Ossimitz, 2002; Ossimitz, 2002; Sterman, 2002). Most people understand that as long as more people are entering than leaving a store, the store gets more crowded, and that as soon as more people start to leave than enter, the store will get less crowded. It appears, however, that the representation of this information failed to speak to the participants failing at this task. A tabular representation proved somewhat easier than line graphs (Kainz & Ossimitz, 2002), a result replicated by Jensen (2005a).

In the CSIRT task the accumulation process is described by the following equation

$$\text{Stock}(t+\Delta t) = \text{Stock}(t) + \text{Additions}(t, t+\Delta t) - \text{Removals}(t, t+\Delta t) \quad (4)$$

Stock is the level of capacity, Additions(t, t+Δt) are growth of capacity during one time interval, and Removals(t, t+Δt) are the reduction in capacity due to provision of services in the time interval. In the CSIRT task the time intervals $\Delta t = 1$ quarter.

It is not required by the CSIRT to be able to produce equation (4), to know how to use that kind of representation. It suffices to understand the principle conveyed.

This is, however, not an equation of stocks and flows, but an equation of one stock with smaller stocks added and removed from it. This is not the typical way of expressing the accumulation process.

Calculating the effect of flows on the stock they are flowing into or out from, is generally done by integrating the net flow. The analytical formula for obtaining the exact result is:

$$S(t) = S(t_0) + \int_{t_0}^{t} [\text{Inflow}(s) - \text{Outflow}(s)] \cdot ds \quad (5)$$

$S(t_0)$ is the stock at time t_0, S(t) is the stock at time t, and the integral $\int(\text{Inflow} - \text{Outflow})\cdot ds$ is the accumulated stock in the time interval from t_0 to t.

System dynamics modeling uses the principle of graphical integration (see e.g., Sterman, 2000, pp. 232-240). The time is divided into intervals of suitable length (see TIME STEPS below). The stock accumulated in a time interval of the length Δt, starting at time t and ending at time t+Δt is:

(6) $$S(t+\Delta t) = S(t) = [\text{Inflow}(t) - \text{Outflow}(t)]\cdot\Delta t$$

This is an approximation, where the net flow at the beginning of the interval is used as *an estimate* of the net flow during the entire interval.

Note that in system dynamics dt is normally used in the way we use Δt here. We use Δt to indicate an interval of measurable size, and ds (or dt) to indicate an infinitesimal interval (which is really no interval at all).

Now, if we want to know the stock at time t, have chosen a time interval of the length Δt, and the duration in time from t_0 to t is n multiples of Δt ($t - t_0 = n \cdot \Delta t$), then

(7) $$S(t) = S(t_0) + \sum_{k=0}^{k=n-1} [\text{Inflow}(t_0+k\cdot\Delta t) - \text{Outflow}(t_0+k\cdot\Delta t)]\cdot\Delta t$$

S(t), the stock at time t, is the sum of the stock at time t_0, $S(t_0)$, and the accumulations in all the intermediary intervals. The exact analytical solution in equation (5) can be thought of as a special case of equation (7). If the time intervals Δt are made incredibly short, so short that they are close to nothing, although yet existing, then Δt = ds. For such short intervals, the beginning and the end will be so close to each other that the flows are virtually the same during the entire, close to non-existing, interval. In fact, they are exactly the same. The sum of the accumulations in these infinitely many infinitely short intervals is the result of the analytical integration.

To understand how flows affect stocks, fundamental to the understanding of dynamic systems, one has to understand the principle of graphical integration, and how it differs from the exact analytical solution. It is not necessary to be able to produce the analytical solution. It is sufficient to be aware of the existence of an exact solution, of that graphical integration produces an approximation, and in what way this approximation deviates from the exact (true) result.

EFFECTS OF STOCKS ON FLOWS

Self-generated inflow

Renewable resources, such as perennial plants and livestock, create their own inflow. They grow or reproduce themselves. The amount of growth or reproduction depends on the present size of the stock, and with no limiting factors present it tends to be exponential.

Although people estimating future growth from experience with the growing tend to underestimate the exponent (Wagenaar & Sagaria, 1975; Wagenaar & Timmers, 1979), they are able to understand the principle. In a group of first-semester engineering students asked when a pond of daily doubling pond lilies would be filled to a quarter if it would fill up entirely in 30 days, 85 % gave the correct answer (in 28 days, or two days earlier) (Jensen, 2005a).

According to the instructions to the CSIRT, the CSIRT capacity growth will be slow due to insufficient available support (Appendix, p. 4). Support is to be interpreted in terms of available time and energy in the CSIRT, a clarifying piece of information that ought to be included with the instructions.

(Stock-dependent outflow)

There might be cases when the outflow is dependent on the size of the stock, but that situation does not present itself in the CSIRT task.

LIMITING FACTORS

Limiting factors, such as available space or food, are stocks that might be emptied. In stock-flow diagrams these can be stocks that are part of the modeled system or exogenous variables affecting the modeled system (Sterman, 2000, p. 202).

Non-linear growth

Population growth tends to be exponential, the more "individuals" who reproduce the more reproduction, until a limit is reached, it starts to get crowded, for example. Then growth might show an inverted-U-shaped relation to size of stock (population), as is the case with lichen in the reindeer management task.

The CSIRT task instruction informs that the capacity growth is slow at low and high levels of capacity with maximal growth in between these extremes. The inverted U-shape of the relation between capacity growth and present capacity is implicit in this information, as is the case in the reindeer management task.

The CSIRT task instructions inform that at high capacity levels updates and upgrades are the only developmental activities occurring (Appendix, p. 4). At the highest capacity level, the entire staff consists of experts. At this stage they are primarily upgrading and updating their knowledge and tools, and not acquiring much new knowledge. Adding this information would also have made the instructions clearer.

The CSIRT task requires, as does the reindeer management task, understanding that when there is a limit to growth, the relation between the growth and the existing amount of the entity growing will be such that there is a point, a specific amount, at which the growth will be as large as it will ever be.

To understand this is, in all likelihood, a precondition for taking the trouble of constructing a growth curve. This understanding, without the idea of constructing the growth curve, would still permit a simplified solution at a higher level if repeated trials are allowed. (See point 4 in the SIMPLIFIED SOLUTION section under the Analysis of the Solution(s) heading above.)

FUNCTIONAL RELATIONSHIP

There is no mathematical description of the functional relationship between growth of capacity that can be used to provide services and available capacity.

Constructing the growth curve

To construct the growth curve, one has to understand how to use the accumulation equation (6) to calculate to capacity growth in the previous 15 quarters from the given historical information. This would give the information plotted in the growth curve in Figure 2. It is not necessary to construct the graph; the table provides the necessary information. A graph is, however, generally a good visualization aid. If the graph is plotted, we now have a graphical description of the relation between capacity growth and present capacity. (Otherwise the shape of the graph can be inferred from the tabular information.)

It seems unlikely that people would get the idea of constructing the growth curve without prior practice in graphing growth, or at least having witnessed it being demonstrated. To spontaneously arrive at the idea, without having experienced a similar solution, must be a rare event indeed.

Sawicka, Gonzales, and Qian (2005) collected all side notes made by their participants when performing the CSIRT task, but very few calculations or notes were produced. This could be interpreted as evidence that no one succeeded, probably not even tried, to produce the growth curve. Moxnes (1998) found a mildly beneficial effect of presenting participants with a ready-made growth curve as an aid on performance in the reindeer management task.

When presented with a growth curve, there are two possible ways to make use of it, reflecting two separate levels of understanding:

Identifying maximum and minimum values of a variable

Jensen (2005b) investigated how participants performing the reindeer management task (Moxnes, 1998) made use of a ready-made growth curve. Of the 28 psychology students presented with the reindeer management task, and equipped with the growth curve, 21 pointed out the maximum as the desired goal.

This requires the ability to identify a maximum of a graphed function, which is probably related to identifying minima, and to do so with tabular information as well. Basic understanding of most and least are acquired pretty early in life (Bryant & Nuñes, 2002; Flavell, Miller, & Miller, 1993).

With this insight gained from the growth curve, you know what to strive for, but not necessarily how to get there.

Understanding the growth curve as a range of possible equilibria

The growth curve represents a range of possible equilibria, where the capacity growth is an expression of the number of services it is possible to provide while maintaining the present level of capacity. This has to be understood to be able to make full use of the growth curve. This would allow one to:

a) calculate, in a simple way, the optimal adjustments to make to reach the goal as soon as possible. (See "How to reach this goal" in the Analysis of the Solution(s) section above.)

b) understand that there was an excess amount of capacity available in the initial three years. Then it was a correct decision to increase the number of provided services. By

> the 13th quarter, however, the provision of services ought to have been reduced. This means that it is only the last year that shows an unlucky development.

Three of the 24 participants in Jensen's (2005b) study offered good and clear explanations to what was presented in the growth curve and how it could be used. Several participants were clearly confused that the curve went down in the rightmost part of the curve. Even when they concluded that the maximum represented the optimal situation, they were frequently at loss as to how to get there. Several tried to set the size of the reindeer herd to the number that could be sustained under optimal conditions, and were surprised these optimal conditions were not brought about. Moxnes (1998) also reported such behavior among his participants. Another strategy tried by some participants was to set the reindeer herd to the size that could be sustained on the present amount of lichen growth, in the belief that this would bring them "uphill" along the curve towards the maximum. Using the growth curve clearly does not come natural to people, it seems to require some adequate training.

A few comments on the growth curve in the CSIRT task

In the instructions, it is said that all data is perfect. In the light of this information, the shape of the growth curve is a bit surprising. One would expect a smooth curve and not the little dent in the curve at a capacity of 34 average person-hours (see Figure 2). This makes one feel a bit uncertain when estimating the maximum point and inferring the symmetry of the curve. A few more points to the left in the graph would have been nice.

TIME STEPS

Selection of appropriate length of time steps

This is relevant to both of the two different versions, the reindeer management and the CSIRT, of the task analyzed in this paper. In the reindeer-and-lichen system, lichen grows during summer and is eaten by the reindeer during winter when there is no grass. If growth of lichen were plotted in time, it would be zero during winter, start to grow in spring, peak sometime in the summer, be gradually reduced during fall and eventually reach zero in winter. Grazing by reindeer would be practically zero during summer, gradually increase when fall arrives and the grass stops growing, peak during winter, and gradually decline down to zero as grass becomes increasingly available in spring. If the unit of analysis, the time step, is a year, growing and grazing can be summarized over the year, as is done in the reindeer management task. Then it is important to keep using the year as unit of analysis. Taking the lichen growth in a year and divide it by twelve to estimate lichen growth in May, would be just as wrong as multiplying lichen growth in June (or November) by twelve to estimate lichen growth in a year.

Personnel in companies providing IT services tend to attend to customer needs continuously, just as they can perform capacity enhancing activities any time. The time step of a quarter is more arbitrarily chosen. It could just as well have been a month or a week. This means that growth and depletion of service capacity among the personnel in a computer security incident response team have more the characteristics of flow variables.

Estimate flows from changes in stock

It is possible to use the information on Additions and Removals in both the reindeer management task and the CSIRT task to calculate the average growth rate in each time interval ($\Delta t = 1$ year in the reindeer management task, and $\Delta t = 1$ quarter in the CSIRT task).

(12) $$\text{GrowthRate}(t, t+\Delta t) \approx \text{Additions}(t, t+\Delta t) / \Delta t$$

(13) $$\text{DepletionRate}(t, t+\Delta t) \approx \text{Removals}(t, t+\Delta t) / \Delta t$$

Using flow estimates to estimate stock

In the CSIRT task it would be possible to use the estimated capacity growth rate to calculate a reasonable estimate of the growth of capacity in the next month to come, while using estimated growth rate of lichen to calculate the growth of lichen in the next months, or any time interval other than a year, would produce a dramatically flawed result

Selection of time step for graphical integration

When using information on flow to estimate stock by means of graphical integration, it is important to use an adequate length of the time step Δt. This is illustrated by Figure 3 below. If we rewrite Equation (7) using the net flow, we get

(14) $$S(t) = S(t_0) + \sum_{k=0}^{k=n-1} \text{Netflow}(t_0 + k \cdot \Delta t) \cdot \Delta t$$

Let the graph in Figure 3 depict the net flow from time 0 to 2 in an arbitrary time unit. The accumulated stock during this time period corresponds to the area under the graph. With a time step of 1, the estimate of the accumulated growth would correspond to the area of the two large red rectangles, and it would be a vast underestimate of the true accumulation. With the shorter time steps of ¼ time units, the narrower blue rectangles cover an area in better correspondence with the area under the graph, hence producing a better estimate of the accumulated growth. It can easily be seen that ever-shorter time steps would produce increasingly correct estimates of the true accumulation.

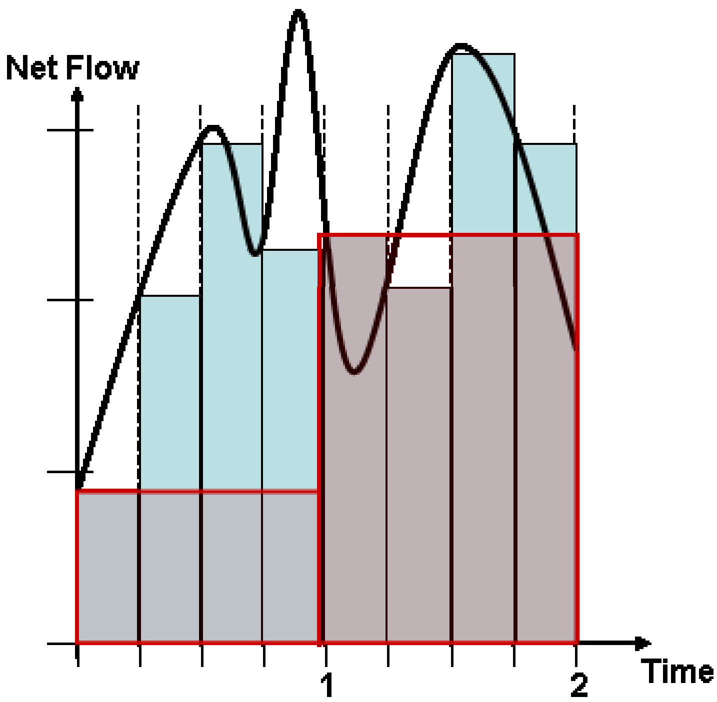


Figure 3. The importance of length of time step when estimating accumulated flow from the net flow.

Note that it is not necessary to understand anything discussed in this section on time steps to produce the correct answer to the CSIRT task, or the reindeer management task, but it is an important part of a thorough and correct understanding of these, and other, dynamic systems.

Table 2. To what degree, and in what ways, the basic the dynamic aspects reported are needed to produce the correct solution (as well as the simplified solution) to the CSIRT task.

Basic Dynamic Aspects	Needed to Solve the CSIRT Task
SYSTEM STRUCTURE	
Model thinking	Model thinking is important, but only at a simple basic for both the full, correct solution and the simplified solution.
Understanding the model concept	Important at a simple level
Model visualization	Important at a simple level
Model construction	N/A (means that it does not apply)
Identify and classify central system variables	Important at a simple level
Identify the relations among the system variables	Important at a simple level (apart from creating the growth curve)
EFFECTS OF FLOWS ON STOCKS	Basic, but not full, understanding needed, (at a very basic level even for the simplified solution)
EFFECTS OF STOCKS ON FLOWS	

Self-generated inflow	Yes, to understand that there need to be capacity for capacity to grow (required even for the simplified solution)
(Stock-dependent outflow)	N/A
LIMITING FACTORS	
Non-linear growth	Yes, to understand that it exists *one* optimal situation (required for a simplified solution at a "higher level")
FUNCTIONAL RELATIONSHIP	
Constructing the growth curve	Obviously for the full, optimal solution
Identifying maximum and minimum variables of a variable	Yes, to use a ready-made growth curve
Understanding the growth curve as a range of possible equilibria	Yes, to make full use of a ready-made growth curve, required for the correct, optimal solution
TIME STEPS	N/A. Not required to solve the CSIRT task, but important for the understanding of dynamic systems in general
Selection of appropriate length of time steps	
Estimate flows from changes in stock	
Using flow estimates to estimate stock	
Selection of time step for graphical integration	

Using the CSIRT Task

In the light of the preceding analysis, what are the prospects of successful use of the CSIRT for the purposes of testing, teaching and communicating with CSIRT managers? The following section offers some answers, but, first, we discuss the description of the CSIRT, because this is important to any use of the CSIRT task, or variants thereof.

THE DESCRIPTION OF THE CSIRT

The CSIRT task instructions were criticized by the university students participating in the Sawicka, Gonzales, and Qian (2005) for being too long. It would probably suffice to inform the participants, as is done on the first page in the instructions (Appendix, p. A1), that a Computer Security Incident Response Team (CSIRT) is a service organization that assists business or governmental organizations with handling computer security incidents, both by dealing with upcoming threats and by taking proactive measures.

Still, the participants requested more and clearer information, while not always able to express exactly what they found lacking, probably because they did not entirely understand the task even after it had been explained to them (Sawicka, Gonzales, & Qian, 2005).

The work capacity, or the capacity to provide services to the customers, is measured in average person hours. One might say that this is the amount of work an average employee is able to provide within an hour. Inexperienced personnel produce less than an average person-hour per working hour, and experienced personnel produce more than one average person-hour's work per working hour. The work capacity of the team in total is the sum of the work capacity of the employees. The example with the CSIRT A and CSIRT B in the instructions (Appendix, p. A2) provides an illustration.

The total work capacity of the CSIRT can be increased either by hiring more personnel or by increasing knowledge in the present personnel, so that each of them are capable of producing more average person-hours. If the personnel are exhausted, having worked too hard providing services, their capacity is probably best enhanced by rest. Competence-building activities are more restful and revitalizing than providing services to customers, meaning that knowledge-enhancing activities are assumed, in the model, to increase capacity either way.

People can only work for so many hours a day producing so many average person-hours depending on their level of expertise. New kinds of malicious attacks on computer systems are invented and spread all the time, so computer security is a fast-evolving field. To remain an expert you need to keep up with the development, old knowledge soon becomes obsolete. This means that unless sufficient time, or capacity, i.e. number of average person-hours, is invested in competence building, the work capacity of the CSIRT will decline.

A novice takes longer to learn new stuff than a more experienced colleague. Learning, or capacity increase, per invested average person-hour increases with the individual's capacity up to a certain level. More knowledgeable employees solve their tasks quicker and are able to take on more difficult and complex tasks than their less knowledgeable colleagues. At some point, however, there starts to be less and less to gain from additional investments in competence building in an individual. Additional knowledge gets increasingly difficult to come by and digest. It is not found it textbooks or available from regular courses any more. More important, perhaps, is that there is not much demand for such advanced knowledge. So, the returns of an average person-hour invested in competence building describe an inverted-U-shaped relation to the present capacity in an individual. (If one wishes people to infer the inverted-U-shape, this last sentence can be left out.)

The effect of investing in competence building is, of course, dependent of staff composition (novice-expert ratio) and how the investment is divided on the team personnel. Hiring new personnel could also increase the capacity. Such considerations are excluded in this model. A "typical" staff composition is assumed, that investments are evenly distributed among the personnel, and that new staff is hired when deemed required.

What one wishes to achieve when running a CSIRT is an optimal division of the available working time, or work capacity in average person-hours, between capacity invested in competence building and capacity used to provide service to customers. One would like to reach and maintain an optimal balance that allows for keeping a high level of capacity in the CSIRT and at the same time proving many services to the customers.

This is what people might need to grasp to understand the CSIRT system, the logic behind the inverted-U-shaped growth curve of the CSIRT capacity, and the goal to reach and sustain the balance between capacity invested in capacity building and capacity used to provide services, and this at a high capacity level. Some of this information is already present, however more implicitly, in the present task instructions (Appendix). This is still quite a lot of information. It would probably be a good idea to highlight vital information in the instructions, and to go through it in class before presenting the participants with the task.

HOW CAN THE CSIRT TASK BE USED?

Test for understanding or performance

In its present shape the CSIRT task tests for several things at once. It would probably be better to test them one by one.

First, with the complete instruction and historical information there is the task to construct the growth curve from the historical information. The simulation is not needed for this. It could just as well be a paper-and-pencil assignment, but then the task has to be one of finding the optimal capacity level, or more explicitly to *construct the growth curve*.

Second, with the growth curve provided, it could be tested if the test takers are able to use the growth curve to a) infer the optimal situation *(identify maximum)*, and b) to calculate what series of adjustments to make to reach it *(understanding the growth curve as a range of possible equilibria*). This could also be done as a paper-and-pencil test.

Third, with the complete instruction and historical information, people could be asked if they believe that there exists *one* optimal situation, a number or range of equally good situations, or that there is no way to know *(understanding of non-linear growth)*.

Forth, given the instructions, the historical information, and explicit information that there exist one, and only one, optimal situation, the simulation could be used to see if the users understand how to search for this optimal situation.

Fifth, given the instructions, the historical information, and explicit information about the optimal situation, the simulation could be used to see if the users understand how to achieve and maintain the optimal situation.

These forth and fifth possible uses both test for understanding of the *EFFECTS OF FLOWS ON STOCK* at the most elementary level.

Teaching about dynamic systems

For students learning about dynamic systems, growth curves might be used, first, to illustrate non-linear growth and to serve as one means to teach the principle of dynamic equilibrium (for a stock to remain constant additions have to equal removals). Later in the course, it might be useful to learn how to construct growth curves from available information. For these purposes, the reindeer management task is probably better suited than the CSIRT task, because its details are more easily understood. People already know about animals grazing and do not need having it explained to them.

If one knows how to construct and use the growth curve, one can calculate the appropriate adjustments to make. There is no need for the simulation, apart from confirming the results of the calculations.

The simulation could be used in two ways. With knowledge about the optimal situation, the users could practice how to reach and maintain the optimal situation. Without any knowledge about the optimal situation, but knowing that there exists an optimal situation, the users could practice searching for the optimal situation. This would hopefully provide the users with a feeling for the workings of the system, and maybe even some grasp on the concept of dynamic equilibrium. There is probably little to gain from simply putting people in front of the simulation to play with it with no prior instructions. This means that the simulation is probably most useful as an illustration when the principles of non-linear growth and dynamic equilibrium are introduced.

A very important possible educational use of the CSIRT task in combination with the one-stock reindeer management task is to demonstrate that even if two systems share the same structure, they still are not necessarily identical. There may nevertheless be very important differences in how the systems are constructed, and how they behave. This could teach the students not to be fooled by superficial likeness. Archetypes are good, but one has to know when likeness ends and differences begin to matter.

Foster system understanding in professionals

If the purpose is to help managers of CSIRTs achieve a better understanding of the importance of continuous investments in competence building in its personnel in order to get good long-term returns, how is this best achieved?

It is probably a good idea, as has been suggested by Sawicka, Gonzales, and Qian (2005) and their participants, not to allow for providing no services at all in a quarter. You cannot just stop providing services to your customers. If you do that you will soon be out of business. It also allows for an extremely large buildup of capacity in a very short time. This is highly unrealistic. By putting a limit on how many services the CSIRT has to provide at least to remain in business, the system would show a more realistic behavior.

It seems unlikely that CSIRT managers will ever be able to estimate a growth curve that provides an *exact* answer, as is the case in the CSIRT task. First, too many simplifying assumptions are made in the CSIRT model, and second, even if CSIRT capacity could be measured it would be difficult indeed to find the relation between present capacity and capacity growth. This means that it is probably not of much use to CSIRT managers to be able to construct a growth curve, as required by the CSIRT task.

The growth curve might be useful as a reasoning tool, however, if properly explained, to demonstrate the vital knowledge to be gained from the CSIRT task (or any of the renewable resource management tasks), that for the capacity (the resource) to grow it is not enough to redirect just a small amount of the present capacity from providing services to competence building. A large enough share has to be redirected for sufficiently long time, for the capacity to grow to a level that will allow provision of many services at the same time as investments in competence building pay off well. It also demonstrates that over-investment in competence building is as bad use of resources as under-investment.

Providing system dynamics insights to professionals presents various difficulties depending on the purpose. If the purpose is to demonstrate the workings of the system of which the professionals are part, or in control of, it is imperative that they understand and accept the model. Here the CSIRT system will most likely not stand to the test. There are too many important aspects missing, and unrealistic assumptions made for the system to gain acceptance as an accurate model of the workings of competence building in an IT company. Increasing competence opens the possibility, not only to provide more services to present customers, but also to enter new markets. Capacity can be increased either by investing in the present personnel or by hiring new, more qualified personnel. Competence building investments pay off differently when directed to novices than when directed to senior employees. Different kinds of investments pay off differently in a short-term or a long-term perspective, or at all. And, if the personnel have been pushed too hard too long rest will have a better effect on performance than competence building, and will be required as an initial investment if subsequent competence building investments are to provide decent returns. These are aspects that need to be included in a model with any claims to realism. This was, however, not the intended use of the CSIRT task[4]. Anyhow, once the model is accepted, and all parameters are agreed upon, how is the resulting behavior of the simulated system best explained?

The systems archetypes (Senge, 1990) have been suggested as one means to provide an understanding of the fundaments of dynamic system behavior, and “flight simulators” (B. E. Bakken, Gould, & Kim, 1994; Sterman, 1992), simple simulations illustrating basic behavior frequently occurring in dynamic systems, another.

The archetypical structure underlying the reindeer-lichen system, as well as the CSIRT system, is a combination of the underachievement and the relative achievement archetypes described by Wolstenholme (2003). Underachievement because the players of the simulations hit upon a resource constraint that would be less constrained if correct actions were taken (less gains now for higher gains later), and relative achievement because the achievement in increased work capacity (or lichen growth) is in conflict with the achievement in terms of provided services (or lichen serving as fodder to reindeer).

The simulation seems to teach the players that cutting consumption enough for the resource to grow will allow for higher consumption later on, the simplified solution (Moxnes, 2004; Sawicka, Gonzales, & Qian, 2005), an insight that they are actually able to transfer from one situation to another (Jensen, 2005b).

Important questions are how archetypical understanding is successfully transferred to real-life complex systems, and how people understand principles without learning the underlying details. As mentioned in the Introduction, we wish people to have a better conceptual understanding of dynamic systems, devoid of technical details. Some answers to how this is best achieved can perhaps be found in the literature on analogical and metaphorical thinking (see e.g., Gentner, Bowdle, Wolff, & Boronat, 2001; Gentner & Wolff, 2000; Hummel & Holyoak, 1997; Lakoff & Johnson, 1980).

Conclusions

It is not surprising that people fail to construct the growth curve, if they have not received appropriate instruction and earlier practice. It is a little more surprising that people fail to use a growth curve, when it is given to them, beyond attending to the maximum (Jensen, 2005b; Moxnes, 1998). People are referred to the simplified solution, to reduce the services provided

[4] Moxnes’s (1995, 1998) initial version of the reindeer management task is far more complex and based on realistic data while still demonstrating the typical behavior reported in this paper.

until capacity starts to increase and then settle for a capacity level they deem appropriate. Since this is as far as most participants get, it is not unexpected that performance in the CSIRT task mimics performance in the reindeer management task (Sawicka, Gonzales, & Qian, 2005). The somewhat more complex setting of the CSIRT task never gets the chance to matter, due to this floor effect in performance.

It is an important observation that the CSIRT task can be solved without proper understanding of some fundamental dynamic systems concepts. When using the task to test for understanding or as an educational devise, it is crucial to be aware of what knowledge might be implied from a correct solution, and what might not.

There are significant differences between the reindeer-and-lichen system in the reindeer management task (Moxnes, 2004) and the CSIRT task, regardless of their identical structure. These differences can, instead of being seen as obstacles, be viewed as educational opportunities. It would in all likelihood be a good experience for students of dynamic systems to compare the two systems and see how two so similar systems can still be different in some quite important respects.

For other educational purposes, the best approach is probably to construct tasks, demonstrations and assignments that treat the aspects one at a time, at least initially, until they have been mastered individually. It is also important to progress in a way that the tasks build upon each other following idea hierarchies discussed in the Introduction.

For testing, as well, it is important to construct tasks and test items that are as specific as possible in what they test for, and ideally, only test for one thing each.

The analysis performed here point to some essential factors to consider in future use of the CSIRT task, as well as some possible improvements of the task, some of which were also suggested by the participants in the Sawicka, Gonzales and Qian (2005) study. The instruction needs to be improved to better explain how the growth of work capacity depends on the present capacity. This may not be necessary for performance in the CSIRT task, but if the task is to be used to demonstrate the workings of real computer security incident response teams, it is crucial.

The most significant contribution of this paper is to demonstrate the importance of a thorough analysis of the demands posed (and not posed) by the tasks we use for testing or education. It also addresses some of the difficulties faced in efforts to enhance dynamic understanding in the public and professionals in other fields than system dynamics.

Good tests of the understanding of basic dynamic concepts are needed, if we want to be able to assess prior knowledge in a population we wish bring new knowledge about dynamic systems. What do they already know? Thorough analysis regarding the *needs* of the populations we are addressing is crucial, as well. What do they *have* to learn? The answer to these questions will differ dramatically for managers of companies compared to university-level students in system dynamics. The answers, however, decides what tasks, demonstrations, and assignments are best suited to meet the needs of the population taught.

References

Bakken, B. E., Gould, J. M., & Kim, D. H. (1994). Experimentation in learning organizations: a management flight simulator approach. In J. D. W. Morecroft & J. D. Sterman (Eds.), *Modeling for learning organizations* (pp. 243-266). Portland, OR: Productivity Press.

Bakken, B. T., & Vamraak, T. (2003, July 20-24). *Misperception of dynamics in military planning : exploring the counter-intuitive behaviour of the logistics chain.* Paper presented at the 21st International Conference of the System Dynamics Society, New York City, USA.

Booth Sweeney, L., & Sterman, J. D. (2000). Bathtub dynamics: initial results of a systems thinking inventory. *System Dynamics Review, 16*, 249-286.

Brehmer, B. (1995). Feedback delays in complex dynamic decision tasks. In P. A. Frensch (Ed.), *Complex problem solving: The European perspective* (pp. 103-130). Hillsdale, NJ: Erlbaum.

Bryant, P., & Nuñes, T. (2002). Childrens understanding of mathematics. In U. Goswami (Ed.), *Blackwell Handbook of Childhood Cognitive Development* (pp. 412-439). Oxford, UK: Blackwell.

Chen, Z. (2002). Analogical problem solving: a hierarchical analysis of procedural similarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 81-98.

Chi, M. T. H., Slotta, J. D., & de Leeuw, N. (1994). From things to processen: a theory of conceptual change for learning science concepts. *Learning and Instruction, 4*, 27-43.

diSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education, 20*, 1155-1192

Dörner, D. (1996). *The logic of failure* (R. Kimber, Trans.). New York: Metropolitan. (Original work published 1989).

Dörner, D., Kreuzig, H. W., Reither, F., & Stäudel, T. (Eds.). (1983). *Lohhausen. Vom umgang mit unbestimmtheit und komplexität [Lohhausen. On dealing with uncertainty and complexity]*. Bern: Verlag Hans Huber.

Flavell, J. H., Miller, P. H., & Miller, S. A. (1993). *Cognitive development* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Gentner, D., Bowdle, B. F., Wolff, P., & Boronat, C. (2001). Metaphor is like analogy. In D. Gentner, K. J. Holyoak & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 199-253). Cambridge, MA: MIT Press.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 45-56.

Gentner, D., & Wolff, P. (2000). Metaphor and knowledge change. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines* (pp. 295-342). Mahwah, NJ: Erlbaum.

Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist, 52*, 35-44.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representation of structure: a theory of analogical access and mapping. *Psychological Review, 104*, 427-466.

Jensen, E. (2005a, July 17-21). *Balancing bathtubs in math class.* Paper presented at the 23rd International Conference of the System Dynamics Society, Boston, MA.

Jensen, E. (2005b). Learning and transfer from a simple dynamic system. *Scandinavian Journal of Psychology, 46*, 119-131.

Jensen, E., & Brehmer, B. (2003). Understanding and control of a simple dynamic system. *System Dynamics Review, 19*, 119-137.

Jensen, E., & Brehmer, B. (2006). Manuscript in preparation.

Kainz, D., & Ossimitz, G. (2002, July 28 - August 1). *Can students learn stock-flow thinking? : an empirical investigation.* Paper presented at the 20th International Conference of the System Dynamics Society, Palermo, Italy.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, G., & Nuñes, R. E. (2000). *Where mathematics comes from. How the embodied mind brings mathematics into being*. New York: Basic Books.

Moray, N. (1987). Intelligent aids, mental models, and the theory of machines. *International Journal of Man Machine Studies, 27*, 619-629.

Moray, N. (1999). Mental models in theory and practice. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application. Attention and performance* (pp. 223-258). Cambridge, MA: M.I.T. Press.

Moxnes, E. (1995). *Reindrift og beitegrunnlag: En simulator [Reindeer management and pastures: A simulator]* (SNF-report No. 83/95). Bergen: SNF.

Moxnes, E. (1998). Overexploitation of renewable resources: The role of misperceptions. *Journal of Economic Behavior & Organization, 37*, 107-127.

Moxnes, E. (2000). Not only the tragedy of the commons : misperceptions of feedback and policies for sustainable development. *System Dynamics Review, 16*, 325-348.

Moxnes, E. (2004). Misperceptions of basic dynamics : the case of renewable resource management. *System Dynamics Review, 20*, 139-162.

Moxnes, E., & Saysel, A. K. (2004, July 25-29). *Misperceptions of global climate change : information policies.* Paper presented at the 22nd International Conference of the System Dynamics Society, Oxford, England.

Ossimitz, G. (2000a). *Entwicklung systemischen denkens: theoretische konzepte und empirische untersuchungen. [Developing systems thinking: theoretic concepts and empirical investigations]*. Wien: Profil.

Ossimitz, G. (2000b, August 6-10). *Teaching system dynamics and systems thinking In Austria And Germany.* Paper presented at the 18th International Conference of the System Dynamics Society, Bergen, Norway.

Ossimitz, G. (2002, July 28 - August 1). *Stock-flow thinking and reading stock-flow-related graphs : an empirical investigation in dynamic thinking abilities.* Paper presented at the 20th International Conference of the System Dynamics Society, Palermo, Italy.

Richardson, G. P. (1991). *Feedback thought in social science and systems theory*. Waltham, MA: Pegasus Communications.

Sawicka, A., Gonzales, J. J., & Qian, Y. (2005, July 17-21). *Mangaging CSIRT capacity as arenewable resource management challenge: an experimental study.* Paper presented at the The 23rd International Conference of the System Dynamics Society, Boston, MA.

Schwartz, B., & Robbins, S. J. (1995). *Psychology of learning and behavior* (4th ed.). New York: Norton.

Senge, P. M. (1990). *The fifth discipline: the art and practice of the learning organization*. New York: Doubleday.

Slotta, J. D., Chi, M. T. H., & Joram, E. (1995). Assessing studet's misclassifications of physics concepts: an ontological basis for conceptual change. *Cognition and Instruction, 13*, 373-400.

Sterman, J. D. (1989a). Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes, 43*, 301-335.

Sterman, J. D. (1989b). Modeling managerial behavior: misperceptions of feedback in a dynamic decision making experiment. *Management Science, 35*, 321-339.

Sterman, J. D. (1992). Teaching takes off: flight simulators for management education. *OR/MS Today, October*, 40-44.
Sterman, J. D. (2000). *Business dynamics : systems thinking and modeling for a complex world*. Boston: Irwin/McGraw-Hill.
Sterman, J. D. (2002). All models are wrong : reflections on becoming a systems scientist. *System Dynamics Review, 18*, 501-531.
Wagenaar, W. A., & Sagaria, S. D. (1975). Misperception of exponential growth. *Perception & Psychophysics, 18*, 416-422.
Wagenaar, W. A., & Timmers, H. (1979). The pond-and-duckweed problem; three experiments on the misperception of exponential growth. *Acta Psychologica, 43*, 239-251.
Wolstenholme, E. F. (2003). Towards the definition and use of a core set of archetypal structures in system dynamics. *System Dynamics Review, 19*, 7-26.

APPENDIX
Task Instructions

Managing CSIRTs [Computer Security Incident Response Team]

Introduction: What is a CSIRT?[5]

Computer networks have become a backbone of any modern business enterprise; they not only facilitate regular business activities, but frequently are central in gaining competitive advantage. At the same time, computer-based information technology has introduced a new and substantial risk to business operations.

Many remember the turmoil and anxiety sparked by the Y2K (Year 2000) problem. Countless publications discussed doom scenarios that could occur given a failure of critical parts of the computer-network infrastructure; costs of recovery from the computer-network downtime were also high on the discussion agenda. While the Y2K challenge was tackled successfully, the lesson was clear: computer information infrastructure is critical for the performance of any modern organization.

Currently, the main threats to the reliability of computer infrastructure are intrusions and malicious attacks by external or internal actors. Most organizations appreciate the need for ensuring the security of their computer networks. Still, even the best IT-security infrastructure cannot guarantee that intrusions and malicious acts will not happen. The resulting damage and recovery costs will depend on the organization's ability to recognize, analyze, and respond to an incident.

A CSIRT – Computer Security Incident Response Team – is a service organization established to assist business or government organizations in handling computer security incidents. A typical CSIRT is responsible for receiving, reviewing, and responding to computer security incident reports. It may also provide a range of proactive services, including issuing early alerts to potential problems or conducting IT-security staff trainings.

CSIRTs range from on-site internal units, which service one particular organization, to national coordination centers, which provide incident handling services to an entire country or region. Most of the CSIRTs do not charge direct fees for their services. Rather, they are funded by their constituency/parent organization (a commercial or government entity) depending on the scope and number of provided services. Development of new services and maintenance of the existing services is therefore crucial to the survival and growth of any CSIRT. However, such growth must be carried out with care. To understand why, it is useful to introduce the concept of '*CSIRT capacity*'.

[5] Based on information published at http://www.cert.org/csirts/ (accessed February 11, 2005)

CSIRT capacity and its service level

CSIRT capacity may be thought of as the ability of the team to provide computer incident related services: the greater the capacity, the greater the challenges that can be handled by the team.

The capacity may be expressed as the total number of average person-hours that a CSIRT is capable of delivering during one hour. The contribution of the individual staff members will vary depending on their proficiency to carry out the tasks. More experienced staff will be able to perform the tasks quicker, thus contributing more capacity than a less experienced or knowledgeable team member.

CSIRT capacity expressed in terms of average person-hours is further elucidated in Figure 1. Both CSIRTs depicted in Figure 1 consist of 5 staff members. Despite the same level of staffing, each team has a different overall capacity. This is due to the difference in the levels of expertise of the staff on each team.

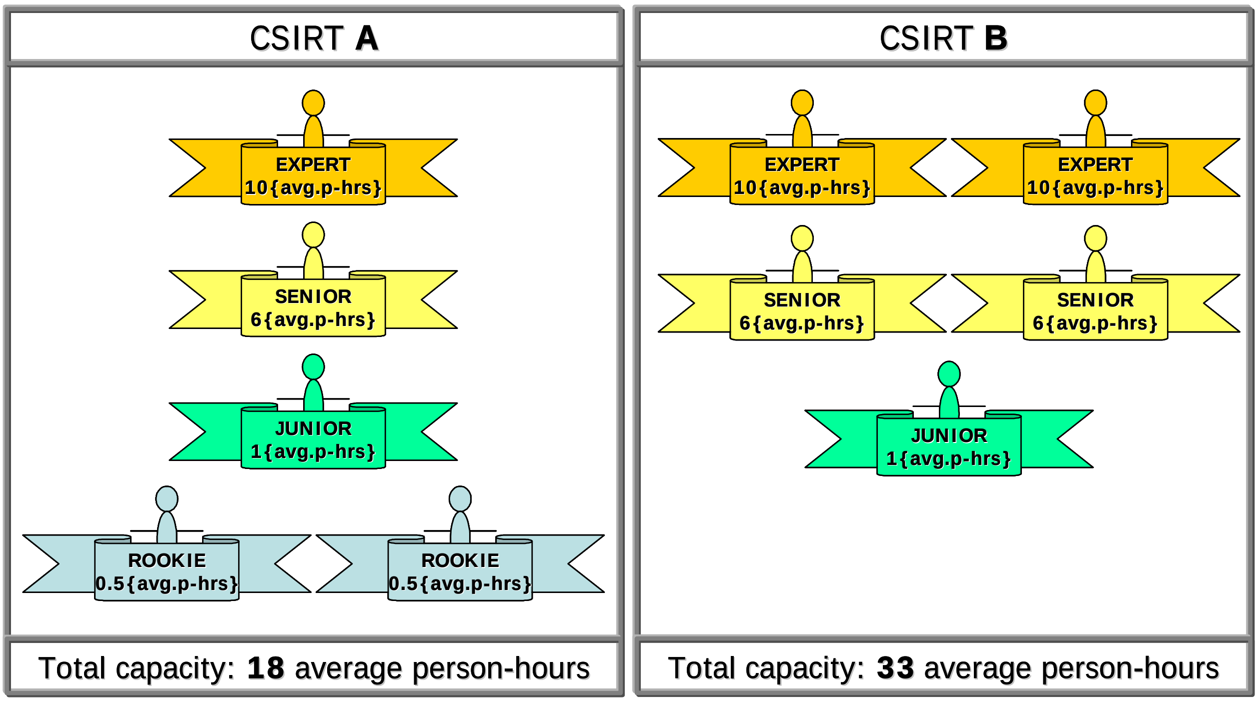


Figure 1 Conceptual illustration of the CSIRT capacity estimation.

The greater the CSIRT capacity, the higher the number of services that can be provided by the CSIRT. However, one should be aware that providing services depletes the overall capacity. This occurs because incident handling is stressful and causes fatigue, decreasing the ability of the staff to respond effectively; consequently, the CSIRT capacity is reduced. The capacity is regenerated when the staff is involved in research, training or software development – these activities are "non-stressful," often creative and exciting. Note that they not only help to reenergize the staff but also prevent the staff's skills and knowledge from becoming obsolete. They may also lead to an increase in the team's capacity (e.g., when new expertise is developed or new tools that allow for more effective incident handling are produced).

These capacity-enhancement activities are indispensable for the survival of any CSIRT; they are the only way to regenerate the capacity that is depleted when providing the services. Moreover, they allow the CSIRT to replace capacity that becomes obsolete or to increase capacity through acquisition of new capabilities. However, one should not forget that these capacity-enhancement activities also utilize some of the existing CSIRT capacity. Therefore care should be taken for the capacity not to be depleted completely by excessive service volume. (Given no capacity, the team will neither be able to provide services nor enhance its capacity.)

Achieving an appropriate and sustainable service level is one of the main challenges faced by CSIRT managers. A sustainable service level means providing the highest possible number of services (ensuring the highest possible funding) while maintaining the CSIRT capacity in the long-run – i.e., the CSIRT is able to restore all the capacity that is depleted when providing services.

In this challenge you will be asked to play the role of a CSIRT manager and to restore a sustainable service level for your team. Detailed instructions for carrying out this task follow.

INSTRUCTIONS

You will play the role of the CSIRT manager. Your funding depends on the number of services provided by your team. Your main objective is to provide as wide a range of IT-security services as possible. Note, however, you should make sure that your operation is sustainable. This means that you should aim for the highest possible sustainable service level.[6] You should also try to reach this desired state as quickly as possible.

The only decision you will make is to set the number of services that will be provided by your team during the next decision period. You make your decisions quarterly for 4 years, i.e. you have 16 decisions to reach the desired state. You will have 3 trials. However, do the best you can in each of them. The participant who gets the best results (i.e., reaches the sustainable service level quickest) will receive a symbolic prize.

You are managing a fixed size CSIRT team that is the only entity of this type operating in your region – i.e., you are not competing with anybody else for constituency support. The primary goal of any CSIRT is to assist its constituency/parent organization. The limiting factor in providing the services is the CSIRT capacity to handle and prevent security incidents.[7]

The net growth of the CSIRT capacity depends on the current capacity level:

- When the CSIRT capacity is low, its net growth will also be low due to insufficient support available for capacity-enhancement activities.
- With high CSIRT capacity levels, the net capacity growth again tends to zero: the only development occurring will concern update and upgrade activities.

In between these extremes, the net CSIRT capacity growth reaches its maximum.

The CSIRT operations may be further characterized by the following piece of information:

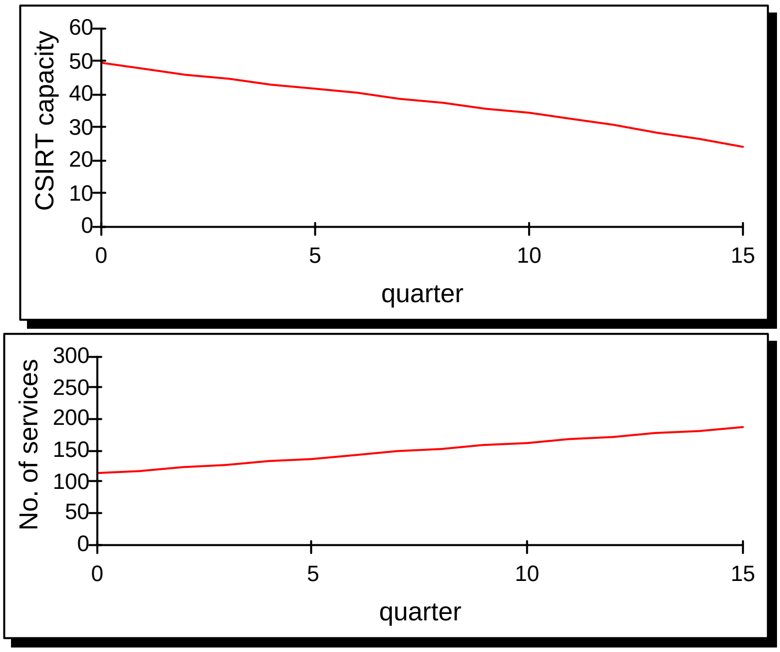
- In one quarter, providing 100 services reduces the CSIRT capacity level by 4 average person-hours. We simplify and assume that the capacity loss estimate is uniform for all services and is independent of the CSIRT capacity level: as long as the CSIRT has any capacity, services are provided (e.g., given the capacity level of 2 average person-hours, the CSIRT will be able to provide only 50 services in the quarter).
- For the purpose of this task you can vary freely the number of services provided by your CSIRT: You can increase or decrease its size arbitrarily at any point in time. The changes will be in effect immediately for the next decision period and will not affect the future funding in any way (i.e., even if you cease providing services for the constituency, the funding will be restored automatically as soon as the services are reestablished).

[6] See p. 3 for definition of sustainable service level
[7] See p. 3 for details

All data regarding the number of services provided by your CSIRT and the CSIRT capacity level are “perfect” (no “noise”, no “distortions”) and there are no random variations in between the decision periods in the service number or the growth of the CSIRT capacity.

Before you take over your CSIRT, you need to know that the previous manager has increased steadily the number of services provided from 115 to 185 over the past 4 years. As a consequence, the CSIRT capacity [average person-hours] has dropped from the initial 50 average person-hours to 24.4 average person-hours at present. This development is shown in the diagrams and table below.

Historical development



CSIRT capacity	No. of services
50.0	115
48.2	120
46.5	125
45.0	130
43.6	135
42.1	140
40.7	145
39.3	150
37.8	155
36.3	160
34.7	165
32.9	170
31.1	175
29.1	180
26.9	185
24.4	*Your 1st decision*