AIAA Space 2007
18-20 September 2007, Long Beach, CA

AIAA-2007-6015

Responsive Space: Concept Analysis, Critical Review, and Theoretical Framework

Joseph H. Saleh[*], Gregory Dubos[†]

Georgia Institute of Technology, Atlanta, Georgia 30332

Customers' needs are dynamic and evolve in response to unfolding environmental uncertainties. The ability of a company or an industry to address these changing customers' needs in a timely and cost-effective way is a measure of its responsiveness. In the space industry, a systemic discrepancy exists between the time constants associated with the change of customers' needs, and the response time of the industry in delivering on-orbit solutions to these needs. Increasingly, the penalties associated with such delays are becoming unacceptable, and space responsiveness is recognized as a strategic imperative in commercial competitive and military environments.

In this paper, we provide a critical assessment of the literature on responsive space and introduce a new multi-disciplinary framework for thinking about and addressing issues of space responsiveness. Our framework advocates three levels of responsiveness: a global industry-wide responsiveness, a local stakeholder responsiveness, and an interactive or inter-stakeholder responsiveness. We introduce and motivate the use of "responsiveness maps" for multiple stakeholders. We then identify "levers of responsiveness," technical spacecraft- and launch-centric, as well as "soft" levers (e.g., acquisition policies) for improving the responsiveness of the space industry. Finally, we propose a series of research questions to aggressively tackle problems associated with space responsiveness.

Nomenclature

a_i	=	activity in the space industry following the issuance of a Request for Proposal
b	=	individual item reduction exponent
τ_{ai}	=	time required to complete activity a_i
c_i	=	cost component in the Total Cost of Ownership
C_{TFU}	=	cost of the Theoretical First Unit produced
$C_{n^{th}}$	=	cost of the nth identical unit produced
MR	=	mission reliability
p	=	pressure
r	=	resources
R	=	traditionally referred to as the learning curve, or learning curve slope, or learning rate
S_i	=	stakeholder in a given industry's value chain
T_1	=	production time of the first unit
$T_{n^{th}}$	=	production time of the nth identical unit
V	=	volume
β	=	compressibility
τ_0	=	total time elapsed from the instance when the need for a given on-orbit capability is identified and formalized to the time when the asset is operational on-orbit and ready to

---

[*] Assistant Professor. School of Aerospace Engineering, Georgia Institute of Technology. AIAA Senior Member. Corresponding author, jsaleh@gatech.edu.

[†] Graduate Research Assistant. School of Aerospace Engineering, Georgia Institute of Technology. AIAA student member.

fulfill the function it was intended for

τ_{LS} = time constant associated with the responsiveness of a local stakeholder

τ_{Si} = total time elapsed for the stakeholder S_i ("supplier") to satisfy its customer's needs

$\tau_{intrinsic(LS)}$ = time component of τ_{LS} associated with the self-responsiveness of a stakeholder (more details in the text)

τ_{inter_resp} = time component of τ_{LS} associated with the interactive responsiveness of a local customer with its suppliers

I. Introduction

Customers' needs are dynamic: they emerge in time and evolve stochastically, prompted by unfolding environmental (political, economic, or technological) uncertainties and network externalities. The ability of an industry to address these needs in a timely and cost-effective way is indicative of its responsiveness.

In the space industry, a systemic and persistent discrepancy exists between the time constants associated with the emergence and change of customers' needs, and the response time of the industry in delivering solutions to address these needs. The needs can consist of a new capability on orbit for a military or a commercial customer, or a modification and repositioning of an existing on-orbit asset. From the moment when the need is identified and requirements are formalized to the time when an operational asset is delivered on-orbit, several years would typically elapse. Increasingly, the penalties associated with such time delays are becoming unacceptable: for example the opportunity loss or failure to secure the first mover advantage in a commercial competitive context, and loss of lives, or lives not saved, in a military context when a needed capability is not delivered on time to the battlefield.

Although different in details, other industries have struggled with conceptually similar issues, and management approaches such as Just-in-Time were developed in part to address the discrepancy between the rate of change of customers' needs and the ability of the industry to deliver timely solutions to these needs (better inventory management also played a role in the Just-in-Time emphasis). "Responsiveness" or "Responsive Space" is the space industry's close analog to the Just-in-Time in other industries.

The word "responsiveness" has become increasingly popular among decision-makers and users of space assets. However, despite its popularity and the emphasis on this concept over the last several years, limited progress has been achieved to-date in practical sense (i.e., limited industry impact) and in the theoretical underpinnings of responsiveness. The conceptual disarray that stymies this subject is best illustrated by the following anecdote: in 2003, an annual Responsive Space Conference was established to bring together various stakeholders in an effort to address problems associated with the lack of responsiveness in the space industry. However, three years later, the conference chair James Wertz acknowledged, "there seemed to be more definitions of Responsiveness than participants."[1] In the same vein, Oderman described the term as being "fuzzy", since "it can mean several things depending on whether you are a developer, operator, or customer for those systems."[2] It is a truism that in order to discuss any subject matter clearly, it is necessary to begin with a clear set of definitions. Much can be gained through careful and consistent definitions. "When words are used sloppily, concepts become fuzzy, thinking is muddled, communication is ambiguous, and decisions and actions are suboptimal, to say the least."[3] In addition to clear definitions, it is important to keep in mind the purpose of these

definitions, that is, why and for what objective(s) they have been formulated. Controversies about definitions often arise from a failure to keep in mind the purpose a definition is meant to serve. In the remaining of this work, we will attempt to provide definitions related to space responsiveness that are actionable and can lead to measurement and quantification.

But while clear and agreed upon definition of space responsiveness is lacking, the need for space responsiveness is unambiguous and deeply felt. Consider the following example reported by Doggrell: "When it became obvious in September 1990, during the planning for Desert Storm, that existing satellite-communications capacity would not support the war effort, we made an urgent attempt to launch an additional Defense Satellite Communications System III spacecraft. That mission finally launched on 11 February 1992, missing the war by over a year!"[4] In addition to the military communications needs, there exist other military space assets that can help save lives and shorten the duration of a military operation or engagement; lack of space responsiveness implies the needed capability is not delivered on time, the consequence of which can be mapped into loss of lives or lives not saved.

It is probably true that space responsiveness was first conceptualized in a military context. But the need for responsiveness pervades the space industry as is equally relevant for military and commercial applications. In a commercial context, responsiveness is needed to gain and sustain a competitive advantage, for example by securing the first-mover advantage against a competing or alternate (not space-based) technology. Conversely, lack of responsiveness results in an opportunity loss and hence, a loss of potential revenue and value to the shareholders. Responsiveness is also relevant and important for scientific space missions. Shotwell and Chinnery[5] argue that more responsive space would allow scientists to observe and study transient phenomena (atmospheric or astrophysical) shortly after they are initiated, instead of many months or years later. And Webb adds that "the desired result gained by [more responsive space] for science missions is the increased return of science in much shorter time horizons at a lower cost."[6]

In short, improving space responsiveness is important for military, commercial, and science applications and end-customers.

In this paper, we tackle the conceptual and theoretical underpinnings of responsiveness. Our objective is to eliminate the "fuzziness" and conceptual disarray that researchers have complained about in the growing literature on space responsiveness. The remaining of this paper is organized as follows. In Section 2, we draw an analogy from controls theory to place space responsiveness on firm conceptual grounds and introduce a metaphorical chronograph for capturing the time dimension of responsiveness. In Section 3, we further refine our preliminary discussion of space responsiveness by drawing another analogy from the concepts of cost structure and Total Cost of Ownership (TCO), and introducing the concept of a "schedule structure" of a space asset. We refer to the "schedule structure" of a space asset as the temporal breakdown of each activity from the time the need for a new (or modified) on-orbit capability is identified and formalized to the time when the asset is operational on-orbit and ready to fulfill the function it was intended for. We show that understanding space responsiveness, and identifying credible ways for improving it, requires that we thoroughly understand, like the TCO concept, the schedule structure of the asset. In Section 4, we address one of the main reasons why responsiveness has remained "fuzzy" to-date, namely who, or what, is responsive in the space industry. To this effect, we introduce three levels of responsiveness, 1) a **global** industry-wide **responsiveness** as seen from the perspective of the end-customer, 2) a **local stakeholder responsiveness** as seen from the perspective of each individual "local customer" in the value-chain, and 3) an **interactive** or **inter-stakeholder**

responsiveness. In Section 5, we further expand on the idea of time compressibility and introduce the concept of "**responsiveness maps**" to help guide responsiveness improvement efforts for every stakeholder in the space industry. Section 6 discusses various levers for improving space responsiveness, and Section 7 concludes this work.

II. Responsiveness: From the control theory to the space industry

A system is commonly characterized as responsive if it can rapidly and with ease react to stimulations and exogenous inputs or events. In so doing, the "responsive" system is in effect coping efficiently with changes and uncertainty in its environment. This informal characterization of responsiveness suggests that time, effort[‡], and uncertainty are essential ingredients for our understanding of the concept of responsiveness in the space industry. Let us first consider the time factor and draw an analogy from control theory: in a controls' context, a (stable) system's response to a stimulation or input consists of a transient response and a steady-state response. The response time of the system is defined as the time elapsed between the onset of the input or stimulation and the time when the system's response gets "close" enough to the steady-state level (often taken at 95% or 98% of the steady-state value). Intuitively, responsiveness measures how "quickly" a system reaches the steady-state response to a specific stimulation; faster systems being more responsive.

In the space industry, the initial stimulation according to which we can start our hypothetical chronograph for capturing the time dimension of responsiveness can be for example the instance when the need for a new on-orbit capability is identified and formalized (i.e., when the Request for Proposal, RFP, is issued). The space industry's response to this "stimulation" is a series of events, design activities, and review gates by which a space asset is developed and readied to address the new identified need—this space asset can be a new one, or an upgrade, a reconfiguration, or a repositioning of an already existing one. Continuing our metaphor and analogy with controls, we can stop our hypothetical chronograph when the asset is operational on-orbit (at the required "location") and ready to fulfill the function it was designed and intended for.

In short, from the perspective of the end-customer or the stakeholder with the need for the space asset, responsiveness is related to the total time elapsed from the instance when the need for a given on-orbit capability is identified and formalized to the time when the asset is operational on-orbit and ready to fulfill the function it was intended for. The other non-temporal dimensions of responsiveness will be discussed later in this work.

III. Responsiveness: From cost structure to schedule structure

In addition to controls theory, we can draw an analogy with another discipline, cost analysis, to help us better understand space responsiveness, and lay the ground for a framework that credibly and practically identifies levers for improving space responsiveness.

The analogy requires that we first understand what a cost structure is. Cost structure is sometimes taken to mean total Life-Cycle Cost (LCC) or Total Cost of Ownership (TCO), the former being commonly used in a government procurement context. For the purpose of this analogy, we do not distinguish between cost structure, LCC, and TCO and use these expressions

---

[‡] Required to initiate and complete the system's response to the unfolding environmental uncertainties, which in turn trigger the previously mentioned stimulations and exogenous inputs.

interchangeably (the minor differences between them will be indicated shortly. These differences however are not relevant for our purposes).

What is the Total Cost of Ownership? As anyone who has bought a capital good (e.g., a car, or an apartment) knows, there is more to the cost of an asset than just the "acquisition cost." Any buying decision that is solely based on acquisition cost is likely to be flawed. It is for example estimated that the 5-year TCO of a "major computing system can be five to eight times the hardware and software acquisition cost."[7] The Total Cost of Ownership includes, in addition to the acquisition cost, the costs to operate and maintain the system, as well as other non-obvious and indirect costs (the condo fees for example for an apartment, the insurance policy, the utilities, etc.). For an IT system, these other costs would include for example the cost to train the support personnel, the maintenance cost, the (opportunity) costs associated with system failure or non-availability, and the cost of power consumption[§] of the system. Figure 1 illustrates a typical Total Cost of Ownership build-up. The figure starts to the left with the cost of the raw materials that go into making the product (e.g., tires). The cost of labor is then added to this first cost category, followed by overhead and other cost categories until the acquisition cost is obtained. Traditionally, this portion of Figure 1 to the left of the acquisition cost is referred to as the cost structure of the product and/or the supplier. The Total Cost of Ownership (TCO) or Life-Cycle Cost (LCC) includes, in addition to the acquisition cost, the cost to operate and maintain the system, as well as other non-obvious costs that will be incurred during the service life of the system.
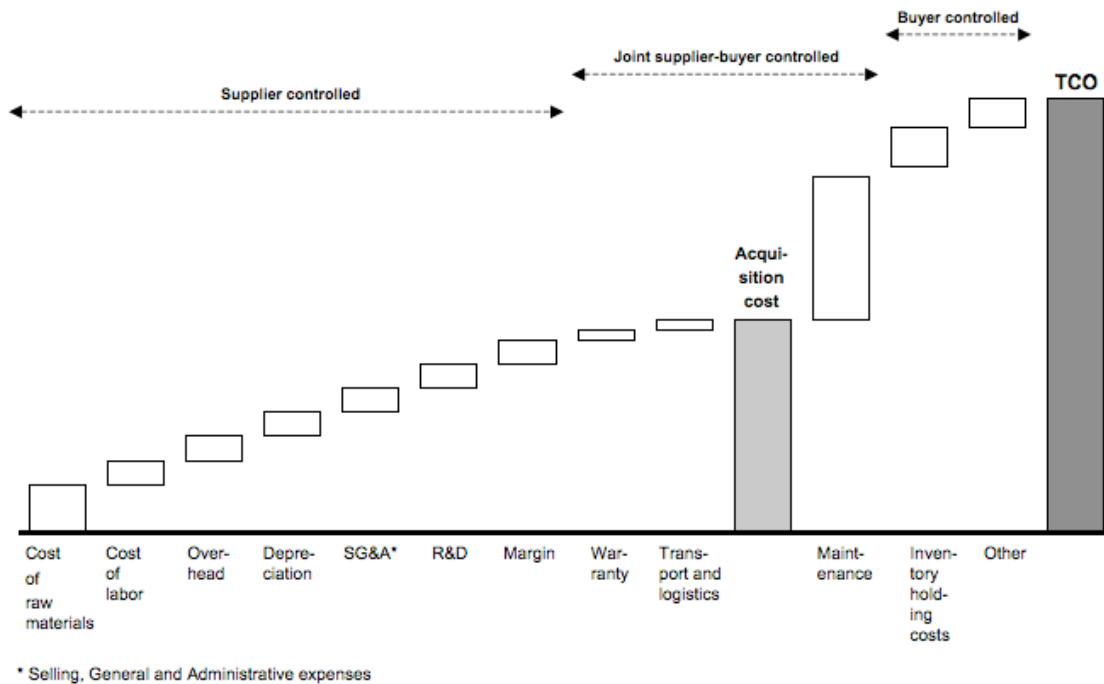

Figure 1. Illustrative Total Cost of Ownership (TCO) build-up.

[§] This is particularly relevant for industries with heavy power consumptions, and in which a power efficient installation can offset a high acquisition cost with a low(er) operation cost compared with a cheaper but less power efficient installation.

What analogy can we draw with the concept of the Total Cost of Ownership to help us better understand and practically improve space responsiveness? In the previous section, we argued that the time dimension of responsiveness can be measured from the instance when the need for a given on-orbit capability is identified and formalized to the time when the asset is operational on-orbit and ready to fulfill the function it was intended for. Let τ_0 be this total time. Conceptually, the TCO analysis suggests an analog breakdown be undertaken, not of cost but of the total time τ_0. The idea is to understand the various activities in the space industry (e.g., design, production, reviews, integration, testing, various bottlenecks and waiting times) following the issuance of an RFP for a new or modified on-orbit capability and how much time each activity contributes to the total time τ_0. This is akin to creating a new accounting system with the various activities as line items, except that instead of "cost accounting" we have a "**time accounting**" of each activity.

In short, understanding space responsiveness, and identifying ways for improving it, requires that we thoroughly understand, like the TCO concept, the schedule structure of a space asset or the temporal breakdown of each activity and how the total adds up to τ_0.

There is however a difference between a schedule structure and a cost structure that should be highlighted. While the TCO is constructed by simple summation of all costs incurred for the acquisition and ownership of an asset, τ_0 cannot be simply taken as the sum of the durations of all activities undertaken to ready the asset for its new on-orbit function. The reason being that not all activities are carried out sequentially and overlap between various activities often occurs and should be accounted for (the difference is one between total "calendar time" and individual activity's "clock time"). Assume that each activity a_i requires a time τ_{ai} to complete, it is easy to see that

$$\begin{cases} \tau_0 \neq \sum_i \tau_{ai} \\ \\ \tau_0 < \sum_i \tau_{ai} \end{cases} \qquad (1)$$

whereas the total cost of ownership of an asset can be written as

$$TCO = \sum_i c_i \qquad (2)$$

In order to improve space responsiveness, it is critical to thoroughly understand, 1) what are all these activities a_i and how do they contribute, or not, to the overall development and readiness of the system, 2) what determines the duration of each activity τ_{ai} and the degree of overlap between these activities—the first and second point bring about the idea of streamlining the workflow and of "time compressibility" as will be discussed later—and 3) where are the bottlenecks, why do they exist and how can they be eliminated. These issues are discussed in the following sections.

It is worth stating explicitly that the activities we have alluded to in the previous paragraphs, a_i, are not only "technical" in nature, but include as well legal, organizational and procedural activities (e.g., procurement, design reviews). Our purpose is to account for all activities that contribute to τ_0, from the ones that are determined by the technical characteristics of the system

under considerations, to the ones that are driven by the organization designing the system, and the ones that are determined by the interactions between and hand-over from one stakeholder in the space industry to another. These issues are discussed in the next section.

IV. Who, or what, is responsive anyway? Three levels of responsiveness

It has probably become evident at this point in our discussion that responsiveness, unlike reliability for example, is NOT a characteristic of a component or a subsystem, but rather a higher-level attribute of an industry's value-chain. More specifically, responsiveness is not an attribute for example of the thermal subsystem or of a particular satellite but rather the characteristic of the space industry at large (hence, "space" responsiveness). Furthermore, although the technical characteristics and architecture of the system under development, a satellite in our case, are key drivers of an industry's responsiveness, or lack of, they are not its sole determinant. Efforts to improve space responsiveness should not only focus on the architecture and technical characteristics of the artifacts created by said industry, but also include the legal, organizational, and managerial aspects of "doing business" in this industry[**]. These issues are further discussed below.

If beauty is in the eye of the beholder, responsiveness is in the eye of the "customer"; it characterizes the reaction time of "suppliers" to an external stimulation (e.g., a new order for product X). However, in an industry value-chain, one stakeholder's customer is often another stakeholder's supplier. Figure 2 provides a symbolic representation of an industry value chain, and Figure 3 illustrates the particular case of the space industry. S_i in Figure 2 are the various stakeholders in this industry and are affected when the end-customer issues a new order for a product or a service[††] (note that in order to avoid cluttering Figure 2, not all the possible links among the various stakeholders are represented). As the end-customer identifies a new need or opportunity and issues a RFP for a new asset, that RFP stirs the industry and propagates upstream its value-chain. Figure 2 illustrates the fact that there are multiple sets of "customers–suppliers" in an industry. Furthermore, one stakeholder's customer is often another stakeholder's supplier. For example, S_{22} is the "customer" of S_{221} and S_{222}, but S_{22} is also the "supplier" of S_2. So who, or what is responsive in this industry value-chain?

[**] One of the authors (JHS) was involved in an effort to develop a product that took two years to complete and fourteen months to finalize an IP (Intellectual Property) and licensing agreement between the various stakeholders. Improving "responsiveness" in this case would more beneficially focus on the legal and managerial aspects of this effort rather than on the requirements and technical characteristics of the product.
[††] Although not important for our purposes, we do make a distinction in this work between an end-customer (who issues the RFP and "pays the bill" for the whole space asset), and the end-user who pays service fees for temporary access to some on-orbit capability (e.g., a transponder).
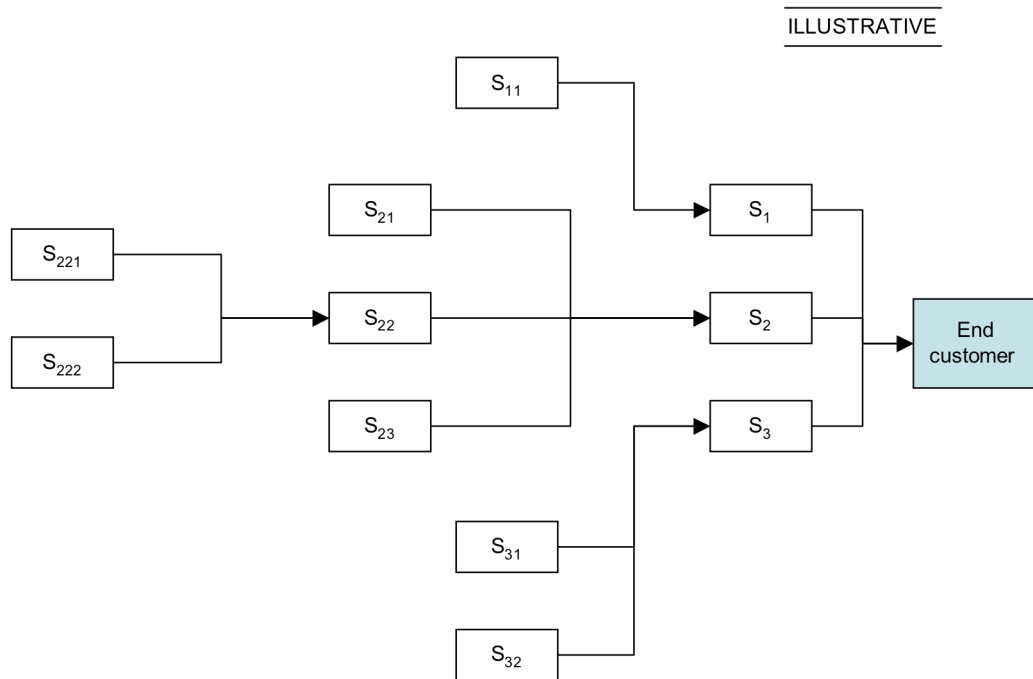
ILLUSTRATIVE

```
                                                                    ┌──────┐
                                                    ┌──────┐        │ End  │
                                 ┌─────┐            │  S₁  │        │      │
                    ┌─────┐      │ S₁₁ │            └──────┘        │cust. │
                    │ S₂₁ │      └─────┘                            └──────┘
     ┌──────┐       └─────┘             ┌──────┐
     │ S₂₂₁ │                           │  S₂  │
     └──────┘       ┌─────┐             └──────┘
                    │ S₂₂ │
     ┌──────┐       └─────┘             ┌──────┐
     │ S₂₂₂ │                           │  S₃  │
     └──────┘       ┌─────┐             └──────┘
                    │ S₂₃ │
                    └─────┘
                    ┌─────┐
                    │ S₃₁ │
                    └─────┘
                    ┌─────┐
                    │ S₃₂ │
                    └─────┘
```

Figure 2. Symbolic representation of an industry value-chain. End customer and various stakeholders (S_i). Not all links are represented.

If the reader agrees with our statement that responsiveness is relevant for "customers" (or stakeholders with needs), and characterizes the reaction time of "suppliers" (or stakeholders addressing those needs, in whole or in part), then the idea of multiple levels and different types of responsiveness becomes evident. In the following, we introduce three levels of responsiveness, 1) a **global** industry-wide **responsiveness** as seen from the perspective of the end-customer, 2) a **local stakeholder responsiveness**, and 3) an **interactive** or **inter-stakeholder responsiveness**. In order to discuss these three levels of responsiveness, it is useful to briefly review the space industry value-chain. Figure 3 illustrates the different stakeholders that partake in and define the space industry value-chain.
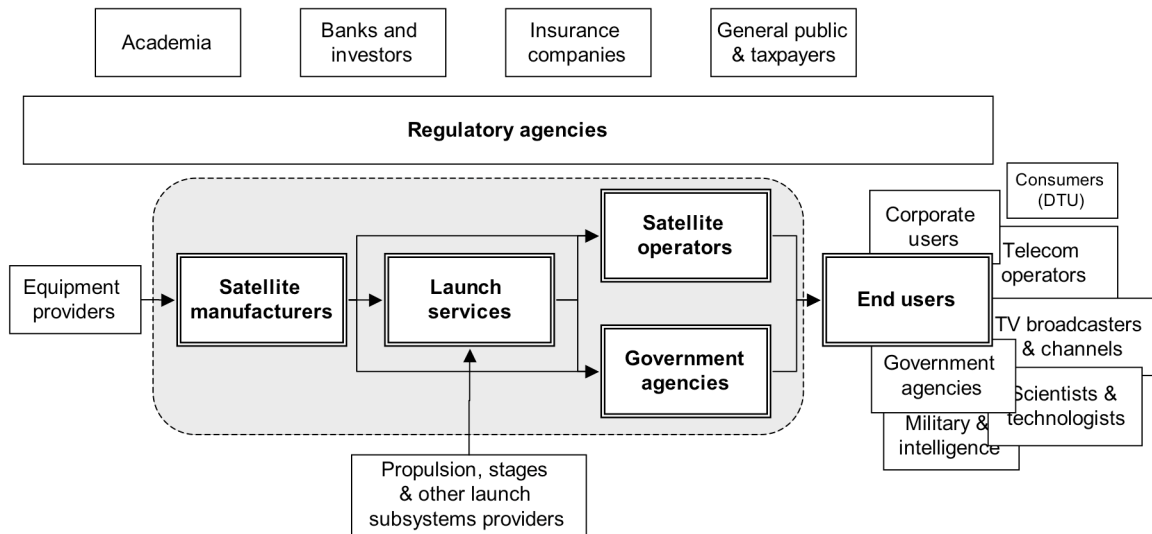
Figure 3. Stakeholders and the space industry value-chain (adapted from Euroconsult 2004)

While the satellite itself is the lifeblood of the space industry, satellite manufacturers are not the only players in this industry. The trends and financial health of this industry are to a large extent defined by the stakeholders downstream the value-chain: government agencies (military and civilian), satellite operators, and end-users. Starting to the left of Figure 3, satellite manufacturers work with a host of suppliers and equipment providers. For example, batteries, solar cells, thermal subsystems, attitude control equipment, and payloads are sometimes acquired from suppliers rather than being developed in-house by the satellite manufacturer—for simplicity, Figure 3 shows only box for the various "equipment providers". Launch providers and range operators, subsumed under "launch services" box in Figure 3, also work with a host of suppliers from engine providers and other propulsion related equipment to launch facility developers and other ground support equipment suppliers. Downstream the value-chain, Figure 3 shows the satellite operators such as Intelsat-Panamsat, SES, or Eutelsat, and government agencies such as the National Oceanic and Atmospheric Administration (NOAA) or the National Reconnaissance Office (NRO). Further to the right, Figure 3 shows a host of end-users, military, consumers, and a variety of corporate users of on-orbit capabilities. Furthermore, banks, investors, insurance companies, and regulatory agencies all participate in the space industry and influence it to varying degrees. And last but not least, the general public and taxpayers are "stake-"holders in the space industry since their money contributes to funding the science and military ventures in space.

Now that we have briefly covered what the space industry value-chain is, we can more clearly discuss where responsiveness resides or what does it characterize in this value-chain, and what the three levels of responsiveness we introduced previously are. We have argued that responsiveness is relevant for the doublet customer–suppliers and characterizes the reaction time of suppliers in addressing customer's needs.

A. Global responsiveness

The global or industry-wide responsiveness is seen from the perspective of the end-customer. This is a "macro-level" attribute of the whole industry. We refer to the end-customer in the space industry as the one who issues the RFP for a given space capability and "pays the bill" for the

whole space asset (as opposed to an end-user who might lease some on-orbit capacity (e.g., a transponder) for a given period of time). Regardless of how the industry is structured, whether there are hundred of suppliers or just a couple of them, from the perspective of the end-customer, what matters is the time elapsed from the issuance of the RFP for a space asset until the asset become operational on-orbit; this is what we referred to previously as τ_0. Symbolically, we can represent this relationship as a block diagram in which the "black box" contains all the suppliers ($\Sigma(S_i)$) that interact with the end-customer (Figure 4). Improving global responsiveness of an industry implies among other things reducing or compressing τ_0 as shown in Figure 5.
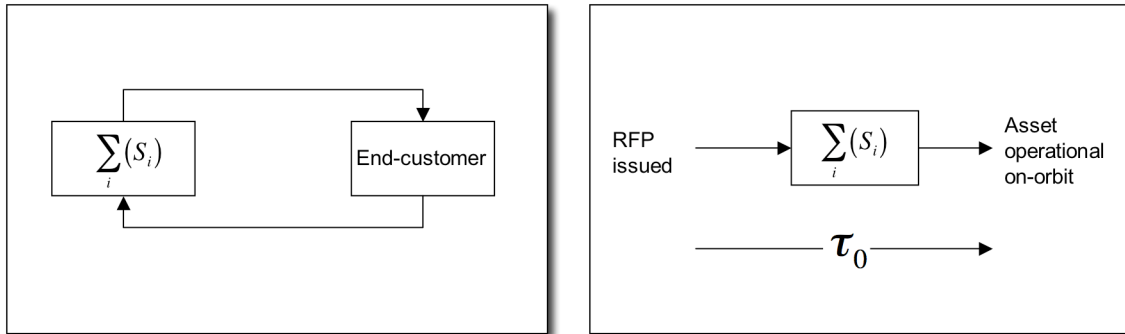
Figures 4. Global responsiveness, end-customer, and block diagram

But how do we do that? Conceptually, compressing τ_0, and consequently improving global responsiveness, can be achieved by three different types of actions, as shown in Figure 5: 1) eliminating bottlenecks in the value-chain and minimizing waiting periods, 2) maximizing overlap, to the degree possible, between different streams of activities at different suppliers, and 3) compressing the "response time" of each supplier. In practice, in order to identify levers for improving responsiveness, we first need to define lower level responsiveness—the constituents or components of this global responsiveness—where practical improvement actions can be taken. Two additional levels are introduced next to this effect, local stakeholder responsiveness, and interactive or inter-stakeholder responsiveness.
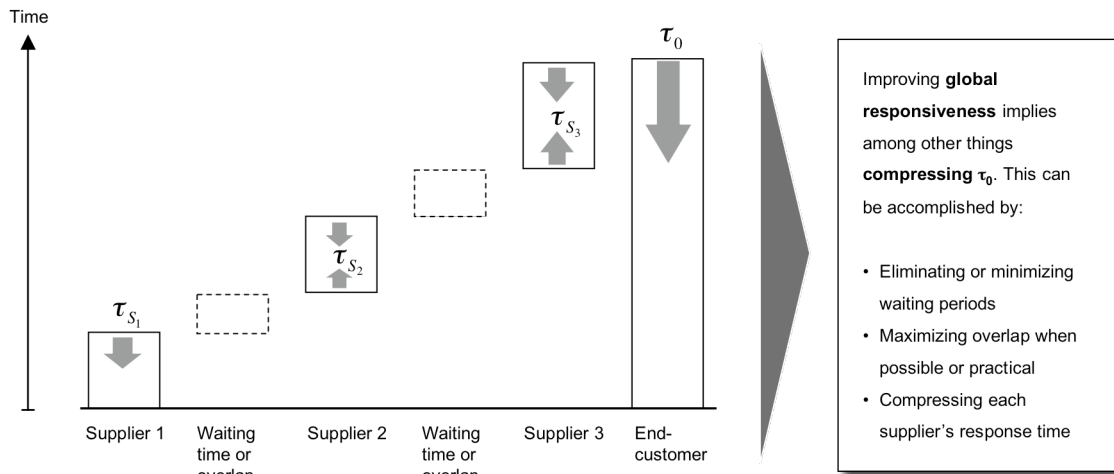
Figure 5. Conceptual improvement of global responsiveness

B. Local stakeholder responsiveness

Beyond the global responsiveness, we can define recursively the responsiveness of each stakeholder as we go upstream a given industry value-chain. The local responsiveness of an individual stakeholder S_i is seen from the perspective of its local customer (as opposed to the "end-customer"). For example, in Figure 2, the local responsiveness of S_{11} is seen from the perspective of its customer, S_1, and is related to the total time elapsed from the instance when S_1 formalizes its needs with respect to a given supplier, here S_{11}, to the time when S_{11} delivers the required product and/or service and fulfills its customer's needs. Let τ_{Si} be this total time. Improving local responsiveness implies among other things reducing or compressing τ_{Si}.
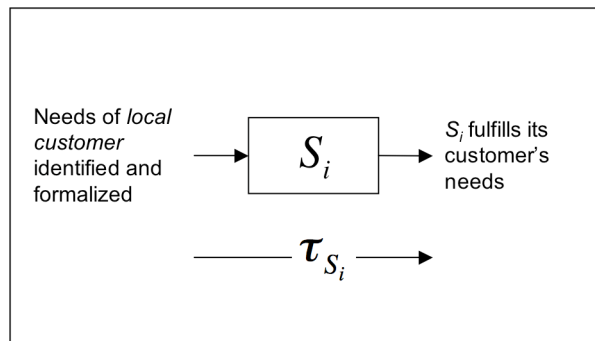

Figure 5. Local responsiveness, individual stakeholder, and local customer in an industry value-chain

A local customer often interacts with multiple suppliers. For example, S_3 in Figure 2 interacts with S_{31} and S_{32} and is their "local customer." A different kind of responsiveness therefore can be defined than the one presented previously: it is related to the total time elapsed from the instance when a local customer formalizes its needs to the time when ALL its suppliers deliver the products and/or services requested by this customer. We prefer our first definition of local

stakeholder responsiveness as it is more actionable than the "cumulative" second one (which "cumulates" the response times of all suppliers of a local customer, and thus dilutes the ability to identify bottlenecks and act upon them).

C. Interactive or inter-stakeholder responsiveness

Who controls, or who has the ability to improve responsiveness? From the previous discussion, it might be concluded that each stakeholder, by improving its own local responsiveness, contributes to improving the global responsiveness. This is true, but is only part of the story. Let us consider another analogy with the Total Cost of Ownership (TCO) to illustrate this point. Who controls the TCO of a system, and who has the ability to reduce it? As Figure 1 shows, there is a significant component of the TCO related to the Acquisition Cost that is controlled by the supplier. Thus, a supplier that improves its cost structure and reduces the acquisition cost will contribute to reducing the TCO. Similarly, a supplier that improves its local responsiveness (τ_{Si}) contributes to improving the responsiveness of its own customer. But Figure 1 also shows there is a part of the TCO that is **jointly controlled** by the customer and the supplier. Thus their joint efforts can also contribute to reducing the TCO. In other words, the TCO is not only the responsibility of the supplier nor is it only controlled by the supplier, but it also includes a joint responsibility of–and control by–the supplier and the customer. Similarly, the responsiveness of a local customer is not only dependent upon and determined by the responsiveness of its suppliers, but also by how well (or efficiently) the customer interacts and works with its suppliers. This is what we refer to as interactive or inter-stakeholder responsiveness; it characterizes the time-efficiency of the interaction between any two stakeholders in an industry value-chain. For example in Figure 2, the responsiveness of S_{22} is not only determined by the intrinsic responsiveness of its suppliers, S_{221} and S_{222} (and the time constant associated with their responsiveness, τ_{S221} and τ_{S222}), but also by the time-efficiency of the interaction between S_{22} and its two suppliers. For example, a customer that can finalize procurement agreements with its suppliers in a few weeks has a better interactive responsiveness than one requiring several months to set up such agreements. We define τ_{inter_resp} as the time constant associated with this interactive responsiveness.[‡‡] We emphasize again that this interactive responsiveness is jointly controlled and determined by the customer and the supplier. Symbolically, we write the time constant associated with the responsiveness of a local stakeholder, τ_{LS}, as a function of the response times of all its suppliers S_i ($i = 1$ to n) plus the interactive responsiveness as shown in Eq. 3:

$$\tau_{LS} = f\left(\tau_{S_1}; \tau_{S_2}; ...; \tau_{S_n}\right) + \tau_{inter\_resp} + \tau_{intrinsic(LS)} \qquad (3)$$

$\tau_{intrinsic(LS)}$ is a time component of τ_{LS} that captures the speed and efficiency by which a local stakeholder (LS) can address its own customer's needs irrespective of, or following its suppliers' responsiveness and the interactive responsiveness, as shown in Eq. 3. *Intrinsic(LS)* can be termed the local customer's "self-responsiveness," and is function of the internal technical skills within the company as well as the managerial skills and organizational structure that facilitates or hamper lean operations and decision-making (Eq. 3 is only a symbolic representation; a more

---

[‡‡] In reality, this time constant should be indexed to reflect the interactive responsiveness of a local customer with each of its suppliers. We have omitted the index in this work to alleviate the notation, and because it does not contribute much to the focus of this paragraph. If we want to be precise however, $\tau_{inter\_respsvns}$ can be defined as the AVERAGE time associated with the interactive responsiveness of a local customer with ALL its suppliers.

formal representation would have to include in addition to f(.), two functions g(.) and h(.) that are monotonically increasing with their arguments τ_{inter_resp} and $\tau_{intrinsic(LS)}$. Such formalism however is not relevant for the purposes of our discussion).

The functional dependence of τ_{LS} on various parameters (Eq. 3) is further discussed in the following two sections.

V. Improving responsiveness: From cost reduction to time compressibility and responsiveness maps

In the previous paragraphs, we suggested that efforts to improve responsiveness (when unqualified, global responsiveness is implied) bears some conceptual similarities to cost reduction efforts in a company or an industry. In this section, we expand on this idea with a brief theoretical presentation, followed in the next section, by a more practical discussion of various levers for improving space responsiveness.

A. Cost reduction

Cost reduction efforts are a fixture in every industry, and they reflect one way by which a company attempts to remain competitive in ever-increasing competitive environments. While cost reduction efforts appear sporadically to the general public in dramatic job cuts following an industry restructuring (e.g., in the U.S. aerospace industry in the 1990s), or heightened competitive pressure (e.g., in the U.S. automotive industry recently), such efforts occur nevertheless on an on-going basis in well-managed companies, and involve many more levers than job elimination (and sometimes none). Many cost drivers can be targeted during these cost reduction efforts. Recall the concept of an industry cost structure and the Total Cost of Ownership; job cuts reduce the "cost of labor" and "overhead," but as Figure 1 shows, there are many other levers to reduce the TCO and improve a company's cost structure.

In short, cost cutting happens on an on-going basis as part of learning and continuous improvement efforts in well managed companies. And it involves a broad panoply of levers to achieve the desired cost reductions.

B. Responsiveness and "time compressibility"

A conceptual parallel can be drawn between cost reduction and responsiveness improvement efforts. We argued in the previous sections that the time dimension of responsiveness is related to τ_0 for the global responsiveness, and τ_{Si} for the local stakeholder responsiveness. Improving a company's or an industry's responsiveness implies among other things compressing these time metrics. Furthermore, responsiveness improvement, like cost reduction, is best accomplished when it is carried out on an on-going basis as part of a learning and continuous improvement process (i.e., when it becomes engrained in a company's performance culture) and not just as a one-time effort. Furthermore, as Eq. 1 and 3 suggest, responsiveness improvement, just like cost reduction efforts, can and should target multiple levers for better results. Some of these levers are discussed in the next section.

In fluid dynamics, a compressibility metric is defined as the relative volume change of a fluid per unit increase in pressure, and is written:

$$\beta \equiv -\frac{1}{V}\frac{\partial V}{\partial p} \qquad (4)$$

More precise definitions of isothermal compressibility and adiabatic compressibility are not relevant for the purposes of this work. By analogy, we can define a time compressibility metric as the relative change in say τ_0 per unit increase in effort or resources. Symbolically, we write:

$$\beta \equiv -\frac{1}{\tau_0}\frac{\partial \tau_0}{\partial r}$$

or $\qquad (5)$

$$\beta \equiv -\frac{1}{\tau_0}\frac{\Delta \tau_0}{\Delta r}$$

This time compressibility metric can be defined for every activity a_i in the space industry that follows the issuance of an RFP for a new or modified on-orbit capability (introduced in Section 3). What use can we make of this metric? This question is discussed in the following subsection.

C. Responsiveness maps

The time compressibility metric of activities helps the analyst and decision-maker to think explicitly about the functional dependence of the schedule for developing a spacecraft on the resources that can be allocated to the various activities in the development and manufacturing process. Consider the following anecdote. In a recent conversation with a satellite manufacturer, we were informed that a major satellite operator has set its decision criteria for selecting the satellite manufacturer for its next satellite as the delivery schedule (i.e., how fast can the manufacturer complete and deliver the satellite) followed by the acquisition cost of the satellite (i.e., at what cost). This situation illustrates the mounting pressure for more responsiveness in the commercial space business, and highlights the possibility of various trades between cost/resources and development and manufacturing schedule.[§§] But how are we to make these trade-offs, or what guidelines can we rely on in our choices? The time compressibility metric for each activity can be one such guide. This metric need not be considered with the analytic rigidity that Eq. 5 may unintentionally imply, but can be assessed qualitatively (e.g., low, medium, high) through the solicitation of experts' opinion and judgment of engineers and program managers. For a given stakeholder, S_i, in the space industry, $\{a_{i,j} \mid j = 1 \text{ to } m\}$ are all the activities performed by said stakeholder to satisfy its customer's needs, and $\tau_{ai,j}$ is the duration of each activity, we define a "**responsiveness map**" as follows (Figure 6): the x-axis is constituted by the compressibility of each activity undertaken by said stakeholder, and the y-axis is the normalized duration of each activity with respect to the total response time of the stakeholder (i.e., how important in duration is each activity, or how much of a chokepoint or bottleneck it is—this can

---

[§§] This is related to the manufacturer's responsiveness, not the global space industry's responsiveness, which would have to include the responsiveness of the launch ranges and the spacelift.

be half-jokingly termed a measure of each activity's "bottleneck-ness[***]"). Each activity undertaken by S_i is then placed on this responsiveness map.
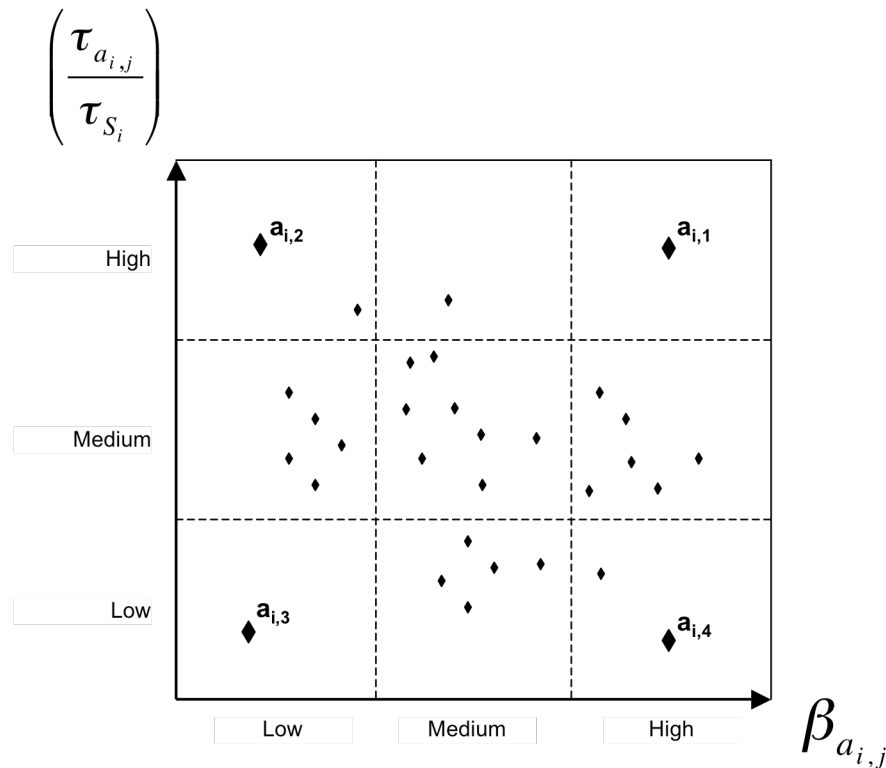

Figure 6. Responsiveness map for a given stakeholder S_i

Figure 6 can be read or interpreted as follows:

1) The **upper-right corner** contains activities that are highly compressible (e.g., $a_{i,1}$), i.e., with limited additional effort or resources (people and/or money) their time to completion can be dramatically reduced. Furthermore, these activities are major contributors to the total response time of the stakeholder, i.e., they constitute important bottlenecks. Therefore, these activities in the upper-right corner should be tackled first in a responsiveness improvement effort.

2) As anyone who has been involved in engineering design and development efforts knows, there are some activities that cannot be easily completed in shorter amounts of time despite additional resources (people and/or money) allocated to them. These activities can be found to the left of Figure 6 in the low compressibility sector. The **upper-left corner** of Figure 6 contains activities that cannot be easily compressed (e.g., $a_{i,2}$), but that constitute important bottlenecks for the company. In other words, the time reduction sought in tackling these activities are more difficult to obtain than in streamlining the activities in the upper-right corner.

[***] This is slightly different from the "critical path" analysis described in the Industrial Engineering and Operations Management literature.

3) The **lower-left corner** contains activities that are neither easily compressible nor do they constitute bottlenecks in the overall workflow of the company (e.g., $a_{i,3}$).
4) The **lower-right corner** contains activities that are easily compressible but they do not constitute bottlenecks in the overall workflow of the company (e.g., $a_{i,4}$).

Responsiveness maps can be developed for every stakeholder in the space industry. And multiple layers or levels of detail can be included on these maps. Once such maps are developed, a company can prioritize its responsiveness improvement effort by tackling activities in the higher priority sectors as shown in Figure 7.
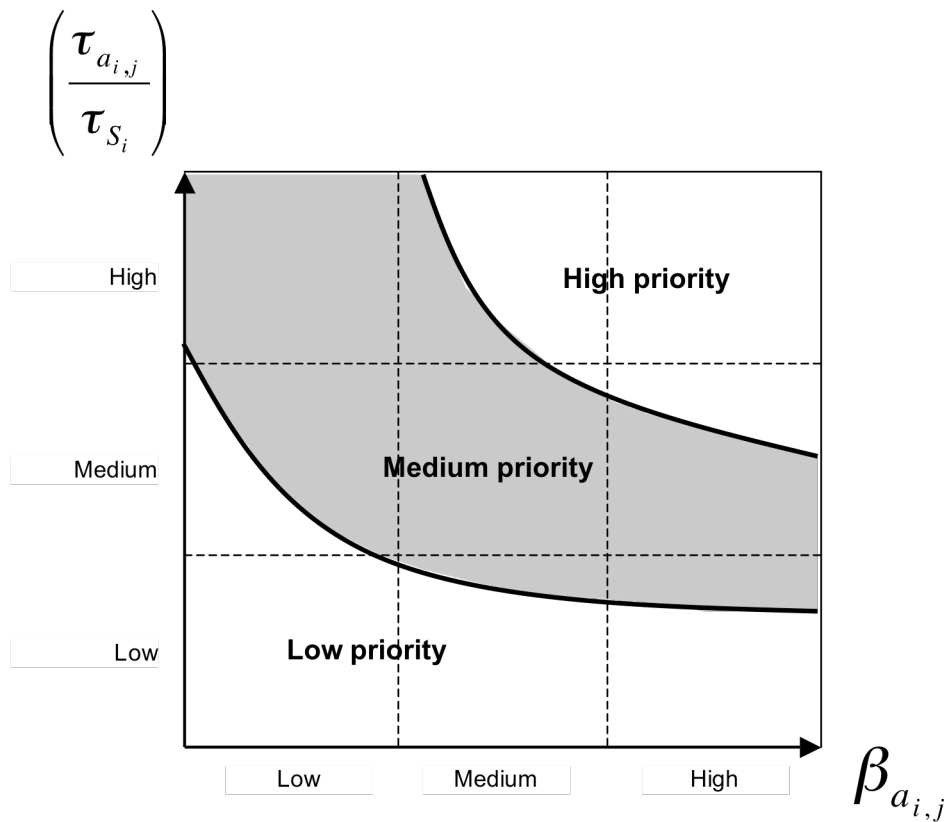

Figure 7. Responsiveness map and prioritization for improvement efforts

Let us now step back for a moment and reflect upon the time compressibility and responsiveness map concepts. The time compressibility metric, we argued previously, is a useful device that helps systems engineers and program managers think about and identify what the schedule for developing a complex system, e.g., spacecraft, is dependent upon. The time compressibility metric, as defined in Eq. 5, captures one important functional dependence, namely the relationship between schedule and resources. The development schedule however, and more generally τ_{LS} and τ_0, are not only dependent upon resources, but also on other "structural" considerations: for example, a change in development process, a modification of program reviews, a change in the architecture of the system under development, or a change in the procurement practices can significantly impact τ_{LS} and τ_0, by modifying or eliminating some

of the activities $\{a_{i,j} \mid j = 1 \text{ to } m\}$. In the following section, we discuss these "structural" considerations, which we refer to as "**levers of responsiveness**" in the case of the space industry.

VI. Levers of responsiveness

Where shall responsiveness improvements be found? What in the space industry can be improved upon, or done differently, and that would result in increasing the industry's responsiveness? The answer to this latter question defines what we refer to as **levers of responsiveness**. To further clarify our terminology, we recognize for example system complexity as a lever of responsiveness, since its modification or change (as a system-level design choice) impacts the responsiveness of the space industry, in particular, the responsiveness of the satellite manufacturer. As an illustration, consider the fact that a highly complex system such as the Cassini spacecraft for example required approximately seven years to develop and launch, whereas the lower complexity Mars Global Surveyor required less than three years to develop and launch. In other words, lowering system complexity, or making design choices that result in lower system complexity, improves the industry's responsiveness. This statement however should not be interpreted as making the case for lower complexity spacecraft design: whether this lever is to be pulled—and to what extent—or not is dependent upon numerous considerations that are not within the scope of this work. A more detailed discussion of this complexity lever will follow shortly.

In a broad sense, improving the (global) space industry's responsiveness can be achieved, on the one hand by improving each or any local stakeholder's responsiveness[†††] (i.e., having more responsive satellite manufacturers, launch providers, and/or launch ranges, and in general more responsive "suppliers"), and on the other hand, by improving each or any inter-stakeholder responsiveness across the space industry value-chain. In the following, we discuss some ways in which this can be achieved. Our discussion is not meant to exhaustive or normative, and in some cases, we simply identify the responsiveness lever and highlight trade-offs between responsiveness and other attributes.

A. Levers related to design choices and architecture of the system under development

As discussed in the previous section, providing additional resources for a given task is not the only way, and sometimes not the most efficient way, by which said task can be accomplished more speedily. In addition to resources, the development and manufacturing schedule of a system also depends on the nature and characteristics of the system under development. These characteristics or system-level attributes, in turn, propagate throughout the system (to the subsystem and component level, and their interfaces), and determine to a large extent the duration of each task necessary to design and field the system. For example, the development and manufacturing schedule of a toaster is different from, and intrinsically shorter than, that of a washing machine, which in turn is different from and shorter than that of a satellite. This difference in the development and manufacturing schedule is dependent upon the architecture and system-level characteristics of these three artifacts. In the following, we discuss some of these characteristics as levers of responsiveness for the space industry.

[†††] Including the suppliers' and the intrinsic responsiveness of a local customer, as shown in Eq. 3 and discussed in Subsection IV-B.

Complexity

Complexity is an increasingly appealing subject for scientists and engineers, as evidenced by the growing number of journals dedicated to the subject,[‡‡‡] and the exponential growth of papers addressing various aspects of complexity.[§§§] Engineers are interested in design complexity, its measures, and implications on schedule, cost, and risk among other things. Both theoretical work and empirical research have shown that design complexity dramatically impacts the development and assembly time of a product. Braha and Maimon for example developed, in a very interesting paper, a complexity measure that can be used as a predictive tool in the early design stages to "estimate the approximate [manufacturing / assembly] times, allowing the comparison of competing concepts."[8] A similar work by Bashir and Thomson[9] demonstrated empirically than the complexity of a system is the dominant parameter in estimating the development schedule (not just the manufacturing time) of an engineering product.

But what is complexity? The following anecdote helps set the stage for an answer. Saint Augustine, in Book XI of his *Confessions*, asks what is Time. "Provided that no one asks me, I know; if I want to explain it to an inquirer, I do not know […] yet what do we speak of, in our familiar everyday conversation, more than of time?" This same answer is applicable if we replace "time" by "complexity." In other words, complexity, and its opposite, simplicity, are easy to understand intuitive concepts, but are difficult to formalize and quantify. Lloyd[10] provides a list of over forty measures of complexity, grouped under three headings: 1) difficulty of creation, 2), difficulty of description, and 3) degree of organization. The Nobel laureate Murray Gell-Mann[11] provides a qualitative discussion of complexity, which is related to Lloyd's classification—Lloyd and Gell-Mann have been collaborators for many years—and Summers and Shah[12] provide an extensive discussion of measures of complexity in the particular case of engineering design. For the purposes of this work, we use *complexity* as indicative of, 1) the total number of subsystems or components used in an engineering system, including identical ones, 2) the number of different kinds of subsystems used, and, 3) the number of interfaces and connections between these subsystems.

Space systems are highly complex artifacts by nature. Their complexity however is not homogeneous and some spacecraft are far more complex than others. The choice of, or the specifications of the requirements that define spacecraft complexity dramatically impacts the responsiveness of satellite manufacturers, as illustrated in Figure 8. There are three ways by which complexity acts as a powerful lever of responsiveness. First, a decrease in system complexity reduces the number (and diversity) of subsystems and payload instruments to be developed, as well as their connections and interfaces; in so doing, lower complexity results in shorter design and development times for the different "parts" of a spacecraft. This first argument can be termed the component-centric influence of the complexity lever on space responsiveness. The second line of reasoning is a systems engineering argument: a decrease in spacecraft complexity reduces the amount of time required to assemble and test the whole spacecraft. Third, a reduction in spacecraft complexity is likely to result in fewer stakeholders and suppliers involved in delivering "parts" to the spacecraft (components, subsystems, and or payload instruments); since fewer suppliers are easier to manage than a plethora of them, reduced

---

[‡‡‡] There are at present over half a dozen journals with "Complexity" in their title, e.g., *Complexity International*, *Computational Complexity*, *Journal of Complexity*, *Progress in Complexity, Information and Design*, and *Complexity* to name a few.

[§§§] Approximately 1200 journal papers per year since 2005, double the average number published a decade ago, and three times more than the number of papers published on the subject fifteen years ago (Source: Web of Science ®).

spacecraft complexity has the potential to improve the interactive responsiveness. In short, spacecraft complexity influences all the parameters identified in Eq. 3, in which the local stakeholder (LS) is the satellite manufacturer.

More research is necessary to quantify the impact of complexity on the design and development schedule of a spacecraft (recent work by Bearden[13] and Bitten *et* al.[14] constitute a very useful step in this work; the authors emphasize the relationship between complexity, schedule constraints, and risk or mission success in planetary missions). Such research can be both theoretical in nature and empirical (involving data collection and analysis); both aspects complementing each other. It would be useful for example to develop parametric schedule models, similar to the traditional cost models, relating schedule to spacecraft parameters such as complexity, mass, power, TRL, payload characteristics, and or other parameters. While satellite cost models are pervasive throughout the space industry, schedule models are practically non-existent. Although satellite cost and schedule are correlated, we believe it is important, when investigating matters relevant to responsiveness, to develop parametric models in which schedule is the independent variable, as opposed to cost. Such a research effort is particularly valuable given the emphasis in recent years on space responsiveness, as it would allow to identify and quantify trade-off—and articulate them to decision-makers—relating spacecraft attributes such as complexity (and other attributes) to the design and development schedule.
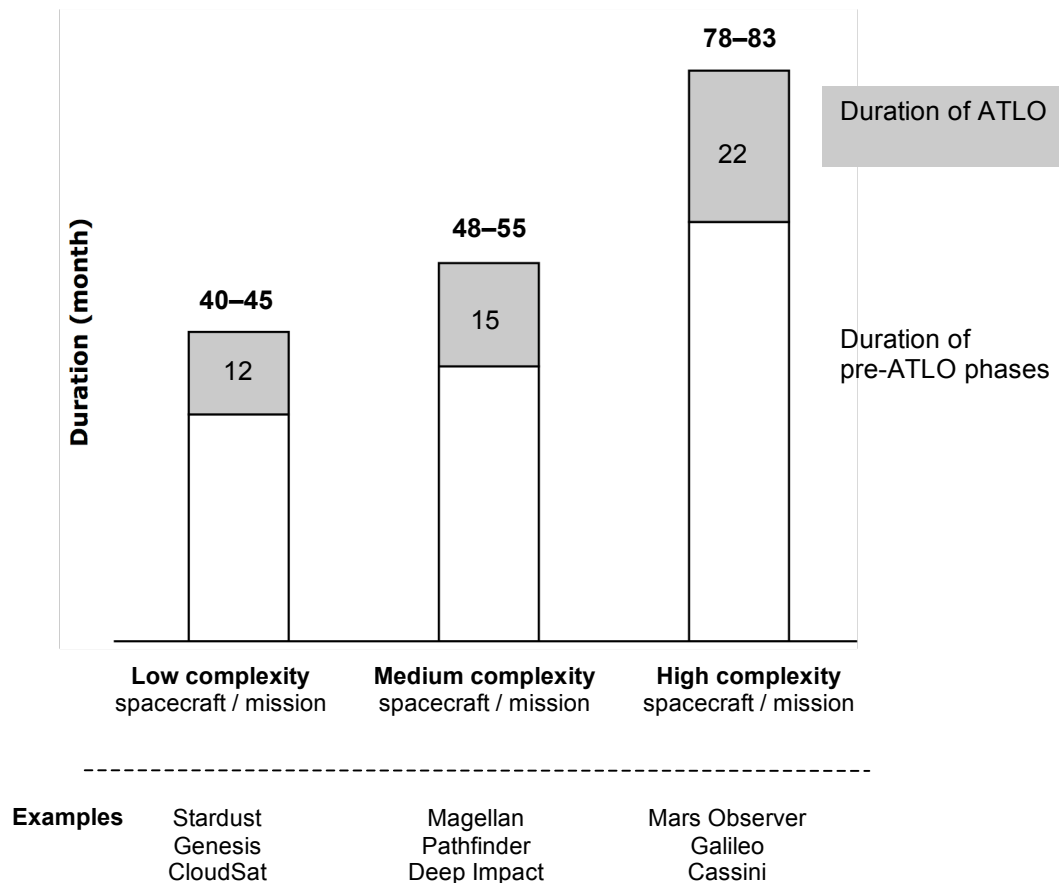

Figure 8. Spacecraft complexity versus design and development schedule (including ATLO)

Heritage, learning curve, and Technology Readiness Level (TRL)

The three terms, *heritage*, *learning curve*, and *TRL*, cover different but closely related concepts in engineering design. They are discussed in this section inasmuch as they act as levers of space responsiveness. The idea of improvement in cost and assembly time resulting from repetitive tasks was probably recognized as early as the American System of Manufacturing—or the use of standardized, interchangeable parts—was devised by Eli Whitney in the late 1790s,[****] and later with the advent of mass production.[††††, 15]

But the idea of such improvements as a function of "repetitiveness" was formalized in 1936 with the publication of a now-famous article by T.P. Wright entitled "Factors affecting the cost of airplanes."[16] Wright proposed a power law to model the improvements resulting from repetitive tasks and production of multiple identical units, *n*; his model has been widely adopted ever since. Wright's model has also been adopted for development and assembly times—see for example Boothroyd *et al.*[17] —in addition to its original focus on cost improvements, and can be written as follows:

$$T_{n^{th}} = T_1 \times n^b \qquad (6)$$

and

$$b \equiv \frac{\ln(R/100)}{\ln(2)} < 0 \qquad (7)$$

The cost analog of this model (Eq. 6) is sometimes written as follows:

$$C_{n^{th}} = C_{TFU} \times n^b \qquad (8)$$

In some publications, the exponent *b* is replaced by, and should not be confused with, the total batch reduction exponent *B* (as opposed to the individual item reduction exponent in Eq. 8). "The learning rate (*R*) for the space and aerospace industry is such that, on average, the nth unit will cost between 87% and 96% of the previous unit."[18]

The application of Eq. 6, and the schedule advantages—or time compressibility—resulting from heritage and learning curve effects are illustrated in Figure 9 and 10. Figure 9 shows the (normalized) reduction in production time for a single unit, the nth identical unit produced, whereas Figure 10 illustrates the cumulative production time for *n* identical units with and without learning effects. *Heritage*, as shown in Figure 10, is the "depth of the past" or the amount of experience in producing identical units (*n*), whereas what is traditionally referred to as the learning curve, or learning rate, *R* is another parameter that determines the improvements (in terms of production time or cost) between two identical and consecutive units produced.

[****] Although credit for the idea of interchangeable parts goes to a French gunsmith, Honoré Le Blanc, who first suggested it in the mid 18th century, but did not go very far with it because other gunsmiths saw this idea as a threat to their livelihood and opposed it.

[††††] Mass production is traditionally associated with Henry Ford and his Model T car. However, although Ford popularized the concept (around 1910), high-volume production of items from standardized parts, i.e., mass production, had been achieved many years earlier. For example, during the American Civil War, the Springfield Armory was producing over 300,000 rifles per year for the Union Army in 1863, almost as many as the peak production of the Model T Ford.[15]

Incidentally, *R* is not appropriately labeled "learning slope" or "learning rate" since, for example when *R* = 100%, no "learning" occurs, and as *R* decreases, more learning effects are reaped. It would be more accurate to designate the complement of *R*, 1-*R*, as the "learning slope" or "learning rate."
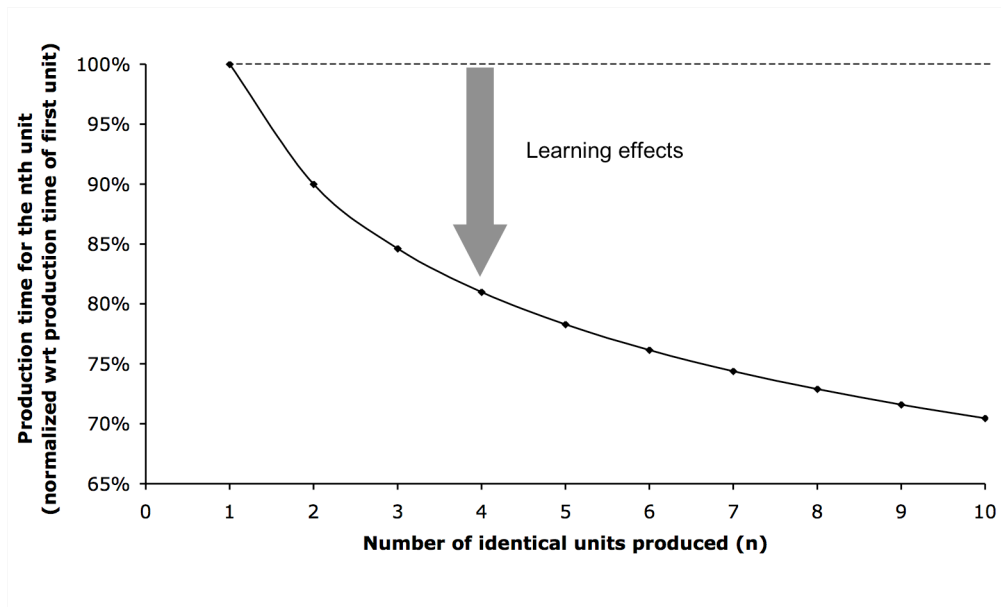

Figure 9. Learning effects and reduction in production time resulting from repetitive tasks (R = 90%)

In addition to *heritage* and *learning curves*, the aerospace community has also developed and widely adopted the concept of *Technology Readiness Levels*, or TRL: "TRL [is] systematic metric/measurement that supports, 1) the assessments of the maturity of a particular technology, and, 2) the consistent comparison of maturity between different types of technology."[19] For example, a TRL 2 refers to a technology concept that has been formulated but no hardware has yet been built or tested; a TRL 6 refers to a system/subsystem that is in prototype stage, its functions and technology concepts having been demonstrated in and outside the laboratory environment; and a TRL 9 refers to an actual system that has been "flight proven through successful mission operations."[19] TRL has been traditionally used to assess the development (and cost) risk of a spacecraft. It is easy to see the close connection between *TRL* and *heritage*. A low TRL for example is conceptually concerned with issues occurring to the left of "1" on the x-axis in Figures 9 and 10, that is, a low TRL refers to the levels of readiness/maturity of a technology **before** a system/subsystem based on said technology is actually built, whereas a high TRL, say 9, becomes analogous to *heritage*. In other words, the concept of TRL is a high-resolution view centered on this "1" and its neighborhood on the x-axis in Figures 9 and 10.
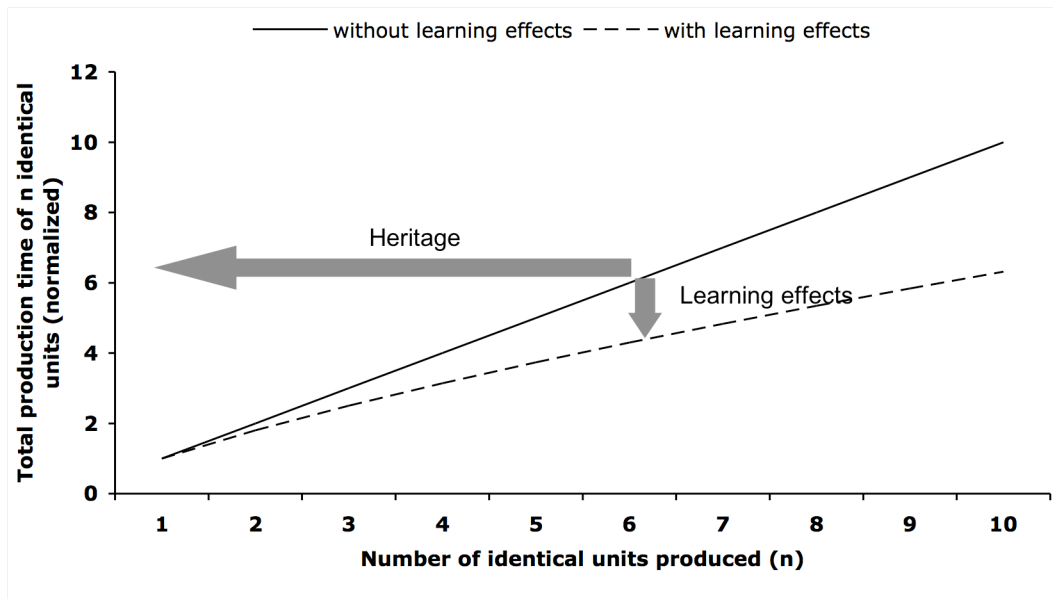
Figure 10. Heritage and learning effects (R = 80%)

It is intuitive to conceive of the TRL, learning curves and heritage as levers of space responsiveness. In the following, we "bundle" to these three concepts under the heading "technology heritage." Technology heritage is an important choice with significant implications on the design of space systems in general, and space responsiveness in particular. For example, whether only in-flight proven technologies should be admitted in response to an RFP, or not, has clear implications on the design and development schedule of a spacecraft.

Technology heritage, like spacecraft complexity, is likely to impact all the parameters identified in Eq. 3, which determine the responsiveness of a local customer—in this case, primarily the satellite manufacturer. No empirical research that we are aware of has been conducted to assess and statistically analyze and model the impact of technology heritage on the development and deign schedule of space systems. Such a study would be particularly relevant, although challenging, in the current emphasis on space responsiveness, as it would help quantify the extent to which technology heritage acts as a lever of responsiveness.

Finally, it is worth noting that technology heritage is a lever of responsiveness that should be manipulated with caution. On the one hand, new or untested technologies (low TRL) might cause schedule delays, thus penalizing responsiveness. On the other hand, high TRL or significant experience (or *heritage*) with a given technology, which has a positive impact on responsiveness, can be indicative of technology obsolescence resulting in uncompetitive value of the system in which said technology is embedded. This argument is conceptually related to the traditional problem of "exploration versus exploitation" in several disciplines such as computer science and machine learning, decisions analysis, and organizational behavior to name a few. It would be

interesting to analytically explore this problem in an engineering design context in general, and spacecraft design in particular.‡‡‡‡

Modularity, Plug-n-Play (PnP), and standardization of interfaces

Modularity in engineering design has been profusely discussed in the technical literature, however its application to, and impact on space systems remain to be further explored. It is discussed in this section inasmuch as it constitutes a lever of space responsiveness.

The many definitions of modularity—see for example Gershenson *et* al.[20], for a good survey of various definitions of modularity—are directly related to and derive from the notion of *module*. A module is a component or group of self-contained components with the following characteristics:

  i.   it has well-defined interfaces to a platform, a system, and/or other modules,
  ii. it provides a specific self-contained function within the system in which it is embedded[21],
  iii. it can be "removed (or interchanged) from a product non-destructively as a unit"[22],
  iv. The following characteristic is not always mentioned or agreed upon in the literature, but we feel it is important for the purposes of our discussion: a module can be easily "plugged" into a system, and both its presence and the function it provides are directly recognized by the system and put to use accordingly.

Modularity, in turn, becomes the property of a system that supports, to varying extent, the use of modules in the design of the system. This definition, when applied to the design of spacecraft, has two major and distinct impacts on the responsiveness of the space industry. The first impact operates at the system or architectural level, and the second impact operates at the module level itself. These are discussed in the following paragraphs. It is worth recalling before getting into our discussion that traditionally, the opposite of a "modular architecture is an **integral architecture** [in which] modifications to any one particular component or feature may require extensive redesign of the product."[23] This property of an integral architecture is to be contrasted with the characteristics (i), (ii), and (iii) of modules/modular architecture discussed previously.

The first impact of modularity on responsiveness operates at the system or architectural level. In the following, we contrast, in a simplified manner, the design, assembly, and test time of an integral design with that of a modular deign in order to clarify and highlight how modularity acts as a lever of responsiveness.

Consider first an integral design (the "opposite" of modularity) in which things are tightly coupled, physically and functionally. Because of the lack of physical and functional separation, the system's development cycle is constrained to a large extent to be sequential, with limited or no possible overlap between different development phases. For example, assembly of interdependent parts can only start once their entire design is complete. Conceptually, we can write the total development time of an integral design as follows:

$$\tau_{integral} = \tau_{design} + \tau_{assembly} + \tau_{testing} \qquad (9)$$

---

‡‡‡‡ This problem in the context of space system design admits several aspects. For example, 1) does an optimal TRL (or level of technology heritage) exist for the space industry? 2) In what context is a low or high TRL more appropriate? And more importantly, 3) what are the relationships and trade-offs between TRL, responsiveness, and competitiveness?

By contrast, modularity gives the opportunity to reorganize the development phases in a non-strictly sequential way. Decoupling of functions between different modules allows a certain degree of parallelism among the tasks performed during the development of the whole system. Since modules are separate providing specific and self-contained functions, they can be designed, assembled and tested separately, that is, simultaneously. Since consecutive operations for an integral design can, for a modular design, be accomplished, to varying extent in parallel, significant time-savings are possible, which translate into responsiveness improvement. Consequently, the first impact of the modularity lever on responsiveness operates at the system or architectural level by enabling overlap between various activities in the development cycle of a "modular" system design. Symbolically, this can be represented by a negative term, $-\tau_{overlap}$, that can be added to Eq. 9.

The second impact of modularity on responsiveness operates at the module level itself. Modularity is sometimes designated in the literature as a "plug-n-play" (PnP) approach, in particular when the characteristic (iv) of modules discussed previously is assumed. Interfaces between modules (and/or between modules and platform) need to be designed in advance, and modules must comply with the standards thus pre-defined in order to be re-used and easily connected, through these interfaces, to the platform or overall system. This upfront investment in modularity and standardization of interfaces presents many benefits, including long-term cost savings. The cost benefits resulting from modularity have been studied by Enright *et* al.[24] in the case of spacecraft design. Similarly, we contend that schedule reduction, or responsiveness improvement, result from the adoption of modular designs and standardization of interfaces in the following manner.

First, certain tasks performed once on a given module do not need to be performed again when this module is re-used (or a similar module being built). By contrast, these tasks have to occur every time a new system is developed in an integral design. Thus, the resulting time reduction consists of a clear discontinuity between the time needed to design, qualify, and integrate a module in a system for the first time and the subsequent times. This effect is particularly noticeable for the design and qualification phases. For example, once a module has been tested and (space-)qualified, its subsequent versions will require limited amount of additional testing before it can be integrate into a new system. In short, the first advantage of modularity (at the module level) consists of eliminating certain tasks during the design and development of modules for subsequent systems, and it so doing, the system's development schedule is shortened and responsiveness is improved.

The second advantage of modularity extends more continuously over time, as more systems using the same module are designed. This second argument takes us back to the discussion of heritage and learning curves. As mentioned previously, tasks that are performed repetitively are subject to important learning effects, during the manufacturing phase for example. The use of a given module (or module design) allows taking full advantage of the learning effects associated with the repetitive use of a module (or module design).

Finally, both effects demonstrate that the benefits of modularity in terms of responsiveness lie in the re-use of the same module (or module design) across multiple systems.

Limited research has been conducted to date to quantify the impact of modularity on space responsiveness. We believe this is an important area of investigation for the space community in a period of heightened emphasis on responsiveness.

Schedule constraints, time compressibility, and mission success: Why mission reliability is NOT a lever of responsiveness

In a recent study, Bitten *et* al.[14] indicated that "during the last decade, NASA's missions were developed in roughly half the time of their traditional predecessors: 3.6 years versus 7.1 years on average." This reduction in development time, or schedule compression, can be construed as a significant improvement in responsiveness of the space industry; in some cases it is, but the argument and analyses put forth by the authors tell a different, more subtle story that require careful attention.

First, the authors demonstrate a strong correlation between "mission success and the time allocated to mission development." Missions developed under strict schedule constraints and inadequate (schedule) growth possibility fail at a markedly higher rate than missions developed under looser schedule constraints. The authors show for example that, for their sample of 66 spacecraft, successful missions took on average 11 months longer to develop than missions that failed.[§§§§] While statistically meaningful, this result should not suggest that schedule compression always compromise mission reliability and success. Instead, this result can be cautiously interpreted as follows: given a mission complexity, schedule compression beyond a certain threshold can compromise mission success.

Second, the authors show that in their sample, planetary missions failed at a significantly higher rate that Earth-orbiting missions. While this observation may be due to the higher complexity of planetary missions, the authors propose a different interpretation. First, they argue that planetary missions, unlike Earth-orbiting missions, cannot afford schedule slip because of planetary alignments and stringent launch window constraints. Second, they provide data that indicate planetary missions, in their sample, incurred a 3.9% schedule growth from the initial planned schedule, whereas Earth-orbiting missions had a 38.3% schedule growth. The authors argue that this schedule growth, which allows for more thorough testing of the spacecraft, explains to a large extent the difference in failure rate between Earth-orbiting and planetary missions (see Figure 11).

[§§§§] The mean development time of successful missions in the authors' sample was 46.5 months, with a standard deviation of 19 months (50 spacecraft). Failed missions had a mean development time of 35.1 months, with a standard deviation of 14.2 months (16 spacecraft).
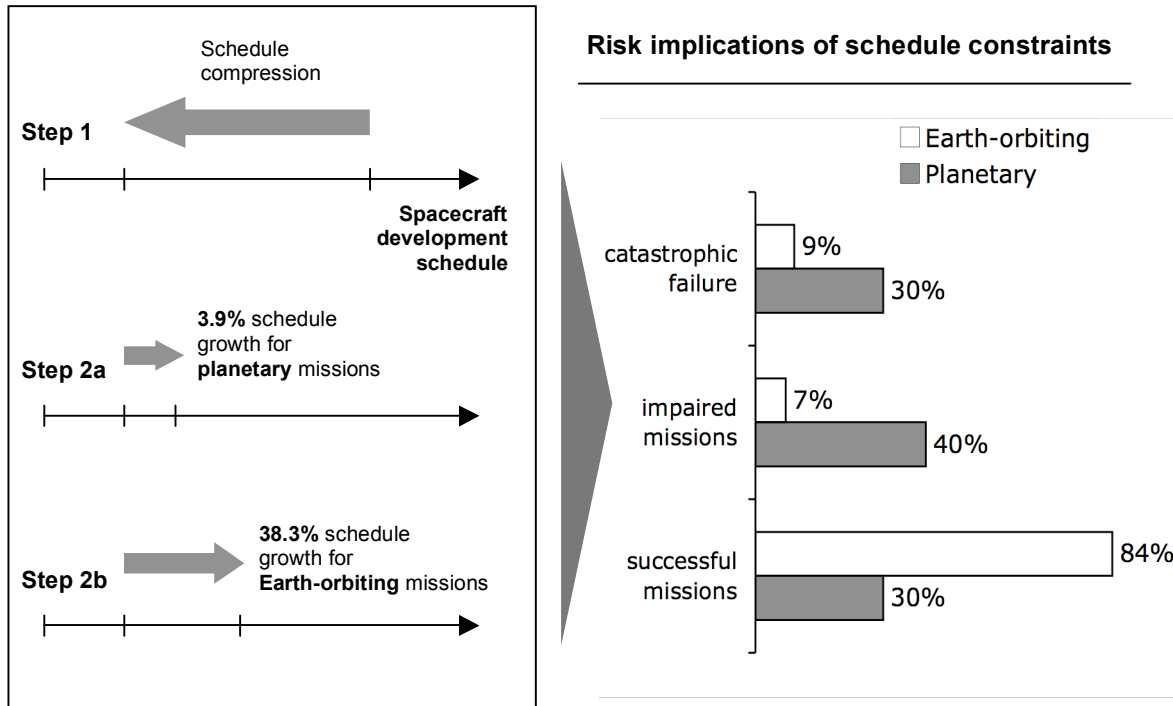
Figure 11. Risk implications of schedule compression and constraints.

In short, the research conducted by the authors, Bitten *et* al.[14] investigated the implications of schedule constraints on spacecraft reliability. Their analyses and interpretation offer a cautionary note against aggressive schedule compression as it might compromise mission success—a result reminiscent of the adage: Faster, Better, Cheaper, choose two!

But the relationship between mission reliability and spacecraft development schedule can be investigated from another perspective than that of Bitten *et* al.[14] In particular, one can ask: can space responsiveness be improved by trading mission reliability? In other words, is mission reliability a lever of space responsiveness?

If the space industry chooses (or can afford) lower mission reliability (*MR*), then a number of activities during the development of a spacecraft can be scrapped or allotted a shorter amount time (e.g., testing). The result is a lower τ_0 than for a more reliable spacecraft. This can be written as follows:

$$\begin{cases} \tau_0 = f(MR;...) \\ \\ \dfrac{\partial \tau_0}{\partial MR} > 0 \end{cases} \qquad (10)$$

However, in our definition of τ_0, we have included a condition that the spacecraft is operational and ready to fulfill the function it was designed and launched for. Compromising mission reliability can be interpreted as violating this condition, and should therefore not be considered a lever of responsiveness in the context of the present work. It should be noted that our argument is confined to mission reliability and does not address subsystem or module reliability.[*****]

B. The launch levers of responsiveness: launch vehicles and launch ranges

The literature on responsive space has early identified launch vehicles as key enablers of the global space industry's responsiveness. This argument falls under what we termed in Section 4 the local stakeholder responsiveness, in which the stakeholder is the launch provider. However, while many papers focused on new launch vehicle concepts and their potential benefits in terms of responsiveness, fewer discussed the required improvements at the launch range level to enable said responsiveness. It is nevertheless essential to acknowledge the respective importance of both launch vehicles and launch ranges—a particular kind of a system-of-systems—as they jointly determine the availability and pace of the overall launch process.

At present, it typically takes several months from the time a spacecraft is shipped from the manufacturer's premise to the launch facility, to the time when it is placed on orbit. This duration is increasingly viewed as an objectionable lack of responsiveness, both for commercial and (especially) military customers. Tackling this problem, the Operationally Responsive Space (ORS) initiative within the Department of Defense aims at launching within 48 hours of call-up. To reach such short turnaround times, one proposed solution requires having launch vehicle parts available "off-the-shelf," so that launchers are built-to-inventory. Such a new approach implies new constraints, among which is the use of propellants capable of being stored at ambient temperatures. Storability of vehicle "parts" is therefore a necessary condition for this solution to ensure responsive launch, and vehicle prototypes such as the SpaceDev's Streaker are currently under development to satisfy this condition.[25]

Procedures governing the assembly of vehicle parts and the transport to the launch pad also determine the rapidity of a launch. Implications of the propellant choice extend to these phases, as toxic and hazardous fluids require special, and therefore time-consuming handling. Also, ease of transport and accessibility of the vehicle can significantly reduce time lags preceding launch. For example, vehicles from the Scorpius family are assembled at or near the primary launch site, transported vertically on cradles or flatbed trailers, and their short size allows them to be serviced directly from the ground, without the need for large service towers.[26] Other design choices translate into quicker pre-launch operations and checkout. The Falcon I by SpaceX is equipped with a fairly unique thrust termination system replacing the more common explosive destruct device installed on most of the launchers. As a result, "Space X estimates at least one day of operations were saved by removing the explosive ordnance from the system."[5] If new technologies and handling methods need to be studied to meet Responsive Space objectives in the long-term, their development also requires an additional time investment. Heritage as a lever of responsiveness discussed previously is therefore as relevant for launch systems as for the payload itself. Thus, vehicles from the Minotaur family created by Orbital Sciences Corporation have been designed to take advantage of previous developments by the company: "heritage, and sibling relationships, [...] make them imminently adaptable to responsive launch requirements."[27]

[*****] Module reliability contributes to, but does not fully determine, mission reliability. System architecture (or *structure*) is an equally important contributor to the overall mission reliability.

In a larger sense, the major problem underlying launch responsiveness, or lack of it, resides in what is referred to in the Operations Research literature as the **build-to-order** versus the **build-to-inventory** production approaches (see for example Holweg and Pil[28]), and the associated financial risks and cost implications, borne dissymmetrically by the various stakeholders in the space industry. In the following, we expand on this observation, which we believe is one of the root causes of lack of launch responsiveness.

What are the build-to-order (BTO) and build-to-inventory (BTI) production approaches? Launch vehicles today are effectively built-to-order, that is, they are built for a specific mission/spacecraft and after a confirmed order–with all the financial guarantees–for the vehicle has been placed[29]; the build-to-order approach is sometimes referred to as "pull" production system in which the market effectively "pulls" the products from the manufacturer. By contrast, the build-to-inventory is a "push" production approach in which products are manufactured (and sent to the "inventory") not in response to confirmed orders, but in the hope that "pushing" said products onto the marketplace will result in them being purchased. It is easy to conceive of hybrid production approaches that lie between these two ends of the spectrum (BTO and BTI) and for which 1) products are built in part to order, and in part to inventory, 2) products are built to (credible) sales forecast, 3) products are built with varying degrees of commitments from the customers (shy of firm orders).

What are the economic implications of the BTO and BTI approaches? The literature on responsive launch advocates a shift from build-to-order to build-to-inventory approach to launch vehicles in order to ensure responsiveness. We do not disagree with this recommendation. However, the absence of discussions in the literature of incentives, risk exposure, and economic implications for the launch providers to shift from BTO to BTI suggests either a lack of understanding of corporate decision-making or an economic immaturity of this literature that gets it ignored or leave limited impact on the industry. The analysis of the risk and economic implications of a shift from BTO to BTI to the launch providers requires particular attention and deserves a treatment of its own; while such an analysis falls outside the scope of this work, its general outline is highlighted herein (and summarized in Figure 12):

1. Launch vehicles are highly complex and costly artifacts—with a price tag roughly around $100 million for many of them. The design of launch vehicles is driven by and matches the present day dominant design of spacecraft as large monoliths (see Anderson and Tushman[30], and Henderson and Clark[31] for the seminal work on dominant designs and disruptive technologies and architectures).
2. Given the high cost of a launch vehicle and the low volume nature of the launch business, launch providers cannot afford the financial risks that come with the build-to-inventory production approach, or the significant inventory holding costs associated with this production approach. The build-to-order approach therefore is both a lower risk and cost approach to the launch providers than the build-to-inventory.
3. From a customer's perspective however, the build-to-order of launch vehicles, unlike the build-to-inventory, is a non-responsive approach and results in significant delays before a needed capability is placed on-orbit. Launch responsiveness, as seen from the end-customer's perspective (e.g., the Air Force), is therefore traded against lower financial risks and inventory costs by the launch providers.

Enticing the launch providers to switch from a BTO to BTI, and hence towards a more responsive production approach, will succeed only when credible (and creative) economic

solutions are found to 1) distribute the financial risks between the launch providers and the end-customers, and 2) have the latter share in the inventory holding costs associated with the build-to-inventory approach. This we believe is an important topic for further research and in-depth analysis.

Another hypothetical solution for the switch from BTO to BTI is to dramatically lower the cost of launch vehicles. This can only happen if the current dominant architecture of spacecraft (large monolith) is significantly challenged (or disrupted) and a new spacecraft design paradigm emerges and proves competitive, such as powerful micro-satellites or fractionated spacecraft.
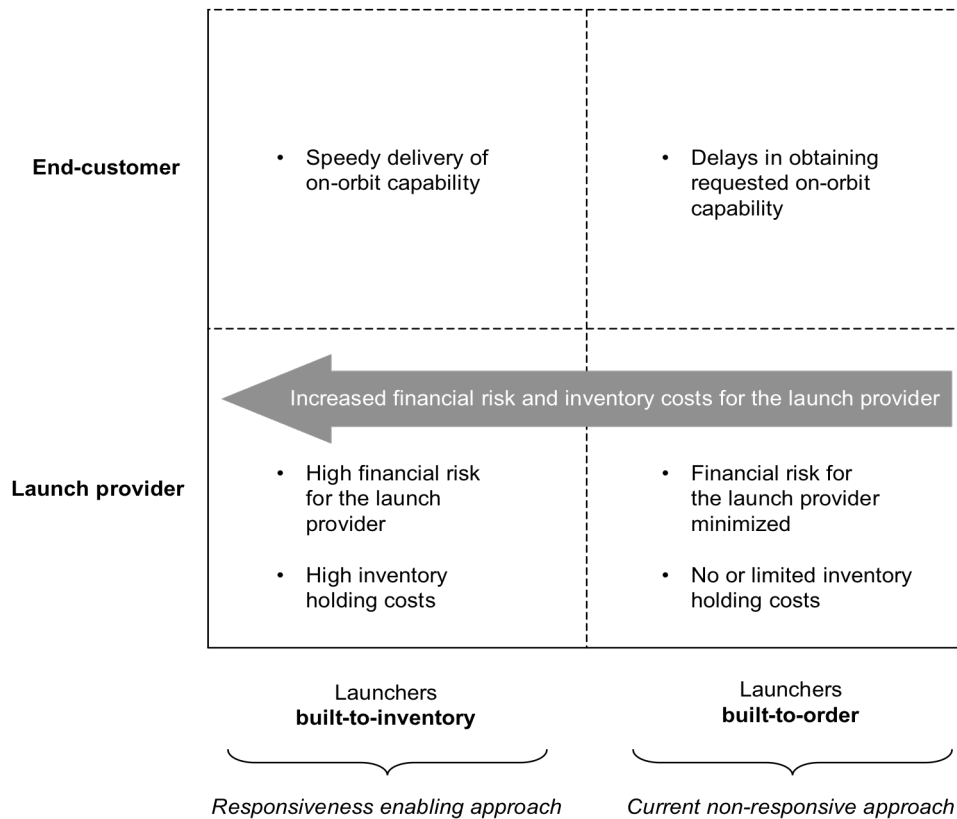

Figure 12. Responsiveness and implications of BTO and BTI to launch providers and end-customers.

We now turn to a brief discussion of launch ranges as they pertain to space responsiveness. Just like airports have a limited capacity to handle air traffic, so do launch ranges have a limited launch turnover rate. Saturation of the launch range capacity can generate an important "bottleneck" representing a challenge for the responsiveness of the space industry. It is often argued that the current number of launch ranges around the world is not sufficient to satisfy the demand without generating waiting periods. In addition, most ranges are government-owned, and function under significant restrictions that often result in delays of the order of months in their operations (i.e., lack of range responsiveness). To tackle this problem, several initiatives have recently emerged to build private launch ranges that would allow more lean operations and would

"un-choke" the current flow of demand in launches. For example, the Mojave Spaceport became the first facility to be certified as a spaceport by the FAA in 2004, and allowed the flight of X-Prize's winner SpaceShipOne. Similarly, Spaceport America, built in New Mexico in 2006, experienced its first successful launch of a SpaceLoft XL rocket in April 2007.

C. Selection, acquisition, and reviews: the "soft" levers of space responsiveness

In discussing the levers of responsiveness, we first focused in Section 6 on the system-level attributes—such as complexity and modularity—of the spacecraft under development, and explored how changes in these attributes impacted the development and manufacturing schedule of the spacecraft. Their impact on the development and manufacturing schedule in turn translates into changes in responsiveness of the space industry, hence our classification of these attributes as "levers of responsiveness," or more specifically as **spacecraft-centric levers of responsiveness**.

Following our discussion of spacecraft attributes as levers of responsiveness, we addressed the **launch levers of responsiveness** as they pertain to changes in launch vehicles (design, operation, and shift from BTO to BTI approaches) and launch ranges.

But as mentioned previously, efforts to improve space responsiveness ought not focus solely on the technical and operational characteristics of the artifacts created by the space industry (e.g., spacecraft and launchers), but should also address the legal, organizational, and managerial aspects of "doing business" in this industry. These aspects pertain to both the interactive responsiveness, as we defined it in Section 4, and the "self-responsiveness" or the $\tau_{intrinsic(LS)}$ term introduced in Eq. 3. In this section, we focus on these **"soft" levers of responsiveness** (as opposed to the technical or engineering levers discussed previously) and briefly discuss how improvements to the following can impact space responsiveness: 1) the selection process of competing proposals, in response to an RFP, for a particular on-orbit capability, 2) the design reviews during the development process of a space system, and 3) the acquisition policies of space assets, in the particular case of military acquisition. Our coverage is not meant to be exhaustive of all soft levers of responsiveness; for example issues pertaining to supplier management are not addressed here, but we believe they are important and deserve their own careful treatment from a management science and operations research perspective.

Selection process

Selection of competing proposals for a space asset is included in our work as a lever of responsiveness because of the way we defined the time constant associated with the global responsiveness of the space industry τ_0—starting from the time when the need is identified and formalized, and not from the time when the program is initiated. We recognize that different definitions of responsiveness can exclude the selection process for the levers of responsiveness; in the following, we argue why we believe its inclusion is relevant.

No research that we are aware of has been conducted to date to benchmark the selection process of competing proposals for on-orbit capabilities in a military acquisition context (e.g., Air Force, NRO), in a government civilian context (e.g., NASA, NOAA), and in a commercial context (satellite operators). Such empirical research would be of significant value to multiple stakeholders in the space industry, especially in a period of heightened emphasis on responsiveness. Academia, with the support of various government agencies for example, can take the lead on such research and provide a thorough and unbiased analysis of selection processes for different programs and within several agencies; the end-objective of such research

would be to identify and share best practices that supports responsiveness across the space industry.

Recently, Bitten et al.[14], in discussing the effects of schedule constraints on planetary missions, touched upon the issues mentioned in the previous paragraph but only in the context of NASA's Discovery missions. They noted for example that

> *"the Discovery selection process is a lengthy undertaking, [and it takes] over 2 years from the development of a draft AO until the start of missions formulation."*

If other NASA or DOD space programs are Discovery-like in their selection process, the corresponding two years that follow the issuance of an RFP for an on-orbit capability and the time when a contractor is actually selected to start the work on said capability are clearly not responsive and unacceptable in an environment where responsiveness is a strategic imperative. It is therefore of primary importance that the structure and various steps within the selection process be thoroughly analyzed and its bottlenecks identified and streamlined in order to improve space responsiveness. We propose that, in a Federal acquisition context for example, the selection process of space assets should not only emphasize fairness and accountability, but also and explicitly, timeliness (not just efficiency). A reduction in time of the selection process from two years to say a few months would constitute a major improvement in space responsiveness, which can be obtained without having to pull on the other (probably more expensive) technical levers of responsiveness (hence the importance of the study we proposed in the previous paragraph). The selection process therefore falls in the upper right corner (high priority) of the responsiveness map introduced in Section 5, and illustrated in Figures 6 and 7.

Design reviews

Spacecraft are developed according to the traditional stage-gate development model. Multiple design reviews punctuate the development process and offer several advantages that justify their existence. For example:

i. They provide **visibility into the project's progress** (compared to initial plans and customer's needs and expectations) for multiple stakeholders, including the customer and senior management.
ii. They allow the (early) **identification of design errors and other problems** with the project; in so doing, design reviews allow the mitigation of development risk—the earlier problems are identified during the development process, the easier and cheaper they are to address.
iii. They allow for a multi-disciplinary **assessment of the level of quality and maturity** of the project before the project is allowed to proceed to the next phase and additional resources are committed to the project.

These reviews however take time, and in an environment where responsiveness is a necessity and a strategic imperative, it is worth carefully exploring other more expeditious and/or less frequent program management assessments and controls approaches. Consider the following: Zirger and Hartley[32] noted that

> *"in the traditional stage-gate model, valuable time can be lost scheduling and preparing for senior management updates. Moreover, during these updates, senior management may also assign additional and non-essential tasks to the project team."*

We heard anecdotal evidence from major satellite manufacturers about the impact of frequent design reviews and intense customer involvement in the development process of a spacecraft that stretches the development schedule by over 50%. On the other hand, a small highly responsive company that builds micro-satellites told us they engage their suppliers with a "handshake and trust," no progress reviews are conducted until the items are delivered. While this approach involves a significant degree of risk and may not scale well for larger companies and major satellite manufacturers, it does indicate nevertheless the possibility of responsiveness improvement through a careful balance and trade-off between risk and the frequency of design reviews on the one hand, and the development schedule on the other hand. The prerequisite to such a delicate balance is trust between customers and suppliers in the space industry, as well as trust between senior management and technical leads (or project team) within a company. As engineers, we would have never expected to come to the conclusion that "trust" is a lever of responsiveness! More research is necessary to carefully explore this delicate and important subject (of carefully balancing senior management and customer trust, support, and input into the project development while minimizing their intrusion on and disruption of the project's schedule).

Acquisition policies
In the following, we focus on acquisitions policies of space systems within the DOD. We first highlight the nature and magnitude of the problem (i.e., how/where space acquisition programs are failing); second, we discuss the relevance of acquisitions policies to space responsiveness; third, we point out what are the likely root causes of these problems and suggest a few possible solutions. As with the previous topics in Section 6, acquisition policies are here discussed inasmuch as they constitute a lever of space responsiveness.

All the government studies by the General Accountability Office (GAO) as well as the report of the Defense Science Board/Air Force Scientific Advisory Board Joint Task Force on Acquisition of National Security Space Programs (also known as the "Young Panel report"[33]) are consistent in their message that the DOD acquisition policies, despite recent reforms, are dramatically failing. Consider the following:

> *"In FY2008 alone, DOD expects to spend over $22 billion dollars to develop and procure satellites and other space systems. [...]. The majority of major acquisition programs in DOD's space portfolio have experienced problems during the past two decades that have driven up costs and schedules and increased technical risks. At times, cost growth has come close to or exceeded 100-percent (hundreds of millions to billion of dollars), causing the DOD to nearly double its investment without realizing a better return on investment. Along with the cost increases, many [space] programs are experiencing significant schedule delays, as much as 6 years, postponing delivery of promised capabilities to the warfighter [...]. In some cases, capabilities have not been delivered to the warfighter after decades of development. [...]. Outcomes have been so disappointing in some cases that DOD has had to go back to the drawing board to consider new ways to achieve the same or less capability."[34]*

In short, the DOD, the world's largest single customer of space systems, has a "poor record of developing satellites within cost and schedule targets and with promised performance."[35] How is this observation, and DOD's acquisition policies, relevant to space responsiveness?

Figure 13 shows the current delays or schedule slippage for a six DOD programs: the Advanced Extremely High Frequency (AEHF) satellite secure communications system, the Global Positioning System (GPS) BLOCK IIF upgrade satellites, The National Polar-orbiting

Operational Environmental Satellite System (NPOESS), The high-capacity Wideband Global SATCOM (WGS) satellites, and the Space Based Infrared System–High for military infrared surveillance. The reader is referred to the GAO-07-406SP[36] for details about these programs.

It is not surprising, given Figure 13 and the assessments by GAO and the "Young report" of DOD's poor track record of developing satellites within schedule that the emphasis on space responsiveness first emerged within the DOD as an effort to first contain or stop the systemic schedule slippage of space programs, and second to compress their development schedule—the former being probably more important to DOD in the near future. While we have focused so far in this paper on responsiveness improvement as schedule compression, it is important to recognize the dual nature of this problem, with the prevention of responsiveness deterioration or averting schedule slippage as the other side of the coin (i.e., one side, schedule compression; the other side, avoiding schedule slippage).
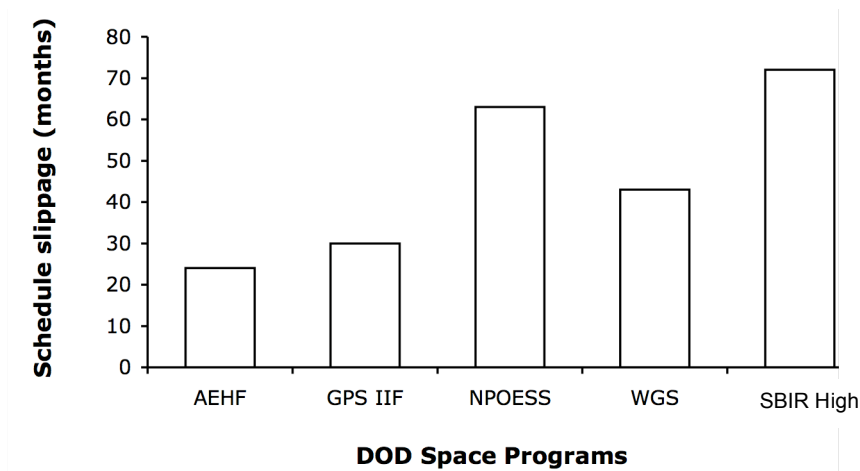

Figure 13. Delays and schedule slippage since program start (adapted from GAO-07-730T[34])

What are the roots causes of these problems highlighted by GAO and the Young report, and how can schedule slippage, or the worsening of responsiveness be averted? In the following, we focus on three particular causes underlying these problems, and recognize that solutions to this schedule slippage problem constitute a special kind of responsiveness levers:

i. The DOD acquisition policies of space systems, unlike most other weapon systems, allows for and supports the development of low TRL technologies within an acquisition program, instead of confining technology development to the "S&T [Science and Technology] environment, which is more forgiving and less costly than a delivery-oriented acquisition program environment."[37] When technologies not mature enough are included in spacecraft acquisition, "programs invariably run into problems that require more time and money to fix."[38] This is one of the major problems underlying the schedule slippage of the programs shown on Figure 13. All the studies by GAO of major weapon systems acquired by the DOD have identified the separation of technology development from system acquisition as a best practice. This has prompted GAO to recommend that the DOD not "allow technologies to enter into a weapon system's

development [and acquisition] until they are assessed [at least at] TRL 7, meaning that a prototype has been demonstrated in an operational environment."[35, 39] This recommendation supports our previous discussion of "Heritage, Learning Curve, and TRL" as levers of responsiveness.

ii. The second major cause underlying the problems illustrated in Figure 13 is the "undisciplined definition and uncontrollable growth in system requirements" occurring after the program has started ("Young report", 2003). In other words, requirements for space systems are not adequately defined at the start of the program, or they are changed significantly after the program started, thus resulting in the significant schedule delays that we observe in DOD's acquisition of space systems.[38] Decision-makers should therefore carefully consider balancing the conflicting urges of changing system requirements after a program has been started with the need for responsiveness.

iii. The third major cause underlying the problems illustrated in Figure 13 is a hybrid of an organizational design and a program management problem. Consider the following: the Young report noted, "as space-based support has become more critical to National Security, the number of users [of space assets] has grown significantly. As a result, requirements proliferate." This is echoed by a GAO finding that in some cases, "more than 30 organizations may have a hand in determining [and changing the] system's performance requirement before a contractor with systems engineering expertise can identify [potential] gaps between the requirements and available resources."[38] Having an array of officials and organizations involved in requirement definitions and system acquisitions may not in and of itself be problem for responsiveness if the program managers are sufficiently empowered and have the authority (and accountability) to make the necessary trade-offs between requirements, requirements growth, and maintaining the program on schedule (and within budget). This unfortunately is not the case, and the present situation is the result of two policies.

    a. First, the "adoption of the Total System Performance Responsibility (TSPR) policy in the 1990s […] lessened the government's program management role in favor of a stronger industry role, [and] essentially eroded the government's ability to effectively manage and oversee space programs […]. Over time, this shift […] resulted in decline in critical capabilities within the government workforce, particularly for systems engineering."[38] This observation is supported by the Young report, which noted that "policies and strategies inherent in acquisition reforms inordinately devalued the systems engineering and acquisition workforce […] and [TSPR] marginalized the government program management role." In short, the unfortunate combination of a growing number of officials and organizations involved in defining the requirements for space systems coupled with a lack of support and empowerment of DOD's program management to make the necessary choices and trade-offs between competing demands result in the problems observed in Figure 13 and the DOD's systemic inability to be responsive.

    b. Second, in addition to (and probably as a result of) the previously mentioned organizational and policy problems underlying DOD's lack of responsiveness, GAO and the "Young report" highlighted both the problems resulting from short tenures among staff critical to achieving acquisition success and responsiveness, and the lack of experienced program managers within the DOD to effectively execute space acquisition programs: "there are not enough experienced program managers to run space programs […] and program managers are not always

experienced enough to stand up to contractors when development is being mismanaged."[38]

In short, effective program management skills, and team empowerment—to make the necessary trade-offs during the requirement definition and system development phases in a timely and judicious manner—constitute as effective levers of responsiveness as the other (technical) levers discussed previously.

We are presently researching the possibility of developing a computational or Modeling and Simulation (M&S) approach to assessing the impact of acquisition reforms, in particular on space responsiveness.

VII. Conclusion

The first objective of this paper was to clarify the definition of responsiveness in the context of the space industry. For that purpose, we started by drawing an analogy from controls theory, and emphasized the temporal dimension of responsiveness as indicated by the total time τ_0 elapsed from the instance when the need for a given on-orbit capability is identified and formalized to the time when the asset is operational on-orbit and ready to fulfill the function it was intended for.

Next, drawing on the concepts of cost structure and Total cost of Ownership, we proposed that an analog breakdown be undertaken, not of cost but of the total time τ_0 to develop and launch a space system. The idea is to understand the various activities in the space industry (e.g., design, production, reviews, integration, testing, various bottlenecks and waiting times) following the issuance of an RFP for a new or modified on-orbit capability and how much time each activity contributes to the total time τ_0. This is akin to creating a new accounting system with the various activities as line items, except that instead of "cost accounting" we have a "time accounting" of each activity. In order to improve space responsiveness, we argued that it is critical to thoroughly understand, 1) what are all these activities a_i and how do they contribute, or not, to the overall development and readiness of the system, 2) what determines the duration of each activity τ_{ai} and the degree of overlap between these activities—the first and second point bring about the idea of streamlining the workflow and of "time compressibility" as will be discussed later—and 3) where are the bottlenecks, why do they exist and how can they be eliminated. These issues were discussed in sections 5 and 6.

We argued that conceptually, responsiveness is an attribute that is relevant for "customers" (or stakeholders with needs), and characterizes the reaction time of "suppliers" (or stakeholders addressing those needs, in whole or in part). Given this observation, the idea of multiple levels and different types of responsiveness becomes evident. To this effect, our review of the space industry value-chain lead us to define three levels of responsiveness:
1) a global industry-wide responsiveness, which is seen from the perspective of the end-customer. This is a "macro-level" attribute of the whole space industry. We referred to the end-customer in the space industry as the one who issues the RFP for a given space capability and "pays the bill" for the whole space asset
2) a local stakeholder responsiveness within the industry value-chain, as seen from the perspective of its local customer (as opposed to the "end-customer")
3) an interactive or inter-stakeholder responsiveness, which is dependent upon and determined by how well (or efficiently) the customer interacts and works with its

suppliers. We emphasize again that this interactive responsiveness is jointly controlled and determined by both the customer and the supplier.

Following our discussion of the three levels of responsiveness, we introduced and motivated the use of a new tool we termed "responsiveness maps" for multiple stakeholders in the space industry as a means for identifying focus areas and prioritizing efforts to credibly improve responsiveness. We argued that efforts to improve space responsiveness should not only focus on the architecture and technical characteristics of the artifacts created by said industry, but also include the legal, organizational, and managerial aspects of "doing business" in this industry; our last section in this paper identified multiple "levers of responsiveness," including technical spacecraft- and launch-centric, as well as "soft" levers (e.g., acquisition policies) for effectively improving the responsiveness of the space industry. Finally, we proposed a series of research questions to aggressively tackle problems associated with space responsiveness.

We believe the framework here developed will prove useful to engineers and decision-makers in thinking about credibly addressing issues of space responsiveness. These issues, as we showed in this paper, offer many exciting opportunities for theoretical and empirical research and contributions. In addition, we believe that academia can provide the necessary thought-partnership that can help the industry and government agencies tackle the problems of responsiveness (or lack of).

References

[1] Wertz, J. R., "Levels of Responsiveness," *AIAA 4th Responsive Space Conference*, Los Angeles, CA, 24-27 Apr. 2006. Technical note. URL: http://www.responsivespace.com/LevelsofResponsiveness.asp [cited 23 May 2007].

[2] Oderman, M. R., "Responsive Financing – the Ultimate Oxymoron?," *AIAA 1st Responsive Space Conference*, AIAA-RS1-2003-7002.

[3] Kaplan, S., "Bayes is for Eagles," *IEEE Transactions on Reliability*, Vol. 39, No. 2, 1990, pp. 130–131.

[4] Doggrell, L., "Operationally Responsive Space: A Vision for the Future of Military Space," *Air & Space Power Journal*, Vol. XX, No. 2, AFRP 10-1, Summer 2006, pp. 42-49.

[5] Shotwell, G., and Chinnery, A., "Space Exploration Technologies' Falcon I Launcher: Towards Operationally Responsive Spacelift," AIAA-2004-3905, *40th Joint Propulsion Conference*, Fort Lauderdale, FL, July 11-14, 2004.

[6] Webb, J. J. Jr., "On-Demand Science Missions," *AIAA 3rd Responsive Space Conference*, AIAA-RS3-2005-1003.

[7] Schmidt, M., "Soft Benefits in a Hard Business Case," Solution Matrix LTD, white paper, 2006. URL: http://www.solutionmatrix.com/total-cost-of-ownership.html [cited 23 May 2007].

[8] Braha, D., and Maimon, O., "The Measurement of a Design Structural and Functional Complexity," *IEEE Transactions on Systems, Man, and Cybernetics*, Part A. Vol. 28, No. 4, 1998, pp. 527–535.

[9] Bashir, H. A., and Thomson, V., "Models for Estimating Design Effort and Time," *Design Studies*, Vol. 22, No. 2, 2001, pp. 141–155.

[10] Lloyd, S., "Measures of Complexity, a non--exhaustive list," URL: http://web.mit.edu/esd.83/www/notebook/Complexity.PDF [cited 23 May 2007].

[11] Gell-Mann, M., "What is Complexity?," *Complexity*, Vol. 1, No. 1, 1995, pp. 16-19.

[12] Summers, J. D., and Shah, J. J., "Developing Measures of Complexity for Engineering Design," *Proceedings of DETC'03, ASME 2003 Design Engineering Technical Conference*, DETC2003/DTM–48633. Chicago, IL, September 2–6, 2003.

[13] Bearden, D. A., "A Complexity-based Risk Assessment of Low-cost Planetary Missions: When is a Mission Too Fast and Too Cheap?," *Acta Astronautica*, Vol. 52, No. 2, 2003, pp. 371–379.

[14] Bitten, R. E., Bearden, D. A., Lao, N., Y., and Park, T., H., "The Effect of Schedule Constraints on the Success of Planetary Missions," *Acta Astronautica,* Vol. 59, No. 8-11, 2006, pp. 1101–1109.

[15] Saleh, J. H., and Marais, K. "Highlights from the early (and pre-) history of reliability engineering," *Reliability Engineering and System Safety*, Vol. 91, Issue 2, 2006, p. 249–256.

[16] Wright, T. P., "Factors Affecting the Cost of Airplanes," *Journal of Aeronautical Sciences*, Vol. 3, 1936, pp. 49–73.

[17] Boothryd, G., Dewhurst, P., and Knight, W., *Product Design for Manufacture and Assembly*, 2nd ed., New York, Marcel Dekker Inc., 2002.

[18] Apgar, H., and Bearden, D. "Cost Modeling," *in:* Wertz, R., and Larson, W., *Space Mission Analysis and Design*, 3rd ed., Microcosm Press, Torrence, CA, 1999, pp. 809.

[19] Mankins, J. C., "Technology Readiness Levels," NASA Office of Space Access and Technology, White Paper, 1995. URL: www.hq.nasa.gov/office/codeq/trl/trl.pdf

[20] Gershenson, J. K., Prasad, G. J., and Zhang, Y., "Product Modularity: Definitions and Benefits," *Journal of Engineering Design*, Vol. 14, No. 3, September 2003, pp. 295–313.

[21] Ulrich, K., and Tung, K., "Fundamentals of Product Modularity," *Proceedings of the 1991 ASME* Design Engineering Technical Conferences - Conference on Design/Manufacture Integration. Miami, FL. September 22–25, 1991.

[22] Allen, K. R., and Carlson-Skalak, S., "Defining Product Architecture During Conceptual Design," *Proceedings of the 1998 ASME Design Engineering Technical Conference*. Atlanta, GA. September 13-16, 1998.

[23] Ulrich, K. T., Eppinger, S. D., *Product Design and Development*, 2nd ed., Irwin McGraw-Hill, Boston, MA, 2000.

[24] Enright, J., Jilla, C., and Miller, D., "Modularity and Spacecraft Cost," *Journal of Reducing Space Mission Cost*, Vol. 1, No. 2, 1998, pp. 133–158.

[25] Williams, G., Macklin, F., Sarigul-Klijn, M., Sarigul-Klijn, N., and Benson, J., "Almost There: Responsive Space," *AIAA 2nd Responsive Space Conference*, AIAA-RS2-2004-8000.

[26] Wertz, J. R., Conger, R., and Kulpa, J., "Responsive Launch with the Scorpius Family of Low-cost Expendable Launch Vehicles," *AIAA 1st Responsive Space Conference*, AIAA-RS1-2003-5001.

[27]Schoneman, S., Amorosi, L., Willey, R., and Laidley, M., "Space OSP-2 Minotaur Family of Space Launch Vehicles for Near Term, Low Risk Responsive Spacelift," *Space 2005*, Long Beach, CA, 30-1 Aug. 2005.

[28] Holweg, M., and Pil, F. K., "Successful Build-To-Order Strategies Start With the Customer," *MIT Sloan Management Review,* Vol. 43, No. 1, 2001, pp. 74-83.

[29] Wertz, J. R., "Responsive Launch Vehicle Cost Model," *AIAA 2nd Responsive Space Conference*, AIAA-RS2-2004-2004.

[30] Anderson, P., and Tushman, M. L., "Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change," *Administrative Science Quarterly*, Vol. 35, No. 4, 1990, pp. 604–633.

[31] Henderson, R. M., and Clark K. B., "Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms," *Administrative Science Quarterly*, Vol. 35, No. 1, 1990, pp. 9-30.

[32] Zirger, B. J., and Hartley, J. L. "The Effect of Acceleration Techniques on Product Development Time," *IEEE Transactions on Engineering Management*, Vol. 43, No. 2, 1996, pp. 143–152.

[33] Report of the Defense Science Board/Air Force Scientific Advisory Board Joint Task Force on Acquisition of National Security Space Programs. (also known as the "Young Panel report"), May 2003, (updated July 2004).

[34] Chaplain, C. T., "Space Acquisitions: Actions Needed to Expand and Sustain Use of Best Practices," GAO-07-730T, 19 Apr. 2007. URL: http://www.gao.gov/docsearch/repandtest.html [cited 23 May 2007].

[35] "Defense Acquisitions: Improvements Needed in Space Systems Acquisition Management Policy," GAO-03-1073, Sept. 2003. URL: http://www.gao.gov/docsearch/repandtest.html [cited 23 May 2007].

[36] "Defense Acquisitions: Assessments of Selected Weapon Programs." GAO-07-406SP, March 2007. URL: http://www.gao.gov/docsearch/repandtest.html [cited 23 May 2007].

[37] "Technology Development: New DOD Space Science and Technology Strategy Provides Basis for Optimizing Investments, but Future Versions Need to be More Robust," GAO-05-155, Jan. 2005. URL: http://www.gao.gov/docsearch/repandtest.html [cited 23 May 2007].

[38] "Defense Acquisitions: Incentives and Pressures that Drive Problems Affecting Satellite and Related Acquisitions," GAO-05-570R, Jun. 2005. URL: http://www.gao.gov/docsearch/repandtest.html [cited 23 May 2007].

[39] "Best Practices: Better Management of Technology Development Can Improve Weapon Systems Outcome," GAO/NSIAD-99-162, Jul. 1999. URL: http://www.gao.gov/docsearch/repandtest.html [cited 23 May 2007].