# CLASSIFICATION AND ASSESSMENT OF SLACK: IMPLICATIONS FOR RESILIENCE

Tarcisio Abreu Saurin[1]

[1] Federal University of Rio Grande do Sul, Av. Osvaldo Aranha 99, 5. andar, CEP 90035-190, Porto Alegre, RS, Brazil

[1] saurin@ufrgs.br; Tel: +55-51-9628-2554

**Abstract**

Slack is a key concept for resilience engineering (RE), since it can provide resources for dealing with both expected and unexpected variability. However, in complex systems slack interacts with other elements, and this can imply unexpected impacts, which are not necessarily good for safety and efficiency. As such, slack has an ambiguous nature, and a theory of slack using a RE perspective is necessary. This paper has the objective of contributing to the development of the said theory, by introducing a classification and guidelines for the assessment of slack. An example of using the classification in the pharmacy of a hospital illustrates its applicability. Future studies will focus on the investigation of the extent to which the presented framework is conceptually compatible with RE, empirically justified and motivating for action.

## 1 INTRODUCTION

Complex socio-technical systems (CSSs) are known for tightly-coupled processes, which at the same time account for increased efficiency and facilitate variability propagation in unexpected ways (Perrow, 1984). Tight-coupling usually means the system has little or no slack, which in turn can be a contributing factor to accidents. For instance, NASA´s policy of being "faster, cheaper, and better", ultimately implied slack was gradually degraded, thus introducing brittleness in the system that culminated in the Columbia accident (Woods, 2006). In fact, slack can make systems more loosely-coupled, since it can absorb the impacts of both expected and unexpected variability, providing time and other resources that can support performance adjustment, which is a core characteristic of resilient systems. Furthermore, the need for slack is implicit in the principle of defense-in-depth (Reason, 1997), a well-known safety practice in CSSs. Indeed, by providing multiple barriers and decoupling processes, slack can either slow down the speed of variability propagation or completely block it.

However, slack has its own drawbacks. For instance, it can increase a systems´ opaqueness, disguising small changes and latent hazards which may have non-linear effects (Saurin et al., 2013). Also, badly designed, misused or excessive slack can constitute waste. A number of process improvement methods are focused on the elimination of slack that accounts for waste. From the perspective of these methods, the need for slack is a consequence of unreliable and unstable processes, and therefore the assumption is that slack can be gradually reduced as processes stabilize. Furthermore, excessive slack can be detrimental since the effects of disruptions will not be immediately visible, and thus there will be no pressure to control their underlying causes (Liker, 2004). In other words, excessive slack implies a high threshold for detection of variability.

Regardless of the important and ambiguous implications of slack for CSSs, there is no widely accepted theoretical framework for its investigation from the safety management perspective. The literature seems to be mostly focused on specific forms of slack, such as financial resources and work-in-process in manufacturing plants (Goldstein and Iossifova, 2012), without a broader analysis of the concept and without emphasizing its safety implications and trade-offs. This paper has the objective of contributing to the development of the said theory of slack, by proposition categories of slack as well as guidelines for its assessment in CSSs. It is assumed that understanding how different types of slack can be designed and monitored is relevant for proactive resilience engineering (RE), as it contributes to reducing dependency on opportunistic slack.

## 2 DEFINITION AND CLASSIFICATION OF SLACK

According to Nohria and Gulati (1996) slack is defined as ''the pool of resources in an organization that is in excess of the minimum necessary to produce a given level of organizational output''. While useful, this definition works better for slack formed by resources that can be easily quantified, such as stocks of materials, money, and equipment. For other resources, such as degrees of freedom of employees in relation to standardized work, and cognitive diversity, the definition of what counts as the ''minimum necessary'' is more elusive. Furthermore, the definition by Nohria and Gulati is restrictive as it implies slack necessarily means the addition of extra resources. A less constraining concept is proposed by Fryer (2004), who suggests that slack means available spare resources, of any sort, which can be called on in times of need. Such spare resources do not necessarily mean extra and idle resources, as they may account for existing and strictly necessary resources that may be relocated and used in different ways as needed. This interpretation is in line with the notion that slack can take different forms depending on the nature of the underlying resources (Voss et al., 2008), and on the potential for deploying the resources in various ways and at the time that they are needed. Based on a literature review, ten categorizations of slack are proposed, as follows:

**(i) Origin**: slack may be either designed-in, which usually occurs in tightly-coupled systems, or opportunistic, which usually occurs in loosely-coupled systems, in which slack is often intrinsic to their nature (Perrow, 1984). According to Righi and Saurin (2015), designed slack refers to spare resources whose quantity, place of storage/usage, and nature, are standardized and result from decisions made by groups of individuals and supported by management (organizational level). An example of this type of slack can be stocks of medication with a safety margin that is explicitly sized and visually delimited in stock areas. Opportunistic slack refers to isolated and informal initiatives by staff in times of need – e.g. borrowing a specific piece of equipment from another hospital, and the placement of hospital beds in hallways. This type implies creation of slack via local reorganization (Stephens et al., 2011). Thus, designed slack arises from proactive organizational resilience, while opportunistic slack relies on reactive individual and team resilience, which is often overused (Wears and Vincent, 2013).

**(ii) Nature of the resources**: in principle, any physical or virtual resource may work as slack in a certain context (e.g. astronomers are looking for an Earth 2.0, which may be a slack to Earth in a distant future), although the resources typically considered are time, people, materials, space, and money. Of course, there might be resources which are more elusive and difficult to be quantified, such as perspectives to solve a problem, and degrees of freedom in standardized operating procedures.

**(iii) Availability:** slack may be either immediately available or not. Availability is easier if slack is near to the point of use and decentralized, which tends to favor performance adjustment. Another characteristic of available slack is that the resources are not yet committed to organizational design or specific expenditure – e.g. excess liquidity (Cheng and Kesner, 1997).

**(iv) Strategy of deployment:** five broad strategies were identified. The first strategy, redundancy, may be divided into several sub-categories, such as standby redundancy, active redundancy, and duplication of functions (Clarke, 2005). Clarke provides definitions focused on human redundancy, although the strategy applies to other resources. Standby human redundancy implies the redundant individual is not immediately involved in the task at hand, is typically not present in the operator´s immediate environment, and must be called when necessary (Clarke, 2005) – if the resource in standby is neither loaded nor operational (offline), this strategy may be referred to as ''cold redundancy'' (Hoepfer et al., 2009). Active redundancy means the individual fulfilling a redundant function is involved in the task at hand – e.g. a worker carries out a task while another monitors the performance of that operator (Clarke, 2005). Hoepfer et al. (2009) refer to active redundancy as ''hot redundancy'', to convey that the redundant component is fully loaded and operational. Duplication of functions refers to situations in which two different units perform the same function (Clarke, 2005).

The second strategy for the deployment of slack is through the design of work-in-process (WIP), which refers to the creation of queues between workstations. In fact, this strategy is widely used in manufacturing plants, and it usually accounts for stocks of materials in different processing stages – e.g. stocks of raw materials, partially processed products, and finished products. In manufacturing, the size of WIP is normally a function of the stability of processes; the more unstable, the greater the stocks. In workplaces that adopt the lean production philosophy WIP has a standardized limit, and once these limits are achieved, operations must stop functioning in order to avoid overproduction (Liker, 2004).

The third strategy refers to three types of margins of maneuver, suggested by Stephens et al. (2011). Margin of maneuver type 1 is characterized by maintain local margin by restricting other units´ actions or borrowing other units´ margin. Margin type 2 accounts for autonomous strategies to create margin via local reorganization or expand a unit´s ability to regulate its margin. Type 3 refers to coordinated, collective action of recognizing or creating a common-pool resource on which two or more units can draw (Stephens et al., 2011). The fourth strategy is conceptual slack or cognitive diversity, which refers to a divergence in analytical perspectives among members of an organization. The fifth strategy is control slack, which implies individual degrees of freedom in organizational activity, with some range of individual action unconstrained by formal structures of coordination or command. The fourth and fifth strategies were proposed by Schulman (1993).

**(v) Tolerance**: this refers to the threshold of maximum variability which slack may withstand. A large tolerance means variability is not easily detected, to the extent it does not affect the system´s output. The notion of high tolerance has a parallel with the idea of graceful extensibility, which is characteristic of resilient systems (Woods, 2006). It is also worth noting that tolerance does not necessarily depend on the number of items that form slack (e.g. number of people), but rather it depends on the effectiveness of these items in their role as slack.

**(vi) Visibility:** the status of existing slack should be easily and quickly visible in the workplace (i.e. at a glance), in order to support performance adjustments triggered by scarcity of resources. A study by Righi and Saurin (2015) found a strong correlation between the need for designing slack and the need for giving visibility to processes and outcomes, in an emergency department. For instance, in this type of environment, staff needs to quickly identify physicians on-duty that can be called in times of need, as well as critical medications that have low stocks. There is a substantial body of knowledge on visual management, which can be useful for the design of visible slack (e.g. Galsworth, 2005).

**(vii) Side-effects:** in CSSs, elements are highly interconnected and they influence each other. Therefore, the introduction of slack is not a neutral action, making it necessary to assess side-effects, such as new possibilities of error, increased complexity, and maintenance costs arising from slack. In addition to the impact of context, the side-effects of slack may have a relation with the nature of the resources that constitute slack. For example, slack that is formed by physical barriers is prone to have high maintenance costs (Hollnagel, 2004), and it can be brittle under highly dynamic conditions.

**(viii) Rate of degradation**: while it is not proposed that this category be always quantified, it refers to how long slack maintains its properties even if it is not deployed. A number of factors, which are likely to be dependent on the nature of the resources, can play a role in the rate of degradation. For instance, financial slack may deteriorate due to inflation and unexpected expenditures, slack formed by physical resources (e.g. certain equipment) suffer from wear and tear, and when time means slack, the rate of degradation is expressed in terms of time measurement. Furthermore, it is worth noting that the rate of degradation can be non-linear. This situation may occur, for instance, when unexpected changes in the environment cause an abrupt acceleration/reduction of the pace of consumption of resources (e.g. natural disasters or speculation in financial markets) as well as when technological or organizational changes simply render a given type of slack irrelevant to the intention it was originally devised.

**(ix) Breadth**: it refers to the breadth of sources of variability that slack can match. The more sources of variability can be matched the more general-purpose the slack is. Again, this category seems to be related to the nature of resources, since some of them are intrinsically more general-purpose – e.g. money can be used to purchase and deploy several types of slack, while fail-safe devices in dangerous equipment are specifically built-in to respond to certain types of human error. An important dimension of breadth is related to the adaptability of slack, which is associated with the idea that slack can self-adjust to dynamic variability.

**(x) Hierarchy:** in principle, this category seems to be only applicable when there is a linear chain of defenses. In such cases, hierarchy refers to the position of slack along the chain. Slack that forms the first barrier to a certain source of variability is a first-order slack, and so on.

## 3 GUIDELINES FOR THE ASSESSMENT OF SLACK

In this section, some guidelines for the assessment of slack are presented, as follows:

**(a)** The values and operational goals of the CSS should be identified. In particular, evidence must be sought of how the efficiency-thoroughness trade-off is usually managed (Hollnagel, 2009), and how the organizational policies state it should be managed. This sets a foundation to determine which types of slack are more important, and which are the acceptable levels of slack. For example, critical equipment on standby might be an asset in an intensive care unit, even if this may be seen as inefficiency from a purely financial viewpoint.

**(b)** If possible, the assessment should make a distinction between slack-as-imagined (SAI) and slack-as-done (SAD), similarly with the distinction between work-as-imagined as work-as-done, proposed by Hollnagel (2012). The general equation for the assessment of both types is: $\mathbf{SAI\ or\ SAD_{Ri} = Availability_{Ri} - Necessity_{Ri};}$ where $R_i$ = Resource$_i$. This formula recognizes the possibility of negative slack, when the availability is lower than the necessity and the CSS has to make-do with scarcity of resources. Moreover, the formula makes it clear that negative slack is different from no/zero slack. As an extension of the above formula, the imagined or actual net slack in a CSS may be expressed as: $\mathbf{SAI\ or\ SAD\ Net = \sum Slack_{Ri}}$. The value of these formulas is likely to be mostly conceptual, since their operationalization requires the normalization of all types of slack, so as they have a common unit of measurement. Furthermore, due to the dynamics of slack, these calculations should be made for a snapshot of the CSS.

**(c)** Both the imagined and actual slack should be checked against expected and actual variability, respectively. This may help to identify under and over protected sources of variability.

**(d)** The assessment should account for the dynamic nature of CSSs. Thus, there should be devised means of capturing how the different types of slack interact and how and why they evolve over time under different conditions. The use of the Functional Resonance Analysis Method (FRAM, by Hollnagel 2012) may be useful for this analysis, as each of the six aspects of the FRAM functions may contain slack – i.e. input, output, control, time, resources, and preconditions. It is also worth noting that changes in slack may be due to the changing goals that are typical of CSSs.

## 4. AN EXAMPLE OF APPLYING THE CLASSIFICATION OF SLACK

An exploratory study of the central pharmacy of a hospital (Figure 1) illustrates the applicability of the categories for classifying slack. Data collection involved semi-structured interviews with four employees of the pharmacy, about ten hours of direct observations of the functions being carried out, and an analysis of standardized operating procedures (SOPs). The selected example of slack refers to the degrees of freedom of physicians in the function of requesting medications from the pharmacy. While the formal work system design defines specific channels for requesting medications, at predefined times, about 38% of the daily requests (according to data provided by the pharmacy staff) are classified by physicians as urgent. The interviewed employees reported that they do not question the urgency of the prescription, and limit themselves to comply with the physician´s demand. Nevertheless, the employees also reported that they guess, from their experience, whether the urgency is real. Furthermore, the urgent requests made to the pharmacy do not need to be immediately fulfilled, since they do not involve life-threatening situations. In the patient wards there is specific equipment and medications available to deal with acute life threatening situations, such as a cardiac arrest.

Overall, the high incidence of urgent requests for medications seems has consequences, as follows: (i) in line with the SOP, urgent medications may be dispensed and delivered by the pharmacy without the need for checking the prescription for errors; this check is normally made later, only before the second dose of the medication is given, and therefore patient safety is compromised to some extent; (ii) the frequent urgent requests imply staff involved in the dispensation of medications needs to interrupt their workflow, thus being exposed to errors of prospective memory, and delays in the delivery of the regular requests of medications. It also becomes harder for them to follow standardized operating procedures, which would be useful given that the dispensation of medications is a highly repetitive task. Figure 2 presents a summary of the slack classification.


***Figure 1.*** *Overview of the pharmacy*

| Classification of slack/Description | Requests of urgent medications from the pharmacy |
|---|---|
| Origin | The possibility of requesting urgent medications is designed in the system, although it leaves room for opportunistic use, since physicians can decide when and how frequently urgent requests are made |
| Nature of resources | SOPs, which describe how urgent requests should be made |
| Availability | Physicians can make an urgent request at any time, and the target set by the hospital is to deliver urgent requests in less than 20 minutes. However, no control is made of how frequently that target is achieved |
| Strategy of deployment | There is control slack, to the extent that physicians have freedom to make an urgent request at any time. Furthermore, the medications stored in the pharmacy work as a form of standby redundancy, since they are not present in the workplace where they are needed; they are only moved there when necessary |
| Tolerance | The maximum possible number of urgent requests to be handled by the pharmacy, on a daily basis, is unknown. Currently, about 40% of the daily requests are tagged as urgent |
| Visibility | Information of how many requests are urgent, and which requests are urgent, can only be accessed through the computerized system – a filter can be applied to show the urgencies. This makes it difficult to evaluate, on real-time, how close the system is to its performance limits. Moreover, there is no visual device to identify, at a glance, which dispensed medication is urgent and which is not; the bags containing the medications look all the same |
| Side-effects | As previously discussed, the high use of this type of slack creates hazards for patients, reduces the productivity of staff at the pharmacy, and hinders the credibility of urgent requests, which may sometimes not be taken seriously by pharmacists – in fact, there is a parallel between this situation and the incidence of false alarms (i.e. false urgencies, in this case), which occurs in other sectors. |
| Rate of degradation | The historical evolution of the extent to which this slack is used, was not evaluated. If the percentage of requests classified as urgent has increased over time, this indicates that the consumption of this slack is increasing too. The breaking point of this slack might be an adverse event related to patient safety – e.g. caused by not checking the prescriptions, by delivering an urgent request too late, by delivering and administering wrong medications |
| Breadth | The possibility of making urgent requests can be useful for dealing with a wide range of sources of variability, such as an unexpected evolution of the patient´s condition, the improper handling of medications and equipment in the hospital´s units, and even delays in the processing of the normal requests within the pharmacy – in this case a normal request may become urgent. However, since physicians do not need to justify the reason for the urgency, there is no available data on the relative incidence of these sources of variability |
| Hierarchy | The position of this slack in the chain of barriers depends on the reason for the urgent request. For instance, if the medication given to the patient either did not work or was mishandled (e.g. it slipped out of a nurse´s hand and fell on the floor) , the urgent request for a substitute medication from the pharmacy may be a first-level slack |

***Figure 2.*** *Example of applying the categories for classifying slack*

## 6 CONCLUSIONS

This paper presented a classification and guidelines for assessing slack, so as to contribute for developing a theory of slack using a resilience engineering (RE) framework. An example of applying the classification in a hospital pharmacy illustrates its applicability. Future studies will focus on the investigation of the extent to which the framework presented in this paper is conceptually compatible with RE, empirically justified and motivating for action. Also, the FRAM will be jointly used with the proposed framework, in order to support the investigation of how different types of slack could either facilitate or dampen variability propagation.

**REFERENCES**

Cheng, J. & Kesner, I. (1997). Organizational slack and response to environmental shifts: the impact of resource allocation patterns. *Journal of Management*, 23 (1), 1-18.

Clarke, D. (2005). Human redundancy in complex, hazardous systems: a theoretical framework. *Safety Science*, 43, 655-677.

Fryer, P. (2004). Running an organization along complexity lines. In: Kernick, D. (Ed.) *Complexity and Healthcare Organization: a view from the street*. Abingdon: Radcliffe Medical Press. 289-298.

Galsworth, G. (2005). *Visual Workplace Visual Thinking*. Lean Enterprise Press, Portland.

Goldstein, S. & Iossifova, A. (2012). Ten years after: interference of hospital slack in process performance benefits of quality practices. *Journal of Operations Management*, 30, 44-54.

Hoepfer, V., Saleh, J. & Marais, K. (2009). On the value of redundancy subject to common-cause failures: toward the resolution of an on-going debate. *Reliability Engineering and System Safety*, 94, 1904-1916.

Hollnagel, E. (2012). *FRAM: the Functional Resonance Analysis Method: Modeling complex socio-technical systems*. Burlington, Ashgate.

Hollnagel E. (2009). *The ETTO Principle: efficiency-thoroughness trade-off*. Surrey: Ashgate.

Hollnagel, E. (2004). *Barriers and Accident Prevention*. Aldershot, UK: Ashgate.

Liker, J. (2004). *The Toyota Way: 14 management principles from the world's greatest manufacturer*. New York: McGraw-Hill.

Nohria, N. & Gulati, R. (1996). Is slack good or bad for innovation? *Academy of Management Journal*, 39(5), 1245–1264.

Perrow, C. (1984). *Normal Accidents*: living with high-risk technologies. Princeton University Press, Princeton.

Reason, J. (1997). *Managing the risks of organizational accidents*. Burlington: Ashgate.

Righi, A. & Saurin, T.A. (2015). Complex socio-technical systems: characterization and management guidelines. *Applied Ergonomics*, 50, 19-30.

Saurin, T. A., Rooke, J. & Koskela, L. (2013). A complex systems theory perspective of lean production. *International Journal of Production Research*, 51, 5824-5838.

Schulman, P.R. (1993). The negotiated order of organizational reliability. *Administration and Society*, 5 (3), 353 - 372.

Stephens, R.J., Woods, D.D., Branlat, M. & Wears, R.L. (2011). Colliding dilemmas: interactions of locally adaptive strategies in a hospital setting. *Proceedings of the 4th Symposium of Resilience Engineering*: p 256 - 262. Sophia Antipolis, France.

Voss, G. B., Sirdeshmukh, D., & Voss, Z. G. (2008). The effect of slack resources and environmental threat on product exploration and exploitation. *Academy of Management Journal*, 51(1), 147–164.

Wears T. & Vincent, C. (2013). Relying on resilience: too much of a good thing. In: *Resilient Health Care*, Hollnagel E, Braithwaite J, Wears R (Eds.). p. 135-144. Ashgate, Dorchester.

Woods, D. (2006). Essential characteristics of resilience. In: Hollnagel E, Woods D, Leveson N. (Eds.). *Resilience engineering: concepts and precepts*. Aldershot: Ashgate Publishing; 2006. p. 21-34.