Innovation Pathways in Technology Intensive Government Organizations: Insights from NASA

by

Zoe Szajnfarber

B.A.Sc. Engineering Science, University of Toronto, 2006
S.M. Aeronautics and Astronautics, Massachusetts Institute of Technology, 2009
S.M. Technology and Policy, Massachusetts Institute of Technology, 2009

Submitted to the Engineering Systems Division
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in Engineering Systems: Technology, Management and Policy
at the Massachusetts Institute of Technology

June 2011

Author____
Zoe Szajnfarber
Engineering Systems Division
April 20th, 2011

Certified by____
Annalisa L. Weigel
Assistant Professor of Aeronautics and Astronautics and Engineering Systems
Thesis Supervisor

Certified by____
Edward F. Crawley,
Professor of Aeronautics and Astronautics and Engineering Systems
Thesis Committee Member

Certified by____
Daniel E. Hastings,
Professor of Aeronautics and Astronautics and Engineering Systems
Dean for Undergraduate Education
Thesis Committee Member

Certified by____
James M. Utterback,
Professor of Management and Innovation
Thesis Committee Member

Accepted by____
Nancy G. Leveson,
Professor of Aeronautics and Astronautics and Engineering Systems
Chair, ESD Education Committee

Innovation Pathways in Technology Intensive Government Organizations: Insights from NASA

by
Zoe Szajnfarber

Submitted to the Engineering Systems Division on April 20th 2011 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Engineering Systems

Abstract

Despite a rich legacy of impressive technological accomplishments (e.g., project Apollo, the Hubble Space Telescope) in recent years, the ability of government space agencies to deliver on their promises has increasingly been called into question. Although multiple acquisition systems and organizational structures have been tried, there remains a fundamental lack of understanding of how new technology development can, and should, be encouraged in this unique market structure and product context. This thesis seeks to address that gap by developing a more nuanced, and empirically grounded, explanation.

R&D management practices typically conceptualize complex product innovation as a Stage-Gate process whereby novel concepts are matured through a succession of development stages and progressively winnowed down at each sequential gate. This view implicitly assumes that maturity is a monotonically increasing function of the technology, and that partially matured technologies can be restored from the "shelf" for future maturation, barring obsolescence. However, based on evidence from six detailed process histories of instances of innovation in NASA's science directorate – including more than 100 hours of interviews, 150 archival documents and 2 months of informal observations – this thesis demonstrates how, in practice, the pathways taken by new capabilities do not respect these assumptions, with important implications. For example:

- Rather than being a monotonically increasing function in time, particular innovations draw simultaneously from funding mechanisms targeted at different Technology Readiness Level (TRL) ranges, and loop back to win "early stage" grants decades into their development when system level progress is stymied. As a result, increases to early stage R&D funding may not reach the targeted concepts.
- Rather than being a purposeful management decision, getting shelved is something that happens to innovation teams due to the lack of co-timing of technical breakthroughs and mission opportunities. As a result, maintaining a shelved capability is as much a matter of keeping the team together as it is a question of technical obsolescence.

The thesis argues that the observed dynamics can be better explained by four epochs of persistent, stable, and identifiable behavior, punctuated by transition inducing shocks. Acting individually, or in combination, these shocks can induce transitions from any one epoch to another. This new Epoch-Shock formulation enables a rethinking of the policy problem. Where the Stage-Gate model leads to an emphasis on centralized flow control, the Epoch-Shock model acknowledges the decentralized, probabilistic nature of key interactions and highlights which aspects may be influenced. These findings are considered both as NASA-specific recommendations, and more generally in terms of their implications for complex product innovation in a monopsony market.

Thesis Supervisor: Annalisa L. Weigel
Title: Assistant Professor of Aeronautics and Astronautics and of Engineering Systems

Acknowledgements

This thesis would not have been possible without the constant support and assistance of many people.

First and foremost, thank you to Prof. Annalisa Weigel, my advisor, for giving me the opportunity to work on such an interesting topic. I am grateful for the confidence you have shown in me, the enumerable ways you have stretched me intellectually and the subtle nudges whenever I veered off course. To the rest of my committee, Profs Crawley, Hastings and Utterback, thank you for pushing me to defend every statement, and in so doing, sharpening my work and preparing me for any audience. To my broader academic mentors: Thank you in particular to Drs Rhodes and Ross for providing me with a second intellectual home at SEAri; thank you as well to all the ESD and Aero Astro faculty and staff who have enriched my intellectual experience at MIT.

Second, thank you to my NASA hosts in the Goddard Office of the Chief Technologist, without whom this research would never have left the ground. Peter Hughes, thank you for seeing the potential and giving me the support required to let it flourish; Deborah Amato, thanks for helping make the work a reality and for your support, both formal and informal; and Deanna Trask, thank you for all the coordination and for keeping me on schedule. To all the scientists, technologists, engineering and managers (who must remain nameless), I cannot thank you enough for the hours you sat for interviews, digging up old documents and sharing your stories. You guys made my field work both fun and inspiring.

Third, to my various “labmates” over the last five years: Fellow CASPARites, past and present, thank you for your insights, feedback and camaraderie. Wind tunnel crew, grad school would have been a very different experience had I spent the first years in a different space; thank you for the advice that has continued to serve me well, well beyond your respective graduations. SEAri students, thanks for welcoming me into your community, listening to dry runs and commiserating when progress was slow. Special thanks to Rebecca, for the Thesis editing.

Fourth, thank you to my friends and ESD family for making this journey an enjoyable one. I could not have asked for a better community. Of all the people who have touched my MIT life, a few people in particular deserve individual thanks. Abe, Bruce, Erica and Hidda, I can’t express how important your support throughout my various injuries has been; I doubt I could have made it through without you.

Finally, I would like to thank my family, Mom, Dad, Raffi and Bubby. Thank you for your unconditional love and support; for instilling in me the joy of learning and fostering my curiosity.

The author would like to gratefully acknowledge the funding provided for this work by the Natural Science and Engineering Research Council (NSERC) of Canada and the National Aeronautics and Space Administration (NASA).

Table of Contents

Table of Figures

Table of Tables

1 Problem Formulation

The requirement for innovation is fundamental to NASA's mission: "*to pioneer the future in space exploration, scientific discovery and aeronautics research*." Doing ground-breaking space-based science requires the continuous invention and engineering, of new, better, and more precise instruments. Yet, in recent years, the sector has been heavily criticized for its performance in this respect (Augustine, et al., 2009; Lawler, 2009). In the context of NASA's activities, innovation is defined as the process by which ideas for new, cutting-edge technologies are conceived, developed and integrated into advanced space missions.[1] While NASA has a formal systems engineering process for new project development (NASA, 2007), as governed by the Federal Acquisition Regulations (Government, 2005), to the extent that it has an explicit R&D management process, it is highly decentralized. At NASA, innovation is nominally conceptualized as a three-stage process. Promising ideas are initially explored through basic concept development and funded by individual research centers. Next, the most promising concepts are further matured through applied R&D, which is funded from Headquarters (HQ). Finally, a very small subset of the maturing concepts are infused into flight projects and undergo formal project-specific development. The flow of concepts from conception to implementation on flight projects is controlled by a series of gates – decision points - where progress is reviewed and the set of maturing capabilities that will progress to the next level are selected.

The goal is to develop enough new capabilities now, so that future projects will be able to draw upon adequately mature technologies later. In theory, no new technology can be infused into a flight project unless it has reached a Technology Readiness Level (TRL) of at least "TRL 6"[2] (NASA, 2007). However, there is circular logic fundamental to this system. Projects rely on novel capabilities to accomplish their ambitious mission objectives, but prefer *proven* technology. A new technology cannot be considered *proven* until it has flown; in some cases, applied R&D funding cannot even be secured until interest from a flight project has been demonstrated. This tension is illustrated vividly by the contrasting perspectives of an experienced NASA Center chief engineer and a staff instrument scientist:

> Chief engineer: "*there is not a dearth of ideas; [that's not the problem, the problem is that] there is a sad lacking in the understanding of the ramifications of carrying the idea through to its conclusion. So the ability of a human being to sort through 100s of ideas to find the one or two that might be a useful nugget is a very difficult. […] <u>in general, there are more technologists with ideas looking for a place to apply them than there are people who are flying flight missions looking for ways to solve problems that they have with new technology.</u>*"
>
> Staff scientist: "*Technology takes years to develop - from when you have a good idea to when you have an applicable product even to a single government use, never mind*

[1] In the broader literature, multiple definitions of innovation exist. Most include the concepts of novelty and commercialization. In the space and defense context, where commercialization is not typically the goal, implementation (i.e., use in the field) has been taken as equivalent (Grissom, 2006).
[2] NASA's Technology Readiness Level (TRL) scale is described in detail in Ch2. A TRL 6 technology is one that has been demonstrated in a flight-like environment.

commercial. So, to have a coherent plan, what you need is a vision for the technology needs that is stable compared to that timeline. It's not. We don't know what we're doing for years and years at a time, and by the time we do, the technology that's in the pipeline is misdirected. That doesn't always happen, [but] that happens enough that it seriously detracts from the utility of the program. <u>*And people at my level are essentially reading tea leaves and putting fingers to the wind trying to figure out where the wind's shifting to try and leverage the opportunity towards something useful. And sometimes it works… surprisingly! but a lot of times, [you find that] you built a widget that has no applicability*</u>*."*

In practice, the system works as well as it does because these contradictions are resolved through informal mechanisms. Innovations are furthered by dedicated technologists, funded by programs that do not officially exist, and infused into flight projects via informal social interactions. The current conceptual models of NASA's innovation process, upon which many technology management decisions at NASA are currently based, do not capture the dynamics of the informal structure that has evolved. In fact, in discussing this research, one member of NASA's leadership noted: "*even just drawing a picture of what the system actually looks like would be an important contribution."*

This thesis draws that picture and, in so doing, provides a basis for future strategic decisions. It develops a more complete and nuanced understanding of the formal and informal mechanisms that drive pre-project technology development at NASA. Specifically, it addresses the following three research questions:

1) What is the structure of NASA's innovation system?

To answer this question, Chapter 2 maps out the components of NASA's innovation system. This map includes both the institutional and human elements of the structure, and the connections – explicit and organic – that link them. It considers not only NASA's internal structure, but also the network of small businesses, contractors and academics that play critical roles in NASA's efforts.

2) How do new capabilities traverse the innovation system as they are matured and infused into flight projects?

To answer this question, Chapter 3 traces the pathways taken by six NASA innovation from conception through implementation on a flight mission. It considers the technological trajectory as an anchor for the story, but also uses the expansion and contraction of team personnel, the funding mechanisms employed at different times, and the contextual changes to understand how the path is shaped. Chapter 4 then abstracts the observed patterns and develops a process model of the dynamics of innovation at NASA. It illustrates how, despite their idiosyncrasies, all six pathways can be explained as a small number of persistent, stable, and identifiable system states, punctuated by transition inducing shocks.

3) To what extent can the observed *innovation pathways* be improved through feasible management interventions?

To answer this question, Chapter 5 explores the mechanisms that generate the shocks described in Chapter 4. Specifically, it explores the linkages between controllable elements of the system structure (Chapter 2) and the dynamics captured in the process model. It then discusses the policy implications of feasible interventions.

The rest of this chapter delves one level deeper into each aspect of the problem formulation. Section 1.1 describes an illustrative innovation pathway in detail as a basis for concretizing the phenomenon of interest in this research. Building on the implicitly defined concepts, Section 1.2 provides explicit definitions and illustrates the applicability and limitations of extant theory in explaining innovation at NASA. Finally, Section 1.3 describes the research design used to address the above questions.

1.1 The Phenomenon: An Illustrative Innovation Pathway

What does an innovation pathway look like? To give a sense of the phenomenon we are trying to understand it is useful to present one illustrative innovation pathway in a fair amount of detail. This pilot case describes the development of quantum well infrared detectors (QWIPs). This section describes the innovation pathway of the QWIPS using section headings that anticipate the structure of innovation concepts employed in Chapter 2.

Following more than a decade of development, the QWIPs will fly aboard the Landsat Data Continuity Mission (LDCM) as part of the Thermal Infrared Sensor (TIRS). Over their history, Landsat data have enabled agricultural, forestry, air quality, and geological activity monitoring efforts, among other societal benefits. The new QWIPs technology will continue this legacy in the thermal infrared band, adding improved sensitivity. The planned 2012 launch will mark the first implementation of a QWIPs-based sensor on a space-based platform, and one of the first applications of QWIPs principles in an infrared camera system in the relevant wavelength (8-12 μm) [D12, I1]. The QWIP innovation has potential applications to a wide range of space and terrestrial missions [D11].

Gestation Period (setting the initial conditions)

The QWIPs innovation pathway began in the late 1980s, when the potential for quantum wells to be used as far IR photo detectors was demonstrated through a unique collaboration between scientists at AT&T/Bell Labs and the NASA Goddard Space Flight Center (GSFC) [D16, I7]. During their year-long contract, funded on the order of $100K under the strategic defense initiative (SDI) [I7], they incorporated a QWIPs-based photodetector array into a camera system and used it to perform airborne imaging [D16]. However, despite the early promise showed by the new technology, the project essentially terminated with the contract, and the collaborating organizations went their separate ways. For Bell Labs, the termination resulted from a strategic determination that QWIPs detectors were not aligned with their commercial portfolio. For the GSFC technical team, their other flight project responsibilities won out, leaving the QWIPs detector arrays in the proverbial “sandbox.” [D16, I7, I3]

Although Bell Labs as an organization ended the project, many of the young scientists involved maintained interest in the nascent technology [D16]. Recognizing the enhancements that QWIPs could offer to space-based imaging, the Jet Propulsion Laboratory (JPL), another NASA center, acquired both the technology and many of the original scientists in 1992 [D16, I3]. At JPL, they formed the Infrared Focal Planes & Photonics Technology Group, where they continue to push the scientific state-of-the-art in both space-based and earth IR imaging technology.[3]

Project Initiation (reuniting the old team)

In the late 1990s, the original collaborators were reunited for another reason. Believing that NASA as a whole would benefit from more collaboration among its centers, the GSFC and JPL groups were "encouraged" to find a basis for collaboration [I1, I3, I7]. Encouragement in this context meant that scarce R&D resources were earmarked for collaborative projects. So, the GSFC team spent a week on-site at JPL discussing potential common projects. It turns out that QWIPs detectors were the most promising area for collaboration between the groups, and the outcome of that week was one of the concepts that is now, 10 years later, being developed for the TIRS instrument [I7].

Maturing the Concept

The team progressed under a sequence of back-to-back Earth Science Technology Office (ESTO) grants, supplemented by a short Director's Discretionary Fund (DDF) contract on the GSFC side [I1, I4, I7, D1-2]. The first tranche of funding included $700K over 3 years from 1999-2002 from ESTO's Advanced Technology Innovation Program (ATIP) and $80K for the first year of DDF [I7, D6, 9]. This allowed them to develop a hyperspectral QWIPs sensor array [D 6], useful for remote sounding of numerous geospatial quantities. This work showed sufficient promise to secure another three years of funding, again from ESTO, now under the Advanced Component Technology (ACT) bucket. In this round, they requested $1.2M over 3 years through 2005, and built a 1Kx1K detector array and the corresponding read-out circuit [I7, D2, 7]. As listed in the project report, the first contract matured the capability from TRL (technology readiness level) 2 to 5 [D6] and the second contract from TRL 2 to 6 [D7], illustrating the flexibilities of the definitions of TRL. These contracts enabled the technologies to be matured to the point that there was no new science left to be worked out; the remaining investment would target space qualification and engineering progress, tasks that fell under the purview of engineers and technologists.

Treading Water and Branching Out

However, this type of targeted technology funding is hard to come by. There is no clear path between ESTO development funding and project applications, and, for QWIPs, there was no immediate flight opportunity available within the space context. Nonetheless, after 7 years of significant progress, the technologists were motivated to find practical applications to further their work, including medical imaging and cave mapping, which served as partial bridge funding [I7].

A Chance Encounter

A new chance for space flight soon arose, when the Goddard technologist and the CEO of a small business began chatting at a domain specific technical conference [I1, 3, 7]. It turned out that the company had been doing some groundwork on manufacturing a QWIPs-based camera, but had been struggling to secure funding. In fact, they had already submitted three blind

---

[3] See http://scienceandtechnology.jpl.nasa.gov/people/s_gunapala/ for more details

proposals to NASA's SBIR program (a congressionally mandated innovation funding mechanism) [I3], but despite success with the Army version of SBIR, all three NASA proposals had been rejected; the CEO was about ready to give up on NASA [I3].

Developing a Parallel Technology Branch

By the end of this informal chat, the two had come up with a proposal strategy with a well defined concept [I1]. Another *coincidence* further ensured the success of the company's fourth SBIR proposal. The relevant subtopic manager – the individual in charge of soliciting reviewers and technical contracting officers for the SBIRs – had an office down the hall from the Goddard technologist [I1, 7]. Thus, upon returning to Goddard, the technologist indicated to his friend, the subtopic manager, that there might be a QWIPs proposal coming in, and that he would be happy to review it. The subtopic manager agreed "*because it's really hard to find reviewers. So if someone volunteers it's very hard to say no…*" [I1] and assigned himself as the second reviewer [I7]. Both technologists were suitably impressed; the contract was awarded in September of 2006 and the GSFC technologist became the contracting officer's technical representative (COTR) [D3].

Treading Water Again

The output of the phase I contract was a prototype QWIPs-based camera, an excellent result for the 6 month 100K contracting mechanism [I 1, 3]. Despite the success, the phase II bid was rejected for reasons that seemed baffling to the team [I1]. Outraged by what he saw as a clear failing of the system, the Goddard technologist made a series of phone calls to his colleagues in programmatic roles [I1]. As told from his perspective, within a few weeks, the funding managers realized their mistake and righted the wrong by securing enough funding (~$300K) to "*keep us both [him and the small business] alive for another 18 months, which turned out to be enough.*" [I1, 7] The funding came from a partial SBIR phase II and a partial ESTO grant – redistributed at the discretion of the program office [I 4]. In this case, the role of technology portfolio management for the SBIR phase II awards, and ESTO program manager were held by the same individual, a colleague of the Goddard technologist [I1, 4].

As recounted by the ESTO fund manager, the initial phase II rejection should not have been surprising at all [I5]. The way the transition from phase I to phase II SBIR awards is structured, center-level boards rank their own center's finishing phase Is, and that ranking forms the basis for NASA-wide SBIR portfolio planning for phase II [I5, 11]. In the case of QWIPs, the Goddard ranking was quite low, so it would have been inappropriate, from a process point of view, for the ESTO fund manager to recommend that it receive follow-on funding. In his view, the low ranking was because of a lack of advocacy in the review meeting by the Goddard technologist/project COTR [I5]. While advocacy is not explicitly necessary, a short presentation by the COTR is the primary basis for the committee's decision and, all else being equal, enthusiasm about the recommendation plays an important role.

The Goddard technologist had not appreciated the importance of his advocacy role; he believed that the capability's obvious importance as an enabler of future missions should speak for itself [I7]. This was not the first funding proposal on this innovation pathway that had been turned down for lack of advocacy. A previous Internal Research and Development (IRAD) proposal had been rejected because the link to future missions had not been effectively communicated [I8];

similarly, the earlier SBIR phase Is had not shown clear flight project ties. However, at this time, the Goddard technologist felt strongly enough about the efforts of the company to ‘play the game’ [I1, 7]. Retroactively, he was able to convince the funding manager and secure the follow-on funding. At the same time, he supplemented the SBIR/ESTO combination with Goddard’s internal R&D funding (IRAD) [I1, 4, 7, 8].

Changing Contexts

However, even before the IRAD could be completed, the original SBIR team was drawn into a project-specific development contract to develop a QWIPs-based TIRS (Thermal Infrared Sensor) instrument for the Landsat Data Continuity Mission (LDCM) [I1, 7]. The politically charged LDCM mission was facing major technical difficulties with its baselined TIRS instrument [D12 -15]. To understand the context of this decision requires a brief history of the Landsat project, and LDCM in particular.

The Landsat Data Continuity Mission (LDCM) is a joint NASA-US Geological Survey mission that will continue a 30+ year legacy of geospatial data [D11]. Over its history, data from Landsat(from a seven satellite sequence of missions) have enabled agricultural, forestry, air quality, and geological activity monitoring, among other societal benefits. In 1992, the Land Remote Sensing Act guaranteed a data continuity mission to follow Landsat 7. This Act would include continuity of the thermal band imaging provided by Landsat 4, 5 and 7. However, since these measurements are difficult (i.e., expensive), and interesting only to a small community of specialists, their continuity has never been popular in Washington.

From the perspective of TIRS, LDCM has undergone several major reformulations. In the mid 90s, LDCM was investigated as a series of industry studies, with TIRS as an optional extra; they came up with huge cost estimates based on heritage microbolometers. None of the studies were selected and, following a National Security Council-led interagency review, the LDCM functionality was relegated as one of the many NPOESS instruments.[4] That program ran into difficulties. In 2005, OSTP released a memo directing NASA to reinstate LDCM as a free-flying mission. Thus, NASA released a request for proposals, which included the possibility of a TIRS instrument. By that time, scientists in Idaho had found a way to use Landsat thermal data to resolve water resource disputes, an important and expensive issue in the U.S. Midwest, and had established a powerful lobby in support of thermal imaging.

In 2007, NASA conducted an in-house concept study of a TIRS instrument. Believing that coming up with a reasonable cost estimate was the most important factor for TIRS inclusion, they turned to commercially available microbolometers. However a closer look revealed that microbolometers were not an adequate technical solution and the TIRS instrument was de-manifested to preserve schedule. Then, in 2008 at the Systems Requirements Review (SRR), the project was projected to be at least 6 months behind schedule. This created a new opportunity for a TIRS instrument to be manifested. HQ asked Goddard, the systems integrator, if a new instrument could be furnished. As recalls the LDCM project scientist, Goddard responded with

[4] The National Polar-Orbiting Operational Environmental Satellite System has experienced its share of starts and stops (c.f. http://www.spaceref.com/news/viewpr.html?pid=25719)

an ultimatum: "*yes, but only if we can use QWIPs and develop it in-house.*" HQ conceded and QWIPs was baselined. [I10, D13]

The decision for an operational satellite program to infuse an unproven (read: risky) new technology is never taken lightly. Although the above quote emphasizes the Goddard's power to "strong-arm" HQ, this power is highly context specific and in this case was derived from the confluence of two sequences of events, joined by a timely problem. The first event was the application of significant and unprecedented political pressure: In addition to the 30,000+ publications that have resulted from Landsat science data, the turbulent launch history of the previous 7 Landsat missions brought together a cohesive Landsat lobby. It was the lobby that ensured that an LDCM would fly in the first place, and the decision to de-manifest TIRS sparked another round of heated debate [D14, D15]. In the end, the FY2009 Appropriations Act explicitly included $10M for a TIRS, officially legitimizing the risk [D16]. The FY2010 Appropriations Act provided another $150M for the TIRS instrument, to ensure the schedule was met. The second event was fortuitous technology readiness: Whereas in the original TIRS technology trade studies in 2000 and 2007, QWIPs were not mature enough to even be considered [D12, I7], a year later, following significant investment through ESTO, SBIR and IRAD, the technology was now considered the least programmatically risky choice. Further, while the incumbent had made limited progress during the intervening 8 years, Goddard had developed the in-house capability of manufacturing QWIPs devices, thereby eliminating the need for time consuming procurement and making the extremely aggressive two year timeline realistic.

Parallel Development as Part of a Project

Currently in the TIRS baseline, the original QWIPs team and additional engineers have been pulled into the LDCM project and are operating under a stable, legislatively guaranteed ~$10M, project-specific funding with a clear, near-term mission objective. The Goddard technologist believes that an important difference between the ESTO product, which did not convince the TIRS team, and the SBIR output, which did, was the involvement of a commercial company [I1]. From the project perspective, though, the difference was as much a matter of evolving priorities as technical maturity. By 2008, schedule was the primary consideration and the fact that QWIPs could be manufactured in-house was a significant time-saving argument.

The schedule pressure also drove the decision to pursue two parallel development paths simultaneously; the idea being that at least one was likely to succeed. The first was a collaboration between Goddard and the US Army Research Laboratory (ARL) which sought to design and fabricate a corrugated QWIP array based on the concepts proven in the ESTO contracts. The second was a team consisting of the SBIR small business, with support from GSFC/ARL, to develop a grating based QWIP array [D11]. In the end, both technical approaches met all of the TIRS requirements. The SBIR approach was eventually selected based on a second order uniformity measure. The Goddard group sees the infusion as a win nonetheless, because now that QWIP technology has been flown once, it is now a "mature" technology that will be much easier to justify for use on future missions [I12]. At the time of this writing (November 2010) the TIRS QWIP-based flight focal plane assembly had been built, space-qualified, fully tested and was waiting to be installed on the TIRS instrument. The LDCM mission is scheduled for launch in 2012 [D11].

1.2 Some Partial Explanations

The QWIPs innovation pathway provides a rich illustration of the complexity of the process. The QWIPs trajectory was heavily influenced by bureaucratic decision processes, but structured by a formal systems engineering-inspired institutional architecture. Technological contributions came from a diverse, yet specialized, group of actors – including small technology firms, government agencies and research institutions – linked by common interests and a range of short term, one-off, contracting and granting mechanisms. Decisions were framed by bureaucratic conservatism on the one hand, and extremely ambitious technological expectations on the other.

The challenge of reconciling institutional resistance to change and general risk aversion, with the need and desire to innovate is not unique to NASA, and has been addressed previously by multiple disciplines. Three explanatory lenses that are particularly relevant to innovation at NASA (a bureaucratic project-oriented developer of complex engineering systems) are: 1) Stage-gate processes in project-based firms; 2) Garbage cans and windows of opportunity in bureaucratic organizations; and 3) Strategic trade-offs associated with balancing both exploration and exploitation in the context of a firm. Each view offers a partial framework for understanding a different piece of innovation at NASA, but each has limitations as well. In addition, the lower level dynamics we observe in the NASA context can be explained, in part, by organizational literature focusing at the individual level. These explanations will be used to interpret and position the findings given later in the thesis.

1.2.1 Stage-gates

The stage-gate perspective models the innovation process as a series of stages (during which technology is matured) separated by gates (decision points, where progress is reviewed and the set of maturing capabilities that will go on to the next level are selected, while the rest are shelved) (Cooper, 1990; NASA, 2007). By conceptualizing it in this way, the innovation management problem reduces to a design problem: Given a fixed set of resources, choose (a) the number of stages, (b) the relative resources allocated to each stage, and (3) the gate decision rules, such that the desired flow of new capabilities is achieved. Despite the considerable evidence that the innovation process is neither segment-able nor sequential (Rothwell & Zegveld, 1994), the stage-gate view persists as the *modus operandi* at NASA and other space agencies (e.g., DoD 5000; (Guglielmi, Williams, Groepper, & Lascar, 2008)). This is partially a bi-product of NASA's systems engineering heritage, and partially because a stage-gate view is eminently tractable from the perspective of a technically trained NASA manager. As a result, historically, NASA's efforts to restructure its technology development process have centered on shifts between more or less emphasis on basic vs. applied R&D and the extent to which the two R&D buckets are explicitly linked both to each other and to flight projects (see Figure 1-1). Recently, discussions have focused on redefining the buckets and adding additional buckets to smooth transitions among them (Braun, 2010).

However, the QWIPs innovation pathway described above appears to violate some of the fundamental assumptions of the stage-gate representation. Figure 1-2 overlays the QWIPs progression on a representation of NASA's current technology development stage-gate process. It shows that, while the innovators leveraged each stage of technology support, the progression did not follow the expected sequence. As will be explained more fully in the body of the thesis, scientists and technologists frequently secure funding from multiple buckets simultaneously and are liable to pursue interesting technology tangents which can result in a reset of the perceived maturity of the system.

This thesis will argue that because of these and other *realities* of the system, the stage-gate conceptualization is not just a coarse representation; it is wrong, and its logic suggests interventions that will not have the expected effect. For example, if technologists are drawing from multiple buckets at the same time, regardless of the concept's maturity, then altering the relative resource distribution is a futile pursuit. Similarly, the eventual infusion of QWIPs technology on LDCM was as much a function of external political events as it was determined by an explicit gate decision. A stage-gate conceptualization cannot account for this contextual unpredictability.

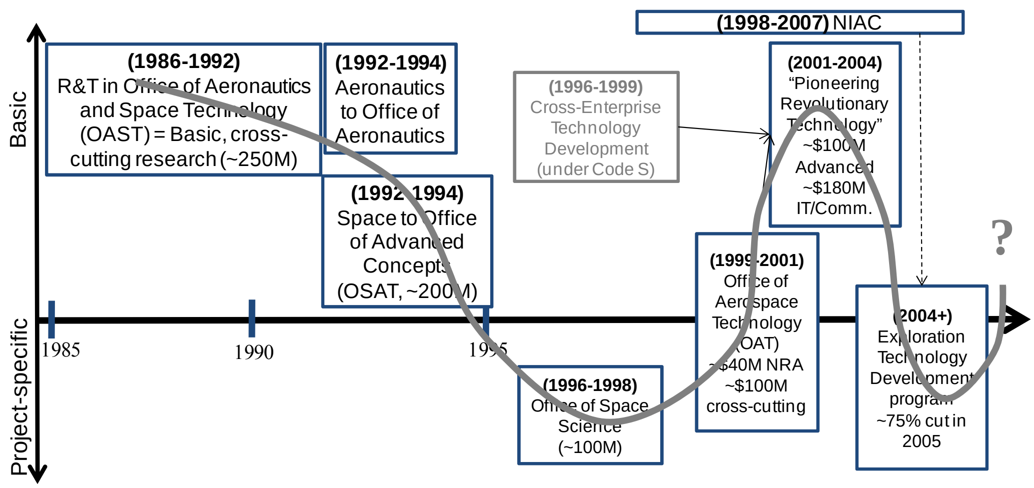

Figure 1-1: History of NASA's Technology Strategy

1.2.2 Windows of Opportunity and Garbage Cans

The 'windows of opportunity' perspective does not try to compartmentalize the process, but it does oversimplify and de-emphasize the importance of structure in the pre-window development for technology-intensive solutions. Popular in policy studies, it argues that separate problem streams and solution streams exist independently. Progress occurs when a window of opportunity opens, allowing a problem and a solution stream to combine and yield a new status quo (Kingdon, 1984; Stone, 2002). In this view, progress should happen incrementally except when an opportunity for a step-change is seized. This notion is similar to the punctuated equilibrium model of organizational change (Gersick, 1991; Tushman & Romaneli, 1985). Conceptualized this way, improving the system becomes a matter of anticipating and using windows effectively, "*like surfers waiting for the big wave*" (p. 165, Kingdon 1984).

The QWIPs infusion on LDCM can be told as a classic tale of a window causing the intersection of a problem and solution stream. Problem: the originally baselined TIRS instrument ran into serious technical challenges. At the SRR (two years before the planned launch), it became clear that the baselined technology would not be flight ready in time. As a "data continuity mission," launch timing was an important mission driver. Solution: in the intervening time since the original selection of the TIRS instrument technology, significant scientific progress had been made on QWIPs devices. Manufactured from mature semiconductor materials (and the corresponding mature processing technology) and boasting a working prototype, QWIP-based TIRS became the lowest-schedule-risk alternative. Window: in addition to the 30,000+ publications that have resulted from Landsat science data, the turbulent launch history of the previous 7 Landsat missions had brought together a cohesive Landsat lobby. It was the lobby that ensured that an LDCM would fly in the first place, and the decision to de-manifest TIRS sparked another round of heated debate. In the end, the FY2009 Appropriations Act explicitly included \$10M for a TIRS, officially legitimizing the risk. This created a new opportunity for a TIRS instrument to be manifested. When HQ asked Goddard, the systems integrator, if a new instrument could be furnished, Goddard responded with an ultimatum: "*yes, but only if we can use QWIPs and develop it in-house*." HQ conceded and QWIPs was baselined.

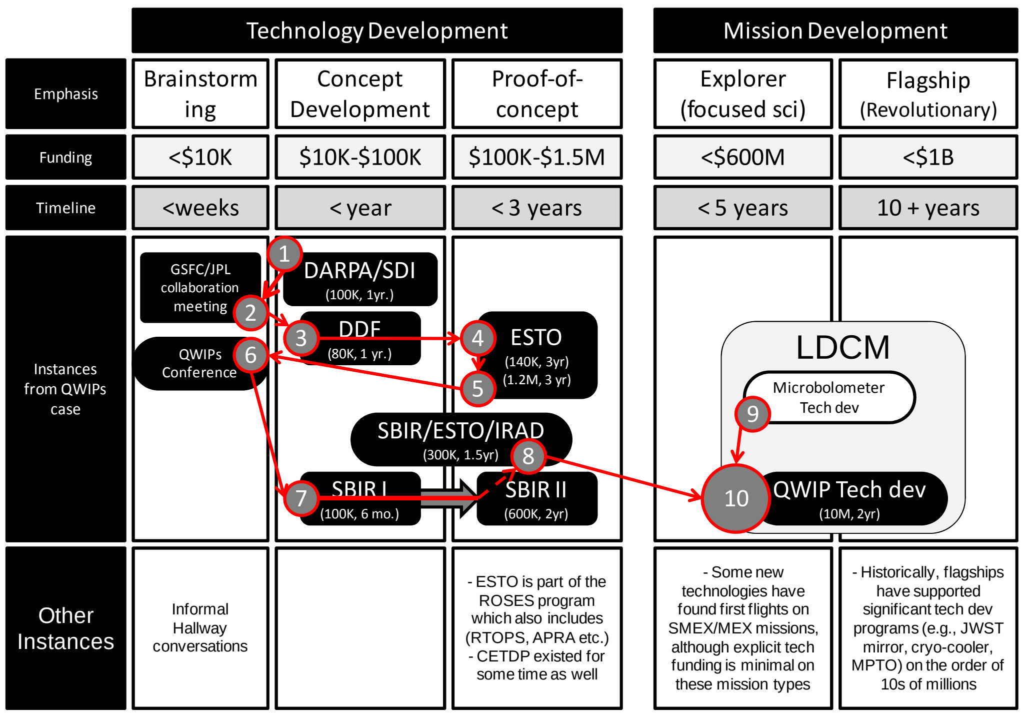

Figure 1-2: Switchbacks in the Innovation Process: QWIPs Illustration

In this case, the opening of the window did in fact allow the problem and solution to find each other. Where the bureaucratic decision making and agenda setting explanations are lacking is in the technology solution side of the story. While it may be impractical, and of little value, as Kingdon argues, to attempt to trace a causal chain of events in a socio-political "primordial soup," there is significant path dependency, and a series of necessary investments, in a technology development process. Further, while the existence of a continuous flow of new concepts and capabilities, generated by independent actors on the supply-side, can be appropriately assumed by decision makers in a competitive market context, these dynamics are fundamentally different in the monopsony markets characteristic of the space sector (Adams & Adams, 1972; Szajnfarber, Richards, & Weigel, in press). Specifically, since many of the new technologies which are critical enablers of future space science missions have limited or no near-term commercially viable applications on earth, the necessary R&D investment will be underprovided by industry without the combination of government patronage and explicitly articulated future needs (Peck & Scherer, 1962; Sherwin & Isenson, 1967).

The concept of a policy window in technology intensive systems has been examined to some extent. Many past studies of innovation in government agencies have been conducted by scholars of military and strategic studies. Their insights have focused on identifying the antecedents of windows (Logsdon, 1970; McDougall, 1985; Posen, 1984) and/or the implications of their opening (Rosen, 1994; Sapolsky, 1972). For example, in his seminal book, Sapolsky (1972) argued that innovation is catalyzed by inter-service rivalry. Based on an analysis of the Polaris missile system development, Sapolsky illustrates how the desire to secure providence over contested mission types (*i.e.,* Navy Polaris versus the Air Force Minuteman missile system), gave the mission sufficient priority to attract top talent and overcome institutional inertia. In a similar vein, Rosen identifies intra-service rivalry and the presence of internal sub-cultures as key catalysts of change (Rosen, 1994). Conversely, Posen argues that the catalyst must be external to the military organization, pointing to the influence of civilian leadership (Posen, 1984). The common theme is that bureaucratic organizations are designed to resist change (Rosen, 1994) and require a shock from outside if major change is to occur (Posen, 1984). This type of intervention has typically been motivated by fear (e.g., the ramp up of missile development following WWII) (Beard, 1976), prestige as an instrument of foreign policy (e.g., the moon race, the International Space Station (ISS) (Launius & McCurdy, 1997; McDougall, 1985)), out of necessity during times of military engagement (e.g., the Manhattan Project), or sometimes idle interest (e.g., Morison's gunfire at sea (Morison, 1966)).

These studies provide a rich understanding of how certain contextual factors can influence the project development process and were helpful in designing this study. However, they largely ignore the component and subsystem innovations that enable system-level radical performance improvements (Sherwin & Isenson, 1967) and have limited prescriptive utility, since sustained innovation cannot wait for the next Kennedy moon speech. An important contribution of this study is thus its extension of Kingdon's model to include the pre-development of technology-intensive solutions.

1.2.3 Balancing Exploration and Exploitation

It is well recognized in the management literature that sustained performance requires both exploration (seeking radical innovation through the pursuit and acquisition of new knowledge) and exploitation (leveraging existing capabilities to enable incremental improvements, and derive value from past exploration) (Greve, 2007; March, 1991). However, studies suggest that characteristics of a firm which enable exploration tend to limit exploitation, and vice versa, (O'Reilly & Tushman, 2007) since they are mutually contradictory and self-reinforcing pursuits (March, 1991).

Two strategies for combining exploration and exploitation have been proposed in the literature. So-called ambidexterity (Tushman & Smith, 2002) advocates for combining exploration and exploitation through loosely coupled organizational sub-units integrated by top management (Smith & Tushman, 2005). Punctuated equilibrium, on the other hand, suggests that the contradictory functions of exploration and exploitation can be balanced through temporal sequencing (e.g., long periods of exploitation, followed by short bursts of exploration) (Brown &

Eisenhardt, 1998; Burgelman, 2002; Tushman & Romaneli, 1985). At the project level, punctuated equilibrium can also be conceptualized as cycles of convergence and divergence (Van de Ven, Polley, Garud, & Venkataraman, 1999). Despite extensive study, there remains limited consensus on how the competing forces should best be balanced.

NASA presents an interesting empirical setting in which to explore these ideas since the agency is ambidextrous in a decentralized, albeit structural, way; but it also undergoes periods of greater exploration and greater exploitation. Flight projects exploit mature technology (incorporating components and subsystems that are at least at TRL 6), while R&D grants support various levels of research and development, usually at the component or subsystem level. The funding mechanisms are administratively distinct (sometimes made explicitly so in congressional budgets) and managed with opposing philosophies. Whereas project managers are responsible for the cost and schedule performance of the projects they run, ROSES[5] program element managers facilitate a competitive peer-review selection process and have essentially no influence once the funding has been distributed. This combination of activities resembles Tushman et al's conceptualization of structural ambidexterity; however, the link between the exploratory and exploitative units occurs at the working level. Specifically, individual scientists and engineers work on some combination of both flight projects and R&D at any given time. The ratio is determined organically by their propensity to secure grant funding, and varies throughout their career.

At the theme-line level (e.g., astrophysics is a science "theme"), a more classic form of structural ambidexterity is maintained by funding a combination of Flagship and Explorer class missions in any given decade (see Figure 1-2). At the agency level NASA has flipped back and forth between periods of greater focus on exploration or exploitation. Figure 1-1 plots this history since 1985. Compared to commercial firms, NASA has always invested in relatively more exploration, especially if Flagship projects are considered exploratory, but the historical differences in magnitude and orientation of the investment are substantial. Qualitatively, these flips in emphasis on the relative proportion of exploration vs. exploitation tend to correspond to changes in administration – both presidential and consequently the NASA administrator – or major environmental changes. For example, the Columbia Accident (Gehman, et al., 2003) precipitated President Bush's Vision for Space Exploration (NASA, 2004) and the reallocation of 75% of all R&D to Constellation-oriented work in 2004. This is consistent with, albeit at a different level than, the findings of (Romaneli & Tushman, 1994).

The language of exploration and exploitation represent a useful framing for making sense of the fundamental tensions in NASA's innovation system occurring at multiple levels. At the same time, NASA provides a unique empirical setting in which to test the proposed combinatory exploration-exploitation mechanisms and to extend the theory in this new context.

1.2.4 Synthesis

In a macro-sense, all three views provide useful lenses for understanding innovation at NASA. In an otherwise conservative bureaucracy, leveraging windows of opportunity as they open is a

---

[5] ROSES is the umbrella R&D funding mechanism of which ESTO is an element (Research Opportunities in Space and Earth Sciences)

critical aspect of infusing new technology. At NASA, however, being ready to "*ride the wave*" is not just a state of mind. It requires more than a decade of pre-window development, which is currently managed as a stage-gate process. While stage-gates are effective in structuring formal project development, they impose an artificial order on pre-project R&D that is not (and likely cannot be) respected in practice. The stage-gate model is least able to capture the transition from R&D to flight projects - the aspect that the windows model describes best. The exploration-exploitation framing has the potential to clarify some of the key contradictions underlying the limitations of both the windows and the stage-gate views. Yet more research is required to understand how the explicit contradictions can, and should be, resolved in practice.

As described above, NASA both time-phases and structurally separates periods and activities of exploration and exploitation. It must thus resolve these activities' inherent contradictions at multiple levels of the organization. Viewed this way, a window opening describes the punctuation that allows exploration and exploitation to alternate cyclically. Stage-gates, on the other hand, seek to provide a framework allowing exploration and exploitation to co-exist in proportions set by the agency's strategic direction. The problem with maintaining both pursuits simultaneously is that, applied independently, their logics lead to fundamentally different and contradictory strategies for improvement. This thesis argues that the differences among the models lie at the levels of informal mechanisms and micro-behaviors, the levels which are least understood about NASA.

1.3 Study Design

This research seeks to understand how innovation happens at NASA. The phenomenon of interest – the so-called *innovation pathway*, which describes the sequence of events, actions and decisions that shape the technology trajectory from conception to first flight – is a difficult one to study. The process unfolds over multiple decades; is subject to extremely high sample mortality (on the order of one out of every hundred funded concepts are actually flown); involves multiple actors, institutional mechanisms and exogenous factors interacting in ways that make it difficult to define a system study boundary *a priori*; there is a paucity of directly relevant existing theory and empirical insights by which to guide the investigation; and, the mechanisms that have been identified (e.g., catalytic events which break down bureaucratic barriers) involve complex interactions not readily measurable with structured (e.g., survey) instruments.

In order to deal with these research challenges, this study adopts a process tracing approach. Researchers have argued that it is more appropriate than other methods for the study of phenomena characterized by complex causality (Hall, 2003) because it allows for the reality of feedback loops in social phenomena to be considered endogenously (Buthe, 2002). Process tracing is both a philosophy about what data to collect, and how to analyze them. In his classic text on organizational theory (Mohr, 1982), Mohr makes a strong distinction between "variance" theory (which explains phenomena in terms of relationships among dependent and independent variables) and "process" theory (which seeks to specify the sequence of events which lead to various outcomes). Specifically, the process tracing method, by going back in time to identify the key events, activities, or decisions that interacted in probabilistic ways to link hypothesized causes to the outcomes of interest, allows the researcher to specify the mechanisms linking causes and effects (Falletti, 2006). Commonly employed in studies of comparative politics (c.f., (Collier & Collier, 1991)), and increasingly in the organizational science literature (c.f., (Langley

& Truax, 1994; Nutt, 1984; Sonenshein, 2010)), it was also the primary method of many of the military innovation studies discussed above (c.f. (Lindsay, 2006; Sapolsky, 1972)).

The rich detail required to generate process theories necessitates a tradeoff of sampling breadth for depth. Hall suggests that, if falling on the small-N end of the sampling spectrum "*is the price we pay to understand complex causality, the trade-off is worth it*" ((Falletti, 2006) citing (Hall, 2003)). Regarding the size of N, Eisenhardt (1989) contends that 4-10 cases is probably close to the right balance; it is a small enough that the researcher can delve into sufficient detail to really understand the process while ensuring enough variability among carefully-chosen cases to generate meaningful theory. Following that advice, this research uses a multi-case, retrospective, longitudinal process study of six NASA *innovation pathways* as a basis for analysis. Figure 1-3 presents an overview of the research approach, with detailed explanations in the sections that follow.

1.3.1 Research Setting and Case Selection

The cases for this study were selected from the population of (1) new technologies (2) enabling NASA science missions, (3) with a strong Goddard Space Flight Center (GSFC, or Goddard) involvement, and (4) with some part of their development occurring during the decade of the 2000s.

We chose to focus on science missions, and particularly Goddard missions, for a combination of theoretical and practical reasons. Theoretically, the goal was to observe differences in how innovations traverse a government dominated innovation system. This required a multi-case embedded design (i.e., cases taking different paths through the same institutional context). Given NASA's "center and directorate" structure (described in detail in Ch 2), this design could most effectively be carried out in a single center-directorate setting. The Goddard Space Flight Center (Goddard), Science Mission Directorate (SMD) pair was chosen for the following reasons:

1) Access: Goddard's Chief Technologist served as a key project champion within the organization. He was in a position to make initial introductions, give priority to the study and potentially implement relevant findings.
2) Range of potential cases in:
    a. SMD (space science is an extremely technology intensive mission area with relatively unique needs, so there is significant directed R&D within SMD)
    b. And at Goddard in particular (Goddard is second only to the Jet Propulsion Lab in R&D awards)
3) Geographical proximity to MIT made frequent visits feasible over an extended period of time.

The population of candidate cases matching the above criteria was elicited from NASA experts involved in the development of these missions. Specifically, the Goddard chief technologist, project managers, theme technologists and funding managers were asked to suggest potential cases of <u>new</u> technologies that had been (or were being) <u>infused</u> into a flight project that had at

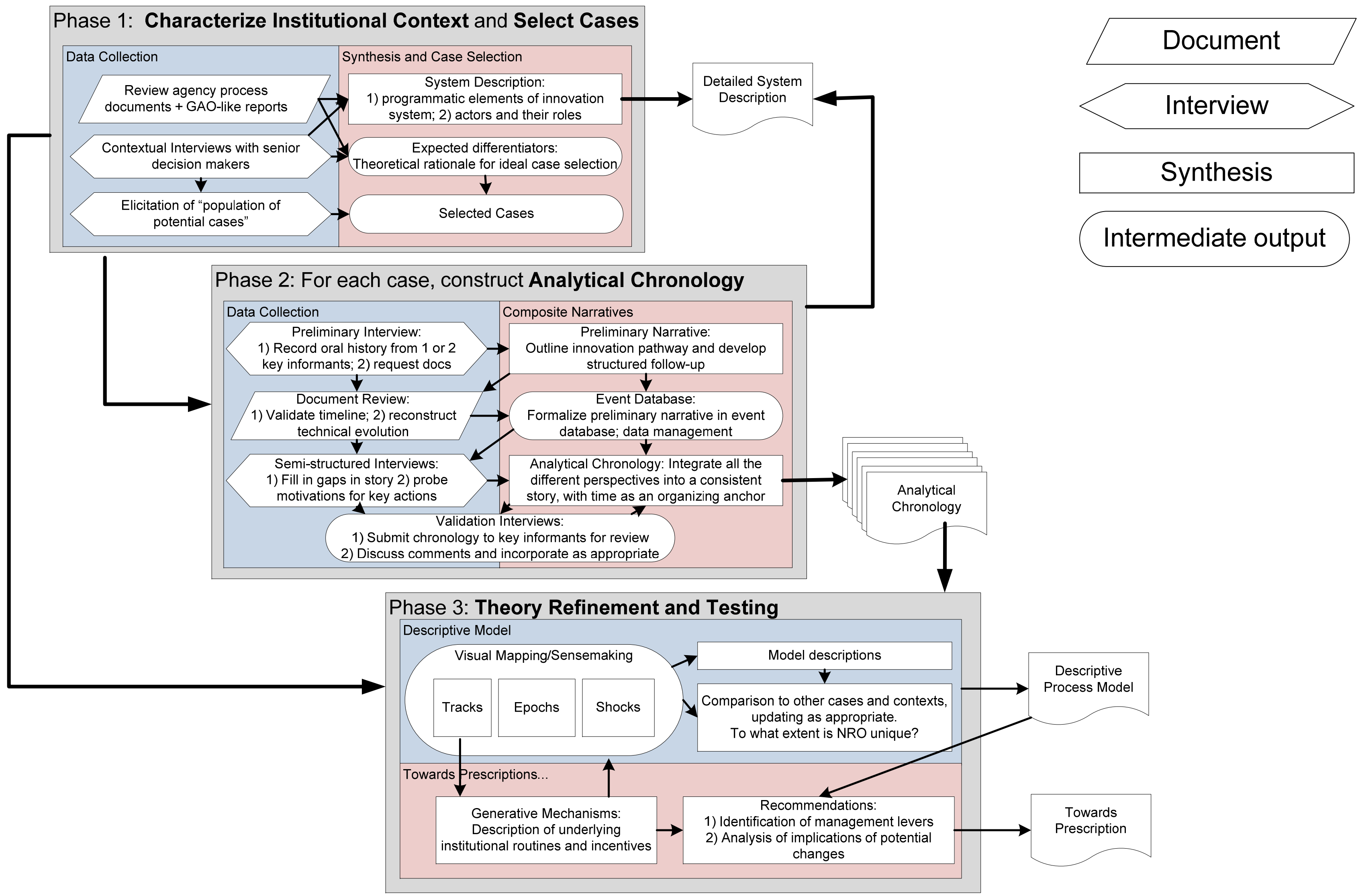

Figure 1-3: Research Approach

minimum passed "Milestone A" (NASA, 2007).[6] We continued asking new informants for potential cases until theoretical saturation was reached (i.e., no new innovations were being suggested) (Strauss & Corbin, 1990). This process generated a list of approximately two dozen potential cases.

Of that list of 24, half were eliminated immediately based on fit. Technologies that were being developed for "directed future studies" and/or implemented on "technology demonstrator missions" were eliminated because they represent special cases of infusion not subject to the normal approval process. Next, innovations that will be flown on the Mars Science Laboratory were eliminated because Mars missions are administratively separate from other science technologies (i.e., there are a whole set of dedicated funding mechanisms separate from the "normal" innovation ecosystem. Mars is a context that merits its own separate study); also, the types of measurement that are of interest on a planet's surface are hard to compare directly to the other science themes-lines. This left 12 potential innovations in Earth Science and Astrophysics all of which had been matured through the "same" institutional system.

The final set of six cases was selected for theoretical reasons (Eisenhardt, 1989) with a goal of variability in the <u>path</u> taken through the NASA innovation system. Since the shape of an innovation path is not knowable *a priori*, selections were made based on multiple expected indicators, namely, Priority, Level, Time and Nature. We also considered matched pairs (e.g., very different technologies tied to similar missions and contexts) where possible. While choosing a tight cluster of cases could potentially limit the generalizability of the study findings, it was believed that there was an equally great risk of missing key interactions (and limiting internal validity) if the cases were spread too thinly. Table 1-1 summarizes the final selection and the expected indicators. The intermediate list of 12 is provided in the Appendix I.

In terms of the indicators, Priority refers to both the political importance and magnitude of resources committed to the target mission. Within NASA SMD there are essentially two dichotomous classes of missions flown. The first consists of billion dollar flagship missions – one or two per theme per decade, selected based on a formal community prioritization process, explicitly pushing the technical frontier in a particular science area and accepting of significant within-project development. The second includes hundred million dollar "explorer class" projects, of which two or three are announced every two to five years, selected through a competitive Principal Investigator (PI)-lead bidding process, leveraging mature technologies to answer focused science questions. Past studies of space and military innovation suggest that Priority has an important impact on how technology development happens (McDougall, 1985; Rosen, 1994; Sapolsky, 1972). The selected cases were infused onto a range of Flagship and Explorer class missions.

[6] Milestone A is the systems engineering decision gate at which projects are formally approved. Although it often takes an additional several years before an approved project reaches its operational phase, the technology baseline is frozen barring unforeseen circumstances, making this an appropriate minimum standard of "infused." The main reason for choosing a relatively early milestone was to mitigate retrospective bias in interview account, and also to be sure that historical records would be obtainable.

“Level” refers to the location in the technical hierarchy (e.g., as defined by a work breakdown structure (NASA, 2007; Ulrich & Eppinger, 2008)) that the innovation changes. Both changes in the components of a system and the linkages among them can have important impacts on the evolution of a technology (Henderson & Clark, 1990). The initial intention was to select pairs of component, architectural and radical innovations; in practice, most of the potential cases were composites of multiple innovations (Sherwin & Isenson, 1967) at multiple levels. Table 1-1 captures both the level of the initiating insight, and the other key changes in brackets.

Over its history, NASA’s philosophy about, and infrastructure to support, non-project R&D has experienced several drastic restructurings (Figure 1-1). While the process tracing approach allows one to consider the effects of complex causality and history directly within each case (Buthe, 2002; Hall, 2003), we felt it important to examine a range of initiation time frames.

Nature (i.e., whether the innovation was competence enhancing or competence destroying) was included as another way to isolate how differences in the innovation might influence its path. At the industry level, competence enhancing innovations tend to be more incremental and perpetuate existing industry structures, while competence destroying innovations lead to periods of ferment and often the failure of existing firms (Anderson & Tushman, 1990). There are no firms within NASA, and its research groups technically do not go bankrupt, but one might expect competence enhancing innovations to exhibit different pathways (e.g., if it were easier to get early-stage grants).

Table 1-1: Overview of Selected Innovation Pathways

Innovation	Description	Priority	Level	Time	Nature
Continuous Adiabatic Demagnetization Refrigerator (CADR)	Solves key limitation of incumbent technology (no hold time). Improvement on traditional performance measures, too.	Flagship/ Explorer	Concept (and component)	Late 90s to present	Competence enhancing
Cadmium Zinc Telluride (CZT) detector	First detector in its class (room temp gamma-ray). Achieved significant position resolution improvement.	Explorer	Materials (and fabrication techniques (fab))	Late 80s to early 2000s	Competence destroying
Microcalorimeter (Semiconductor Thermometer)	Two order of magnitude resolution improvement (non-dispersive x-ray spectroscopy) – fundamentally new approach.	Flagship/ Mission of Opportunity	Concept (and components, materials, physics, fab)	Early 80s to ~2012	Competence destroying
Microcalorimeter (Transition Edge Sensor - TES)	Improved resolution (Si thermometers), enables scalability of array (where Si does not)	Flagship	Materials (and components, fab)	Mid 90s to present	Competence destroying
Quantum Well Infrared Photodetector (QWIP)	New fabrication strategy enabling multispectral thermal IR imaging with mature semiconductor materials	“Continuity” Flagship	Physics, fab	Late 90s to ~2010	Competence enhancing
X-ray Polarimeter	First practical X-ray polarimeter (two orders of magnitude resolution improvement)	Explorer	Concept (and components)	2001 to present	Competence enhancing

1.3.2 Data

Constructing each individual innovation pathway required that the sequence of events, processes, and decisions that link the initial concept to the eventual infused technology be identified. To this end, the data collection was extensive and involved two main sources: (1) semi-structured interviews with all the major participants in each case; and (2) documents produced during the pathway (e.g., grant proposals, scientific publications, personal communications, press coverage, etc.). These data were supplemented by informal observations made during multiple site visits. Table 1-2 summarizes the data sources. A complete list of interviews coded by project and informant roles is provided in Appendix II; the document list is divided by project, with references obfuscated to preserve the anonymity of the informants.

Table 1-2: Data Sources

Interview Stats	Document Stats
- **91 interviews totaling more than 100 hours (mostly individual; some repeat)** o **General (~38%)** o **Case Specific (~62%)** o **Breakdown of roles:** ▪ **Scientists (30*)** ▪ **Project/Technical division Management (19)** ▪ **Funding Managers (13)** ▪ **Engineers (10)** ▪ **Senior Leadership (5)** ▪ **External to NASA (3)**	- ~150 documents o Grant Proposals (~50%) o Journal Publications (~30%) o Project Reports (10%) o Funding records (5%) o Progress Presentations (2%) o Personal e-mails (1%) o Press coverage (1%) o Agency planning documents (1%)
Observations	
- **Approximately 1 week per month over the course of a year (including informal conversations in the shared cafeteria; sitting in on progress reports, impromptu laboratory tours and demonstrations)**	

***The bracketed numbers indicate the number of individuals in the category. For example 30 separate scientists were interviewed about their involvement in one or more innovations. The numbers do not add to 91 since some informants were interviewed multiple times and others were interviewed in small groups.**

More than 100 hours of interviews with key participants were conducted on-site at Goddard, and at some of the small businesses involved with the projects. The interviews were used for two purposes. First, a subset of initial interview accounts served to create a sketch of the critical events in each pathway, helping to focus the document search that followed. Primary records, like contract documents, were then used to validate the details of the timeline, since decade-old memories can be fallible. Some of these documents were publically available in theory, but NASA does not maintain any centralized R&D record system, nor any systematic R&D records dating back longer than five years. So, as part of the initial round of interviews, informants were asked to look through their historical files. Individual records were surprisingly complete, and through the several core scientists/technologists per case we were able to obtain the vast majority of paper work produced during the innovation pathway.[7] Once the more concrete aspects of the

[7] Among the core team, there was a clear sense of the total population of relevant paper-work. As a result, we were aware of the documents that could not be found as well.

pathway had been established, a second round of interviews was used to probe the motivations of the actors and understand why particular decisions were taken at different times.

Interview subjects were selected initially based on introductions from Goddard's Office of the Chief Technologist; as the research progressed, more subjects were identified during the first round of interviews who could provide complementary perspectives. This strategy of snowball introductions was pursued until the key individuals involved in the particular innovation pathway had been interviewed.[8] All interviews were digitally recorded using a livescribe smartpen™, with permission.[9] Transcriptions were made as appropriate.[10] Interviews were semi-structured and lasted for as long as the conversation was productive. The shortest was 20 minutes and the longest was 4 hours. Most were close to 1 hour. Following (Rubin & Rubin, 2005) each interview began casually, transitioning into the formal interview with safe background questions; these focused on educational and professional history prior to joining Goddard, and other projects with which they had been involved. The middle portion of the interview was loosely structured around an on-ramp/process/off-ramp model of a person's involvement. Such questions included: how did you become involved with innovation X, what was your role(s) during the project and how/when/why did you stop being involved? The vast majority of each interview was spent on the process part. Each interview was concluded by offering the interviewee an opportunity to ask questions about the project, or share anything else they thought relevant. The interviewer asked for permission to contact them to follow-up with any further questions.

Early interviews about a particular pathway were more exploratory, with later interviews focused around particular events in which the interviewee was intimately involved. Follow-up interviews were conducted with early subjects to ensure consistent depth in the final account. Depending on the interviewee, more or less time was spent discussing the technical details of the case; while this study focused primarily in the path and the ways that the formal and informal structure constrained and enabled the choices, spending time discussing the technical details allowed the interviewer to establish technical credibility[11] and a rapport with scientists and technologists. "Shop talk" also provided a safe anchor to return to if I perceived any discomfort with the organizational line of questioning. In the end, the technology-oriented events proved to be a critical part of the story.

Interview accounts were triangulated with more than 250 archival documents (Yin, 2009). Access to the relevant material was excellent. As discussed above, the most valuable data sources were grant proposals written on a yearly basis over the course of the entire pre-project pathways. These documents gave insights into what the innovators were thinking at various

[8] In each of the innovation pathways in Ch 3, the interview coverage for each case is specified precisely.
[9] Of the ~90 interviews, only one informant declined the recording, but several interviewees asked that the recording be paused for potentially sensitive responses.
[10] The livescribe smartpen links audio to handwritten notes made during the interviews. This proved extremely useful in going back through searchable notes. As a result only sections used for direct quotes were deemed necessary to transcribe.
[11] The author holds two degrees in Aerospace Engineering and two years of professional experience working as a space systems engineer. I presented myself to the interviewees as an MIT aerospace engineer interested in understanding how innovation happens at NASA. Although I had limited prior specialized knowledge techniques and theory associated with the cases I studied, I had the background necessary to get up to speed quickly, and was identified as a fellow space-trained technical person.

points in the process. Some informants even found old e-mails and memos capturing key correspondence during the brainstorming phase of a project initiation. More traditional documents including journal articles, reports, press coverage, legislative actions (in one case) were also collected. In addition to verifying and providing the context for the chronologies remembered by the informants, the documents also captured a detailed record of quantifiable characteristics of the pathways; for example, the number, and names, of collaborators at all points during the pathway or the grant support, and sources of that support, received by core members of the team.

Although the interviewer could not explicitly make observations of the innovation pathways under study, the extended period spent on-site contributed greatly to the richness of my contextual understanding of the process. I ate lunch in the shared cafeteria (described in the microcalorimeter case), sat in on lunch meetings and open review sessions, and started informal conversations in each of these contexts. Several interviewees offered impromptu tours of their facilities or put on demonstrations of their devices. I never turned down a tour and spent time getting "introduced around" and learning how things worked. Being immersed in the culture both improved my understanding and enabled access in unforeseeable ways.

1.3.3 Data Analysis

The challenge of "sense-making" from process data is one of how to move "*from a shapeless data spaghetti toward some kind of theoretical understanding that does not betray the richness, dynamism, and complexity of the data but that is understandable and potentially useful to others*" (p. 694) (Langley, 1999), while avoiding "*death by data asphyxiation"*(Pettigrew, 1990). Data analysis was conducted in two sequential phases, first building analytical chronologies for each case and then conducting cross-case comparisons.

Within Case Analysis

Three related strategies were used to become intimately familiar with each case as a standalone entity (Eisenhardt, 1989). Following Van de Ven et al. (2000), event data bases were constructed for each case. This allowed me to keep track of the diverse data sources and be clear about the sequence of events. For each event, the relevant date, a brief description of the event and raw supporting evidence (e.g., relevant quotes, sections of documents and/or pointers to websites) were included. The content of each event was also coded, initially using the codes (tracks) defined during the Minnesota Innovation Research Program (MIRP), later tailored to the NASA science context as the analysis progressed. The event database maintained traceability between the raw data and the abstractions that emerged (Yin, 2009). We initially intended to employ the quantitative process methods developed during the MIRP (Van de Ven, Angle, & Poole, 2000) to analyze the event sequence, but found the necessary coding fidelity too coarse to provide the desired level of understanding we were interested in.

Instead, the technique that Langley (1999) calls visual mapping was employed. This strategy of graphical representations allowed a large number of dimensions to be represented simultaneously and concisely. Easily rearranged and reoriented, the visual codes could be viewed, and made sense of, from multiple perspectives (Eisenhardt, 1989). Multiple representations for each case were explored in an attempt to understand the co-evolution and interactions of the different tracks over time. Early representations used calendar time as an anchor, but as the visual coding became more sophisticated, other scales were found to yield more interesting insights. For example, "event time" revealed common patterns across the cases, and "technology maturity level" revealed non-linearities in the current conceptualization. Figure 1-4 provides a graphical representation of the data processing from raw data to the mid-range theory embodied in the process model presented in Chapter 4.

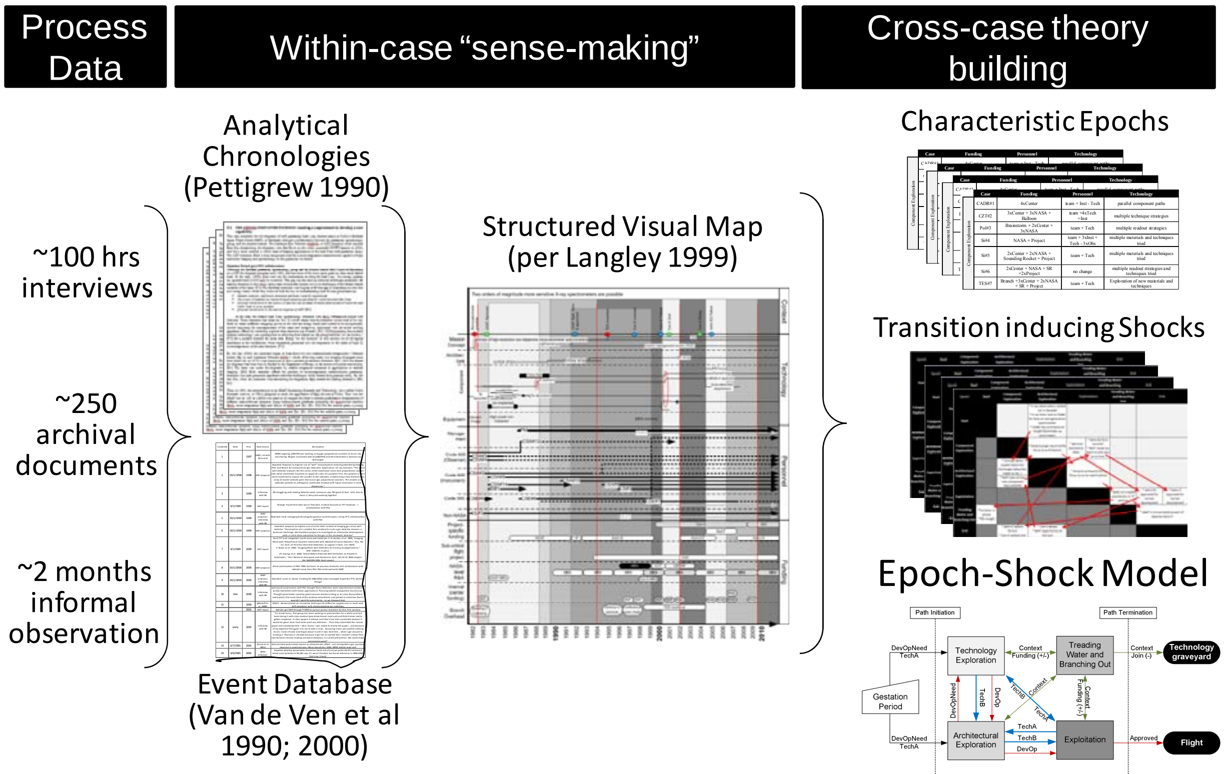

Figure 1-4: Overview of Analysis Strategy

In parallel with constructing event databases and visual mapping, an analytical chronology was developed for each case. As described by Pettigrew (1990), narratives of this form allow the researcher to "*get on top of the data, to clarify sequences across levels of analysis, suggest causal linkages between levels, and establish early analytical themes*" (p.280) (Pettigrew, 1990). Although visual maps generally capture a higher level of abstraction than narratives, each view provided a different type of clarity. Whereas, as Langley (1999) cautions, certain types of (easily visualizable) data were probably overrepresented in the visual maps, the visual nature of the work enforced extreme clarity in other aspects. Writing the narratives, on the other hand, tended to overrepresent certain other types of data. Specifically, the composite narrative (Pentland, 1999; Sonenshein, 2010) provided a framework to integrate the "*fragments of stories, bits and pieces told here and there*" that constitute each individual's perception of the innovation into the researcher's composite interpretation. This allowed multiple perspectives on various issues to be captured, but came at the expense of the temporal and relational clarity enforced by the visual abstraction. For this reason, we chose to preserve the strength of both strategies by constructing the innovation pathways that make up Chapter 3.

Each innovation pathway was validated as follows. A draft of a near-final narrative was provided to each of the core informants for a given case. They were asked if they felt that anything important had been missed, or misrepresented. Feedback back was incorporated iteratively until all parties were satisfied. There was one case with only small changes where consensus was reached after a second iteration.

Cross-case Analysis

The literature does not provide clear direction on any best way to develop cross-case insights; rather, it emphasizes the need for iteration between data collection and analysis (Glaser & Strauss, 1967; Yin, 2009), suggests a need for synergy between systematic abstraction (potentially quantitative) and qualitative richness (Mintzberg, 1979), and cautions that care must be taken to overcome human information processing biases (Eisenhardt, 1989). To this end, Eisenhardt (1989) recommends counteracting these tendencies by looking at the data in as many divergent ways as possible.

Starting from the narratives and visual maps described above, we looked for similar patterns and event sequences across cases, comparing the preliminary event and track codes. The multiple representations for each case facilitated the comparison process. As patterns began to emerge, our focus circled back to the data, checking for the presence of each concept across the other cases in a semi-quantitative fashion. For example, using the event sequence described above, one can easily measure the relative frequency with which a technical breakthrough appears in close temporal proximity to the beginning of a funding shortfall. The results of these "tests" were then used to refine the emerging concepts in an iterative fashion as is common in inductive research (Miles & Huberman, 1994). The emerging concepts were also compared to the extant literature to establish the broader existence of the patterns.

1.3.4 Limitations

In this study design, threats to validity were controlled where possible. Multiple pieces of evidence were used to corroborate the analytical chronologies upon which much of the later analysis is based (Yin, 2009). Interviews were used to understand archival material and documents were used to verify the memories of individuals. A chain of evidences was maintained in an event database to ensure traceability to the raw data. Consensus panels of experts were used to define the study population frame. Further, process tracing as a method explicitly considers time endogenously, controlling for maturation and history. However, some threats to validity cannot be reasonably controlled; instead, their implications for the scope of the generality of the findings must be acknowledged and respected.

One such threat is bias due to sample mortality. In this research, the study frame only includes successful innovations and the process tracing was retrospective. This design choice was made out of necessity; the population of early failures (i.e., concepts that never made it past the basic research phase) are extremely difficult to identify *a posteriori*, and, while a forward-tracing study would overcome this issue directly, with innovation cycle times on the order of multiple decades and with mortality rates as high as 99:100, the resources and time required for conducting a forward-sampling study of NASA innovations are not conducive to PhD research.

This threat to internal validity was mitigated to a certain extent through research framing. Recall that the questions central to this study focus on the differences among pathways taken by successful innovations, not the factors making an innovation successful (i.e., the dependent variable is the pathway-outcome pattern, not the success/failure of the innovation); thus, the reality that some dead-end pathways will be missed is less important to the validity of the conclusions. Nonetheless, there will be no way to say with certainty that particular paths caused the success (since the same path may have also led to failed innovations) based on the present study. Going forward, there is a plan to track the technologies studied herein as they are used on future missions. A decade from now, we can expect variability in terms of the extent of success (i.e., some technologies will have only been used once, or failed during the implementation phase, and others will have become standard on multiple missions). Today's insights can then be revisited and a richer understanding be developed.

Finally, with respect to external validity, six cases are clearly not representative of innovation in space agencies in general. Given the complexity of innovation as a phenomenon, it is believed that, without delving into significant detail on a small number of cases, important details will be missed. To enable some level of generalization, the findings were compared to theory and empirical studies in other domains, published in the literature (Sonenshein, 2010).

1.4 A Brief Overview of the Thesis

The remainder of the thesis is structured in four main sections. Chapter 2 describes NASA's underlying system architecture, considering both the explicit and actual structure as described by our informants. This provides both the background necessary to understand the innovation pathways (Chapter 3), characterization of the model (Chapter 4), and strategic implications (Chapter 5). Chapter 3 captures the empirical basis upon which the later analysis is based. It presents an analytical chronology and visual map for each of the innovation pathways studied herein. Chapter 4 then serves the dual purpose of synthesizing the cross-case analysis and articulating a new Epoch-Shock conceptualization of the process. Finally, Chapter 5 brings all the pieces together, addressing the implications of the new model and highlighting key strategic tradeoffs in terms of our new understanding of the underlying system architecture.

2 Underlying Systems Architecture

> *A lot of the center directors really have a fundamentally incorrect picture of what goes on. They think that scientists analyze data and send requirements over to the engineers and then the engineers go and invent and develop everything. But the illusion that you can go and tell someone, 'hey, can you invent something for me?' is just so far from any kind of reality that it's hard for me to believe that they fall for it, but it's true [that's what they think]! Everything that I've seen that's been developed here, ab initio, has started on the science side and resulted in a big collaboration [with engineering].*
>
> -- Staff Scientist (at Goddard since late '70s)

What is the structure of NASA's innovation system? By way of answer to this question, this chapter maps the components of the U.S. Space Science innovation system, including institutional, human and technological elements of the structure, and the connections – explicit and organic – that link them. It considers not only NASA's internal structure, but also the network of small businesses, contractors and academics that play critical roles in NASA's efforts.

The notion of an "innovation system" was first introduced by Lundvall (1985) and popularized by Freeman (1987), who coined the expression "National Innovation System (NIS)" through his study of Japanese Economy. Freeman defined an NIS as "*…the network of institutions in the public and private sectors whose activities and interactions initiate, import, modify and diffuse new technologies.*" (Freeman, 1987) Similarly, Lundvall focused on "*the elements and relationships which interact in the production, diffusion and use of new, and economically useful, knowledge ... and are either located within or rooted inside the borders of a nation state.*" (Lundvall, 1992) Both definitions take a broad perspective on the set of people, components, institutions, and the flow of information among them, which must be considered in order to understand the productivity of a nation, or, in our case, the well contained space science sector.

Thoroughly mapping the U.S. Space Science innovation system easily merits a volume unto itself. The objective here is merely to provide a sufficiently deep basis for discussing: (1) the dynamics captured in the process model of NASA's pre-project development system; and, (2) the remedies proposed in Chapter 5.

2.1 Funding Mechanisms and Their Cycles

At NASA, as in many bureaucracies, one can most easily observe the structure of the institution by "following the money." According to NASA's annual budget, funding is distributed by mission directorate (i.e., Aeronautics, Exploration, Operations and Science) and then allocated between flight programs and research within each directorate. Since our focus is on pre-project development, flight resources are only discussed in terms of their relationship to R&D.

Table 2-1 presents a conceptual overview of how technology development funding is bucketed within Science, with examples from each of the cases studied in this thesis. The buckets are categorized in terms of emphasis, magnitude and duration of support. There is an expectation that new concepts move through the system from left (novel concept) to right (implemented on scientifically important flight mission), requiring an order of magnitude of more funding resources at each next stage of maturity. The number of funded technologies is expected to be winnowed down at each stage, with only a few of the hundreds of concepts that enter the system finding use on flight projects. The following sections describe each of the funding "decades" and the extent to which there are explicit connections among them.

Table 2-1: Classification of Funding Sources

		Technology Development			Flight Development		
		Branch overhead	Center-level "Internal R&D"	NASA-level "Research & Analysis"	Sounding Rockets & **Balloons**	Explorer class	Flagship class
Focus	Funding level	<$10K	$10K - $100K	$100K - $1.5M	~$2M	$100 -$600M	~$1B
	Timeline	~weeks	<year	<3 years	<3 years	<5years	10+ years
	Cycles	continuous	yearly	yearly	yearly	variable	decadal
Instances from cases	CADR	branch overhead	DDF, IRAD, CTD	CETDP, IPP seed fund	**PAPA/PIPER**	Astro H, ASP	Con-X (IXO)
	CZT	group overhead	DDF, SBIR PI	RTOPs, SR&T, NDE PO	**GRIS (Portia) InFOCuS**	BASIS, SWIFT	EXIST
	Si Microcalorimeter	"free time"	DDF, IRAD	RTOP/SR&T, APRA, CETDP	XQC	Astro EII, Astro H	AXAF(Astro E)
	TES Microcalorimeter	group overhead	IRAD	APRA	XQC		Con-X (IXO)
	Polarimeter	"free time" and "scraps"	DDF, IRAD	APRA		AXP "category 3", GEMS	
	QWIP	collaboration meeting	SDI, DDF, SBIR PI, IRAD	ESTO, SBIR PII			LDCM

Acronyms: Director's Discretionary Fund (DDF); Internal Research and Development (IRAD); Commercial Technology Development (CTD); Small Business Research and Innovation (SBIR) phase I and II; Strategic Defense Initiative (SDI); Cross-Enterprise Technology Demonstration Program (CETDP); Innovative Partnership Program (IPP); Research and Technology Objectives and Plans Summary (RTOPS); Sustaining Research and Technology (SR&T); Non-Destructive Engineering Program Office (NDE PO); Astronomy and Physics Research Activity (APRA); Earth Science Technology Office (ESTO); NASA Research Activity (NRA); The remaining acronyms are project names defined in context.

2.1.1 Branch Overhead

In this context, brainstorming refers to the informal activities that precede formal applications for technology development funding. Since there is a certain cost associated with proposal writing (in terms the opportunity cost of writing the proposal vs. working on other projects) and there is an inherent uncertainty as to whether that proposal will be funded (and if the idea will go anywhere), scientists and technologists tend to spend some time on *back of the envelope* calculations designed to convince themselves that they have a concept worth pursuing. There is no direct funding for this type of activity, but a few mechanisms do exist to encourage these creative interactions. For example, in the QWIPs case, the GSFC technologists were sent to JPL to brainstorm and find an area for collaboration. Similarly, trips to conference meetings are an important (funded) opportunity to brainstorm and cross-pollinate ideas.

2.1.2 Center-level "Internal R&D"

Once an idea has been fleshed out, a small amount of seed funding is needed to develop the concept so that it can be pitched to the larger, more stable technology development funding. Nominal uses for these \$10K – 100K include buying parts for breadboards, machine shop time, and paying civil servants' salaries.[12] Investment in this early stage concept development is currently left to the discretion of the NASA centers. They exist at the discretion of the center director, which means that there's no explicit NASA-level funding line for them; rather, on a center by center basis, management has decided that investing in promising ideas is "*more important than cutting the grass as often as we'd like.*"[I3] This speaks to the reality that early stage funding is currently minimal and highly variable across centers. At Goddard, over the last few decades, this funding has existed in two forms: the DDF (the director's discretionary fund) and IRAD (internal research and development). Although they covered a similar place in the maturity spectrum, they employed quite different operating philosophies. DDF was designed with a high-risk, high-reward mindset, administered by the senior science fellows within the center. There was limited traceability in funding decisions and follow-up, and the perception is that the programmatics emphasized trust in the judgment and expertise of the individual recipients [I1, I3]. IRAD on the other hand, which has superseded DDF, is administered by the Goddard Chief Technologists Office. It serves as strategic funding designed to win future missions for Goddard as a center. As a result, proposers are expected to demonstrate clear links to (potential) future missions, even for early-stage developments. Only civil servants can apply for IRAD funding.

The other main source of concept development funding is from the small business innovation research (SBIR) program. SBIR is a congressionally mandated program[13] which is administered at NASA by the Agency-wide Innovative Partnership Program (IPP). The program requires that NASA technologists work with small businesses on innovative, low TRL concepts.[14] The phase I awards are designed to prove-out the concepts and determine if the relationship is worth pursuing.

2.1.3 NASA-level "Research & Analysis"

Within SMD, the first stage of formal technology funding is covered by NRAs (NASA Research Activities), with the Principal Investigator (PI) led grants administered through the omnibus ROSES (Research Opportunities in Space and Earth Sciences) solicitation. ROSES administers both Research & Analysis (R&A) and instrument development funding (relevant to the types of technologies being considered herein). ROSES encompasses many theme-specific program elements of which a small number are relevant to technology. These include ESTO (Earth Science Technology Office) and APRA (Astronomy and Physics Research and Analysis). In the discussion of the innovation pathways, we use the terminology employed by the informants (interchanging APRA/ROSES/NRA fairly liberally). These are peer-reviewed, multi-year grants, on the order of \$100K to \$2M, designed to support the pre-development of flight instruments for future missions. There therefore needs to be a clear mission application expressed in the

[12] As civil servants, NASA employees are required to account for all their hours with charge codes; and time spent developing new concepts needs to be charged to something. IRADs are one way of covering this cost.
[13] For an overview of the SBIR program see (NRC, 2007)
[14] For an overview of the Technology Readiness Levels see (Mankins, 1995)

proposal. Centers tend to have small amounts of money to support proposal development for these types of awards.[15] Given the science focus of ROSES, only explicitly science-enabling technology is funded through this mechanism. So-called "cross-cutting" enabling technology (products like the CADR or optical communications) have been funded through other mechanisms at different points in history (e.g., CETDP - Cross Enterprise Technology Development Program; or NIAC – the NASA Institute for Advanced Concepts).

Today, these "level 2" grants are openly competed for and any PI (whether they are a NASA civil servant, contractor, or otherwise employed) can submit proposals. However, it has not always been this way. Prior to ~2000, the closest equivalent to ROSES were called RTOPs (Research and Technology Operating Plans) and SR&T (Supporting Research and Technology). While they filled a similar place in the maturity spectrum, their selection criteria and purposes were quite different. Where ROSES are "open" and fixed term, both SR&T and RTOPs were "directed" and flexible. Directed, in this context, meant that headquarters allocated resources to be used at the discretion of the recipient. For example, the Detector Development Lab (DDL) would be given some set amount of money on a yearly basis to invest as the branch head saw fit. In the words of a branch head: "*Back then, the guidelines were very loose. [Grants read] something like we entrust you to go and understand the state-of-the-art in this area.*" [I59] While, in today's NASA, established groups tend to receive a steady stream of ROSES money, the grant writing burden (and associated uncertainty) is significant. As put by one scientist: *it feels like you're applying every year for three-year funding.* The other key difference was that RTOPs and SR&T were only awarded to employees of NASA. Now, NASA scientists compete with the community at large, and technology groups like the DDL are "paid for service" by scientists from their grants.

Phase II of the SBIR program also falls into this category, and provide funding of $600K over 2 years. Companies are often expected to develop prototypes in this phase in order to prepare for follow-on commercialization or infusion into NASA missions (phase III).

2.1.4 Sounding Rockets and High Altitude Balloons: Technology Demonstration

Sounding rockets and high altitude balloons serve the dual purpose of raising the TRL of new technologies and providing a fast and inexpensive platform for certain types of science. Sounding rockets follow suborbital parabolic trajectories. This enables above the atmosphere observing times on the order of minutes. This is adequate to perform some types of "real science" (e.g., diffuse X-ray background, but not focused looks at particular sources). The vibration environment is much more intense than satellite missions (so, from a TRL perspective, it is not a "like environment"), but the qualification requirements are much lower. As a result, "*sounding rockets are about the only way for grad students to learn about building flight hardware… [otherwise, you] need a NASA certified soldering technician to move a wire.*" [I93]

Balloon flights are much longer. Conventional flights last 2 to 36hrs, while long duration flights can last more than 40 days. They reach an altitude of 160 000 ft and can suspend large payloads (up to 1650 lbs). As a result, for some classes of measurements (e.g., for which it is not required to leave the atmosphere or target a fixed source), sequences of balloon flights can serve as the

[15] For example, Goddard has a small proposal support office.

main source of data. For others, balloon flights are a great way to calibrate detectors and to prove operational concepts.

Funding for both programs is managed through a separate element of the ROSES program. Like other ROSES grants, rocket and balloon flights are selected through a competitive peer-review process. In the cases described in Chapter 3, both sounding rockets and balloon campaigns served important roles in the innovation pathway, but were not ends in-and-of themselves.

2.1.5 Flight Specific Development

NASA flies a mix of mission classes, differentiated primarily by funding level (e.g., in the astrophysics context, Flagships are more than $1B in lifecycle costs, explorers (SMEX/MidEX) are low hundreds of millions and medium missions are in between). However, the price differentiation also corresponds to risk acceptance and technology development investment.

The explorer program was designed to enable "*rapid responses to new discoveries and provides platforms for targeted investigations essential to the breadth of NASA's [science] program.*" It supports small (SMEX) and medium (MidEX) missions selected through competitive peer review to be developed in a 5-year timeframe. Since the program was designed to be responsive to developments (both scientific and science enabling technological), there is no macro-level portfolio strategy for the program. However, there is a relatively high expectation of technology readiness. Since it is a capped program (i.e., fixed price) with a tight schedule, for proposals to be competitive they must be ambitious, but also believably feasible. In theory, explorer missions are not *allowed* to require any new technology development (with the exception of one planned "miracle" that will not interfere with overall mission success if it fails [I28]). Proposals are ranked in terms of both scientific value and technical feasibility. Category 1 = high science value and mature technology (usually selected for phase A study); Category 2 = low science, mature technology (not funded); Category 3 = high science, low maturity (sometimes get directed tech development funding). Although there is a yearly budget line-item for the explorer program, the Announcement of Opportunities (AO) come somewhat sporadically at a rate of several a decade.

Flagship missions, on the other hand, are expected to revolutionize a branch of science, and therefore invest heavily in technology development (i.e., they may incorporate a technology as low as TRL 2 during phase A and invest directly in its maturation). With 10-plus year development cycles, they can support fairly major long-lead technology development initiatives, and recent missions have invested tens of millions in this effort. It should be noted that the LDCM $10M investment in TIRS and QWIPs is slightly different. LDCM is not really a flagship mission, and the tech development was not initiated early in the program as a planned innovation mechanism. Nonetheless, it is fulfilling the same function of stable, substantial, project-specific technology development funding. One clear difference between project specific funding and the other stages of funding is that individual scientists and technologists cannot bid for it *per se*. While explorer class mission are Principle Investigator (PI)-lead, with flagships, it is the mission that is being proposed, not the technology. In the Flagship case, in theory, the resources are allocated by the project to baselined instruments that are deemed to need it (although this is not always what happens in practice – c.f. microcalorimeter pathway).

2.1.6 Summary of Funding Structure

It is noteworthy that each of these funding mechanisms proceeds on a different, but relatively fixed, cycle. Brach overhead is spent relatively continuously and on an as-needed basis. Civil servants can propose to center overhead on a yearly basis, around November each year. ROSES proposals are accepted yearly, in the February timeframe, and, if awarded, generally last for three years. Explorer announcements of opportunity (AOs) come out sporadically several times a decade. Five or six Phase A Explorer study contracts are typically awarded; they last short of a year, at which point one or two are selected for full development. Flagship missions (particularly in Astrophysics) tend to follow the decadal cycle. Based on the last decades, there is only enough budget to develop one mission per decade, with another two or so receiving concept and technology development funds.

There are very few structural connections among these mechanisms. In fact, the only explicit transition between any two funding buckets is the progression from an SBIR phase I to an SBIR phase II. As described in the QWIPs case, in the SBIR program, there are a series of review stages where COTRs (contracting officer's technical representatives) advocate for the companies and outputs are ranked. The overall allocation is managed as a portfolio at the agency level. There are no similar mechanisms connecting any of the other funding buckets. In fact, Figure 2-1, adapted from the NASA systems engineering handbook, illustrates the duality between the formal mission development structure and the patch-work of technology development funding mechanisms. Note that more than one project will be under development at any given time and multiple technologies will receive grants. The extent to which new technologies navigate this infrastructure is strongly determined by the facilitating roles played by individuals in the system (discussed in the next section) and their social networks.

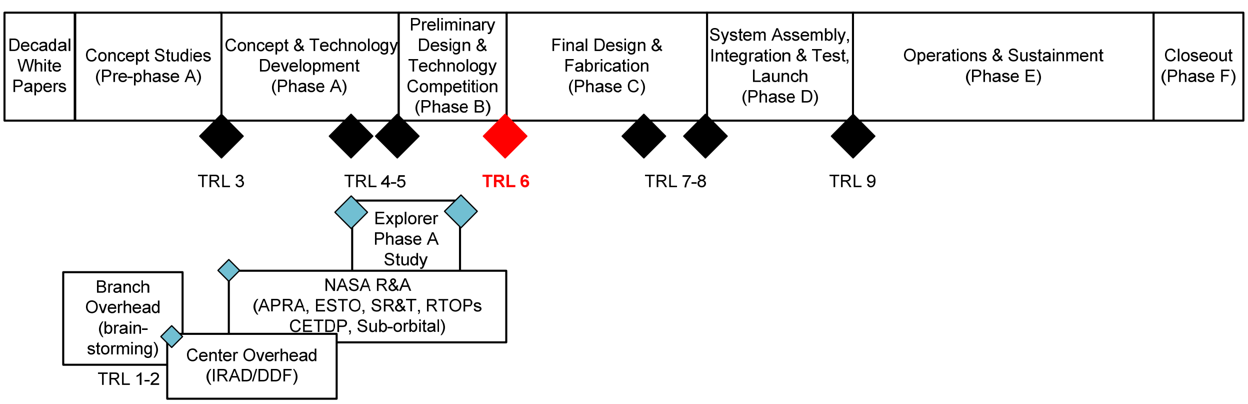

Figure 2-1: Conceptualization of the Nominal Technology Development Structure

2.2 The Human Organization: Roles and Responsibilities

NASA organizes its human resources by "code." Figure 2-2 reproduces a portion of Goddard's organization chart. The names are self-explanatory. Headquarters (HQ) functions are code 100; Management Operations are 200; Safety and Mission Assurance are 300; Flight Projects are 400; Applied Engineering & Technology are 500; Sciences & Exploration are 600; Information Technology & Communications are 700; and Suborbital & Special Orbital Projects are 800. Figure 2-2 also shows the additional level of detail for a representative engineering (500) division and science (600) theme, as these job classes are most relevant to this research. When discussing the level of the organization, we will adopt the labels used by the informants. On the engineering side, this means that directorates are made up of divisions which are made up of branches. On the science side, directorates break down into themes which are further divided into laboratories.

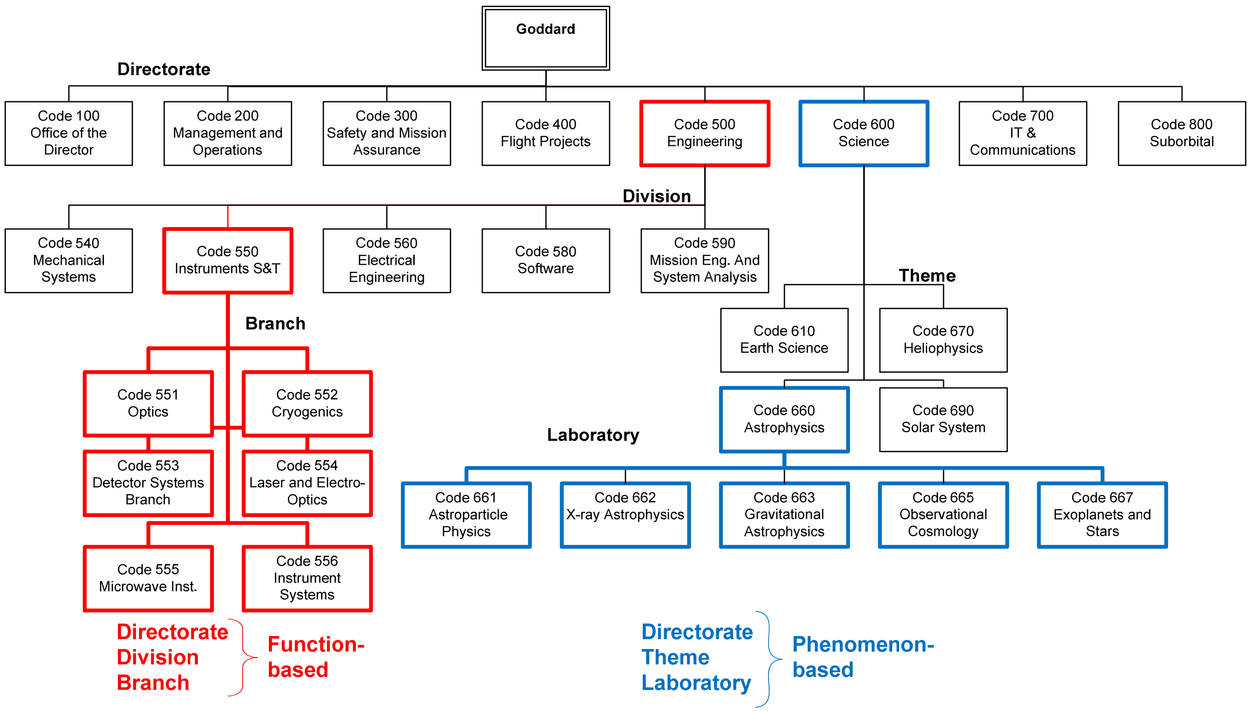

Figure 2-2: NASA Organization Chart

In the narratives and visual maps in Chapter 3, individuals are referred to by a two or three letter code that designates their formal titles. Specifically:

CSA = Civil Servant Astrophysicist (evolved into any civil servant scientist)
CSE = Civil Servant Engineer
CS = Contractor Scientist
BH = Branch Head
PM = Project or Instrument Manager
PS = Project Scientist
St = Student (working at NASA)

The non-NASA job functions were designated:

SB = Small Business
UL = University Lab (we considered the professor and his/her students as a unit)
IR = Industrial Researcher
NL = Researcher at a National Lab

Over the course of the three decades considered in the pathways, some of the informants held more than one job. For example, CSs were hired as CSAs; CSAs "retired" as ULs; UL students graduated and took jobs as NLs or CS/As etc… in most cases each individual had held one job that was most relevant. Were there was any categorization discrepancy, the individual was designated by their self-declared identity.

For the most part, the non-NASA SBs and IRs fulfill similar functions as NASA CSEs, PMs and BHs (i.e., engineers by training); whereas the non-NASA ULs and NLs that we interviewed shared functions in common with NASA CSAs, CSs, PSs and Sts (i.e., scientists by training). Nominally, scientists analyze data and specify requirements for future missions, while engineers develop the enabling technology. However, in practice, particularly in the context of pre-project innovative activities, the boundaries are gray and some scientists share more in common with engineers than their disciplinary brethren. As a result, it is more insightful to describe the roles and responsibilities of individuals in the system in terms of functions rather than job titles.

PhD scientists play a diverse set of roles within NASA and in the broader community, not captured in the above organization chart. As described in the chapter quote, the most traditional scientist role is one of data analyzer, either inside or outside NASA. However, many NASA scientists serve the role of instrument developer*;* some of the most vigorous tinkerers actually take jobs in code 500, where instrument design and testing is their primary function. Tinkering is a less valued skill outside of NASA. There are also several administrative roles that require the expertise of a PhD-level scientist. Many scientists perform work spanning more than one of these roles. The subsections below describe these functions individually and then discuss the relative differences and preferences.

2.2.1 Scientists as Observers

The governing motivation for building science satellites is to collect data that, when analyzed, reveals new insights about how the universe works. Scientists gain recognition and status by making discoveries and publishing journal papers documenting the results of their analysis. A university astrophysics professor can spend an entire career without touching the instrument that makes his or her measurements. Scientists can get access to data in a number of different ways; they can serve as Principle Investigator on large or small, orbital or sub-orbital missions; or, they can win grants to serve as "guest observers" on missions that are currently in orbit. To the extent that there are pure Observers within the ranks of NASA scientists, they still play an important bridging role with respect to new technology infusion. Being collocated with instrument developers (often sitting down the hall), they are much more aware of new developments than they would be from a position at a University. For example, in the semiconductor microcalorimeter case, the X-ray astronomer who "came down the hall" to talk about the future of spectrometers was primarily an observer.

2.2.2 Scientists as Instrument Developers

Observers require the technical infrastructure furnished by instrument developers in order to do their work. Although solving technical instrument issues is sometimes considered scientifically mundane (in part because they result in fewer publications), this function is extremely important to NASA's mission. As described by UL#2, one does not typically get tenure in a physics department for building a radically improved instrument (this yields only one paper), even though it is necessary for someone to do it in order to collect the data which does garner tenure. Fortunately, some scientists love to tinker and NASA welcomes them with open arms. The following quotes capture this sentiment from several perspectives:

> *I like gamma-ray bursts. They're interesting. But I'm not really an analysis person, I guess. I realized that, in the end, I want to build things that go find the answers, and then*

the scientists can then go and determine what they mean… we're also "scientists" because we're physicists… but we're not desk jockeys sitting behind computers. We build stuff in the lab.

-- PhD in Physics, Polarimeter case

I am trained as an Astrophysicist but I have always been very instrumentation oriented. My interests are pretty general, so I found a lot of opportunity at NASA.

-- PhD in Physics, CZT case

My task is to try to solve problems on the development front and actually make solutions... Who wouldn't like to come up with an idea, you actually make it, and then you measure it… and you bring it to a result?

-- PhD in Physics, spends half time in DDL

I mostly had done the basic physics, so it was a big switch to do detector stuff. During the post doc is when I started to acclimatize …I had opportunities to stay in basic physics... but it just seemed to be getting more and more esoteric so I went to visit [a National Lab]… I was really impressed at how they had the ability to work on what they wanted and develop really interesting new technologies and systems to use those technologies and it seemed like an exciting way to go…

-- PhD in Physics, now in code 500, Microcalorimeter case

There are a few university astrophysics departments that value instrument development, but for the most part, the tinkerers work at NASA or in relevant national labs (like NIST). In general, the NASA scientist who invents, or first adopts, the new technology concept will follow it and advocate for it through infusion on a mission. It is common for that scientist also to serve in a key role in the resulting project development (either as a PI, project scientist or technical lead, depending on the destination project and the tinkering/observing preferences of the scientist). This represents an important mechanism for continuity in the system.

2.2.3 Scientists as Administrators

The concept of peer-review is as much a tenet of the mission and proposal selection process as it is in the familiar journal review process. As a result, within NASA, NRA program element managers are typically scientists with PhDs in the relevant sub-discipline. Based on the four element managers interviewed in this study, scientists move permanently into this kind of administrative role for a number of reasons. For example, one was a former university professor whose area of interest had been de-emphasized, so the role of distributing funding seemed like a positive change. Another had joined NASA as a post doc, and when that was done, the only opportunities available were on the review side, so he took the position. While these program scientists do not tend to engage in original research anymore, their expertises seem to be respected by the other active scientists.

There are also part-time administrative roles that scientists take on, both inside and outside of NASA. Administrative in this context means tasks not directly related to advancing science. It includes tasks like proposal writing, but also occasionally serving as project scientists, or sitting on community review panels. The job of Project Scientist is similar to the function of Project

Manager, but where PMs control technical interfaces and mange the product, PSs are responsible for the science requirements, and making sure that mission objectives are met. Also, where becoming a Project Manager is a natural career progression for an engineer, taking on PS responsibilities is more of a temporary commitment that may happen several times over the course of a typical staff scientist's career.

Sitting on review panels (e.g., the decadal surveys discussed above) is another administrative responsibility and can involve a huge time commitment. The scientists see it as a community service –something that they should give back when they reach a certain point in their careers. The personnel overlap between observers and concept selection committees also serves as an important mechanism of information transfer in the system.

Some NASA-based observers see fostering new technologies as an administrative part of their job. "*It's the trade-off you make for not having to teach classes.*" Similarly, tinkerers work hard to protect their R&D time. As expressed by one Physics PhD:

> *As soon as you can't pay your own salary anymore you have to do something else and it's hard to get back… It would be very difficult to fire us, but what they could do is assign you work that you're not interested in... It's a struggle to be left alone, you have to constantly make that happen yourself, otherwise the institution will fall on you like a ton of bricks and you'll get nothing done.*

The general consensus is that, in any job, whether you have a NASA staff position or tenure at a university, about half of your time is taken up by "other" responsibilities. The difference is the content of that "other" half. Over the years, the relative esteem of professor vs. NASA/National Lab staff has evolved. Recently, NASA's director of science has taken steps to make the two seem equivalent. For example, now, NASA promotions are based on publication records. Some people believe this may drive NASA scientists away from the instrument contributions they have made. Regardless, the fact that NASA scientists often take jobs as faculty members at top universities later in their careers is evidence that both tracks are respected.

2.2.4 Engineers as Technologists

In the Instrument Systems and Technology Division (code 550), where most of this study is focused, there are some PhDs (both in Engineering and in Physics), but also many engineers and physicists with masters degrees. They do a significant amount of research, particularly on fabrication issues, and their roles are very similar to the instrument developer scientists. However, because of NASA's emphasis on science, there is an implicit hierarchy in the relative contributions, which is discussed in detail in the microcalorimeter case. As expressed by one long time detector systems branch employee:

> *The way things work here at Goddard, even for technology development you really need a scientific justification for it, so you always want to have a scientist on board with your proposals if you expect them to win. We've got a lot of great ideas, but if it doesn't have some kind of bearing on the science community then it's not going to win.*

NASA does do some in-house fabrication when the needs are extremely specialized and the production is small, but for the most part manufacturing is left to the industrial base. For instrument components, like detectors and mirrors (of interest in this study), which require narrow expertise, the contracting is typically with small, specialized business (e.g., the SB in the QWIPs case). The more traditional engineering functions associated with the large aerospace contractors (e.g., Boeing) become relevant in phase B of formal project development and later, and thus are of limited relevance to the present discussion.

2.2.5 Engineers as Managers

Where scientists can have very successful careers without ever taking on administrative or management responsibilities, management is a natural career progression for engineering professionals. Many of the technologists who play important (to our story) technologist roles early in their careers eventually transitioned to management roles (sometimes not entirely by choice). Management roles included leading an engineering branch or department; or managing a project or instrument. Some informants described their management roles as pure resource management "*my job is to find a way to keep my staff funded.*" Others see their new roles as an opportunity to run administrative interference and enable the scientists and technologists to do their jobs "*they may not even know that I'm the one who goes to bat for them in the funding meetings… I think it's a great effort and we need to keep getting them enough resources.*"

2.3 The Nature of the Technology: Implications for Structure

NASA's mission, to pioneer the future of scientific discovery, requires the continuous development of new satellite systems. In order to manage the inherent complexity of these systems, formal systems engineering processes have evolved over time. In the context of satellite engineering this has resulted in a standardized modular conceptualization of the system, with different disciplines responsible for different subsystems, and with interfaces rigidly controlled throughout. While the principles of satellite architecting are only of tangential interest to this research, their derived norms impose strong constraints on the ways in which scientists and engineers generate, and seek to incorporate, new technologies and concepts, making it impossible to ignore them completely.

This section discusses technology as structure in three ways. First, it focuses on the structure of the technology itself, giving a brief overview of a generic space science satellite architecture. Second, it discusses the way that technology maturity is assessed, giving an overview of the Technology Readiness Level (TRL) ladder. Finally, it describes the relationship between science needs and technology capabilities, linking the different feedback cycles in the system.

2.3.1 Technical Architecture

Satellite systems are typically conceptualized as a (fairly standard) bus and a (highly specific) payload or set of instruments, as illustrated in Figure 2-3. The bracketed numbers denote the NASA branch responsible for that module (corresponding to Figure 2-2). On the bus side, most of the development involves customization to the particular mission, and not innovation per se. The standard subsystems are: propulsion, TT&C (telemetry, tracking and command), structure, GNC (guidance, navigation and control), power and thermal (Wertz & Larson, 2007).

The payload breakdown is highly dependent on the function of the particular satellite. The payload typically includes collecting optics (e.g., mirror and baffle/sun shield), one or more detector arrays in the focal plane, and a separate cryogenic cooling system. In the remote sensing context (i.e., observing from afar) insights are gleaned from the light emitted by different sources. Payload instruments are classified by the wavelength (on the electromagnetic spectrum) and the nature of the measurement (e.g., imaging (position), spectroscopy (energy), polarimetery (polarization)). From a technology point of view, the classification is relevant because the energy and rate of incident photons (which correspond to wavelength and source) determine the types of possible materials and general characteristics of the instrument. Similarly, the nature of the measurement influences which detector development challenges dominate (i.e., for imaging, position resolution is key, while energy resolution dominates in spectroscopy).

Each of the narratives in Chapter 3 provide background specific to particular technology. Figure 2-3 presents an illustrative breakdown of the detector array, using the example of the microcalorimeter. A microcalorimeter is a class of X-ray detector that measures the energy of an incident photon as heat. Each individual microcalorimeter pixel is composed of an absorber, a thermometer and a weak thermal link connecting the detector to a heat sink (typically the thermally isolated backplane). Aspects of each of these components can be designed independently, but there are strong interdependencies among them as well. For example, superconducting transition edge sensors (TES) make much more sensitive thermometers than do ion-implanted silicon. As a result, where silicon-based semiconductor microcalorimeters were constrained to the low heat capacity Mercury Telluride (HgTe) for an absorber, TES microcalorimeters had a much larger head capacity budget, thus provided the option to electroplate the semi-metal Bismuth (Bi) for the absorber.

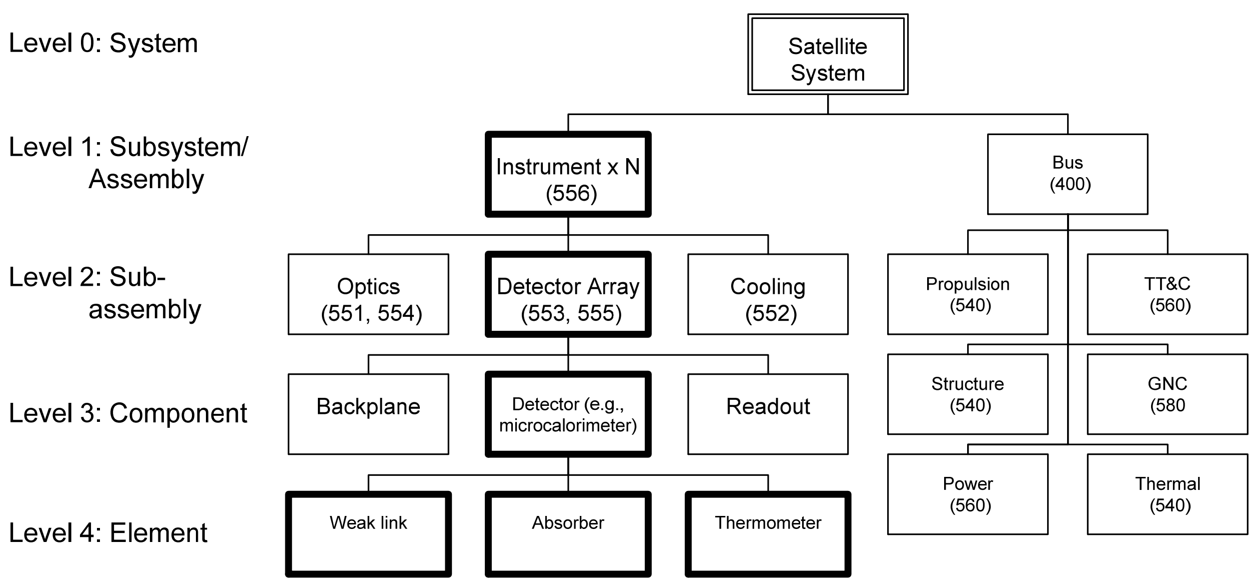

Figure 2-3: Illustrative Satellite Subsystem Breakdown

To understand the implications of the technological architecture for the shape of innovation pathways, it is useful to highlight where each of our innovations fit into the above breakdown. In the product management literature, Henderson and Clark (1990) differentiate between changes to the components and changes to the architecture (linkages among components) and their impact on the innovating firm. Component and Architectural innovations require different kinds of competences and resources. This concept can be equally applied at every level of integration in the satellite architecture. However, in practice, because of the ways that work packages and design authorities are managed, the technology innovations elicited in this study all occurred at the branch/laboratory level (i.e., level 3).

Mapping our six selected cases onto Figure 2-3, all but the CADR fit in the detector box. The CADR is a component of the instrument cooling system. It is worth emphasizing that this selection of innovation pathways is not a function of any choice to include only "component" innovations, or that the "components" we chose are simple. All of the innovations that were suggested to us were at level 3 of the system hierarchy. This is a natural byproduct of the modularity of the system; innovations tend to occur at the unit level. In this case, the branch (i.e., task unit) was level 3, as indicated by the codes. Yet, each had at least one technically meaningful level of further decomposition and multiple complex interactions among the sublevel(s), leaving room for major changes within these modules. In describing the levels of the innovation in Chapter 4, the following convention will be used: as viewed from level 3, activities at level 2 will be considered architectural innovation, while changes to elements (level 4) will be considered component or technology innovations.

The modularity of the system allows multiple improvements to be simultaneously incorporated at any level of the system architecture. For example, as long as the interface is controlled between the detector and readout, innovations can be made on each instrument subsystem independently. Connecting this back to Figure 2-1, there can, in theory, be multiple parallel technology developments merging into a project when it reaches "TRL 6" (defined below). The modularity of the system also requires that multiple disciplinary experts work together to make any major architectural change.

2.3.2 Measures of Maturity: Technology Readiness Levels (TRL)

The TRL ladder was introduced at NASA in the 1980s (Mankins, 1995) and has been widely adopted by other space agencies (including DoD and the European Space Agency). Few in the community believe that TRLs adequately capture the maturation process – e.g., the scale confounds architectural and component innovation (Sauser, Ramirez-Marquez, Magnaye, & Tan, 2008), and the levels are fungible; however, the extent to which they permeate the system structure makes it impossible to talk about innovation at NASA without defining TRL. Based on current definitions:

TRL 1: Basic principle observed and reported
TRL 2: Technology concept and/or application formulated
TRL 3: Analytical and experimental critical function and/or characteristic proof-of-concept
TRL 4: Component and/or breadboard validation in laboratory environment
TRL 5: Component and/or breadboard validation in relevant environment
TRL 6: System/subsystem model or prototype demonstration in a relevant environment (ground or space)
TRL 7: System prototype demonstration in a space environment
TRL 8: Actual system completed and "flight qualified" through test and demonstration (ground or space)

TRL 9: Actual system "flight proven" through successful mission operations

The TRL level of a particular technology nominally dictates what kind of funding it is eligible for and when it can be considered for infusion into a project (e.g., a project shall not baseline a technology that is less than TRL 6). However, as was illustrated in the QWIPs case, there is considerable definitional flexibility in practice.

2.3.3 Concept Selection

The link between science and technology is made through mission concepts. Concepts are typically formulated with respect to particular science questions, and are filtered up through the science community. The decadal survey process is a good illustration of this hierarchical filtering process for large missions. In the year leading up to a decadal survey, any individual or group is permitted to submit a white paper, either outlining a mission concept or advocating for a particular research question. These are reviewed and prioritized by subcommittees, and filtered up to content panels (e.g., star formation) as illustrated in Figure 2-4. The panels do their own prioritization and write panel reports making recommendations to the decadal survey committee (Science Working Group, or SWG). The committee hears presentations from advocates for particular mission concepts, and funds external reviews to get cost estimates. For an example of the Decadal Survey see (NRC, 2001).

While science priority is the dominant factor in deciding which concepts are selected, technological feasibility is acknowledged as a relevant (cost) constraint. In fact, lessons learned from JWST overruns in the last decade have motivated a change in the way missions are prioritized in the decadal survey process. In the past, proposers also included a cost estimate in their bids. Not surprisingly, these estimates were lowballed. For example, JWST, which was originally estimated at $700M, is now pegged at $6.5B, and members of the community expect the cost of JWST to increase before it is launched. As a result, the JWST program has eaten most of the available budget for the astrophysics division as a whole. In this decade's Decadal Survey, independent cost estimates were sought to mitigate against a JWST repeat. An implication of this change is that concepts which relied on major technological advances were penalized significantly in the cost department.

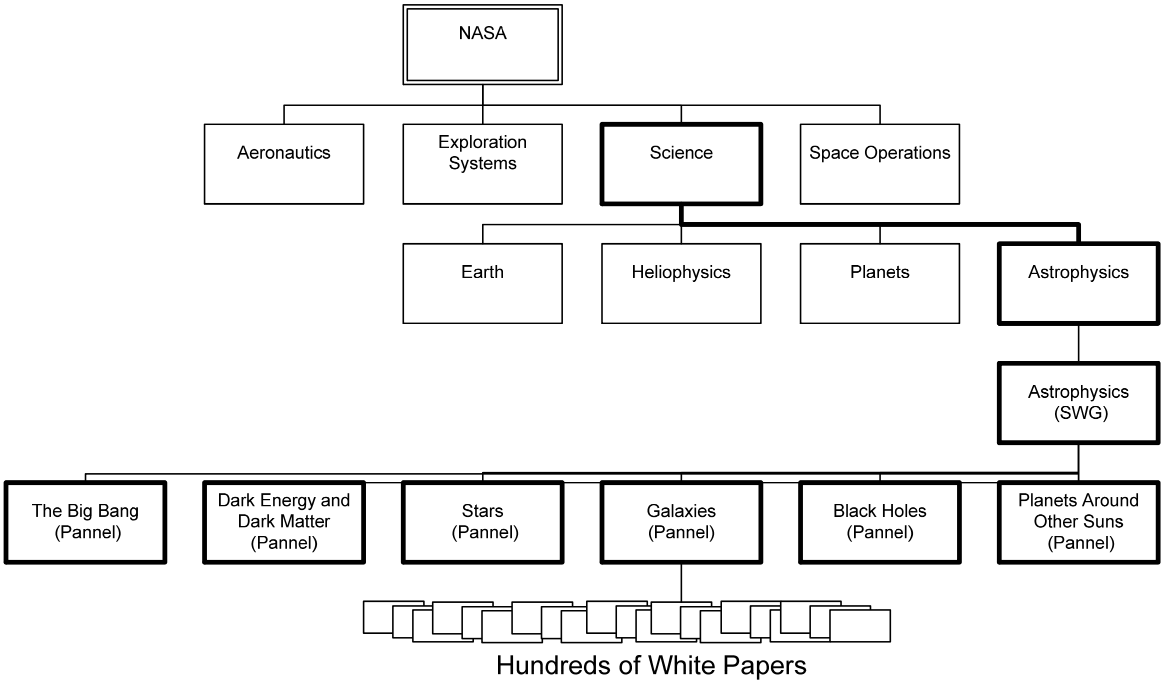

Figure 2-4: Science Concept Filtering Hierarchy

The link between concepts and technologies is captured to some extent through science and technology roadmaps which explicitly recognize that the sequence of missions impacts which domain specific technologies are developed, and when. Figure 2-5 and Figure 2-6 are reproduced from the most recent Astrophysics Science Plan (2005). They show the link between emerging technologies to science questions that they could enable the answering of, and also the infusion plan for technologies currently under development. Of the four classes of Astrophysics innovations studied in this dissertation, three are captured in the road map: Microcalorimeters, CZT (hard X-ray, gamma-ray detectors), and CADR (sub-Kelvin cooling technologies). X-ray polarimeters were not included because they had only been under development at NASA for 2 years at the time of the report's release.

This last point highlights a limitation of the roadmapping process in its ability to serve as a coordination mechanism between concepts and technology development. Specifically, a technology already needs to be known before it can be included. As a result, much of the early formulation happens through informal communication. For example, one of the lead scientists described the TES microcalorimeter co-formulation with Con-X as follows:

> *"[a colleague] who at the time was trying to get a mission concept together, said we need something with a spectral resolution about 2eV and collection area of […] can your calorimeter do that?"* [I89] Her initial response was a non-committal "*maybe? Theoretically it could…"*[I89] Recall that the theoretical limit worked out in 1982 by CSA#2 et al. was 1eV; however at the time, the best single-pixel resolution that had been achieved with any calorimeter was ~7eV. CSA#10 recalls that the discussion evolved through a series of informal gatherings, discussions in the conference room about what was possible and what it would take to get there. They concluded that 2eV was attainable, but would require a major investment in the new technology. "*Although* we *were already*

looking at the TESs as something that we should be developing, we hadn't gotten very far with them in those early days." [I89]

2.4 Synthesis

This chapter has described NASA's innovation system in terms of three interrelated tracks: (1) the organization of funding opportunities (section 2.1); (2) the formal and informal, roles and responsibilities, of individual actors (section 2.2); and (3) the structure imposed by the nature of the technology being developed, and science being investigated (section 2.3). Combined these tracks define the underlying systems architecture, or landscape, in which the innovation pathways unfold. In the remainder of the thesis, this representation will be used extensively to both structure the analysis and interpret the results.

Science Goals	Capability Goals*	Missions	Prime Strategic Technology Challenges
How do black holes form?	Measure X-ray photons with energy resolution >1000	Con-X	Detectors, Coolers, Telescopes
What happens at the event horizon?	Measure a 5-million-km baseline with an accuracy of 10 pm	LISA	Distributed Spacecraft
	Wide-field high-energy X-ray imaging	Black Hole Finder	Detectors
	0.1 microarcsecond X-ray interferometry	Black Hole Imager	Distributed Spacecraft,Telescopes
What powered the Big Bang?	1 uK√s CMB polarization sensitivity (100x better than WMAP)	Inflation Probe	Detectors, Coolers
	10,000x better strain sensitivity than LISA	Big Bang Observer	Distributed Spacecraft
What is the nature of dark energy?	Wide-field imaging with large optical/near-infrared focal planes	Dark Energy Probe	Detectors
How do galaxies form?	Photon-limited detectors with a diffraction-limited telescope	JWST SAFIR LUVO	Telescopes Detectors, Coolers
How do gas and dust form planetary systems?	100–100 milliarcsecond far-infrared interferometry at the photon-noise limit	Far-Infrared and Submillimeter Interferometer	Distributed Spacecraft, Detectors, Coolers
How common are terrestrial planets?	Measure a contrast ratio of 1 part in 10^{12}	TPF-C	Telescopes
	>50-m baseline infrared interferometry	TPF-I	
Is there life on planets outside our solar system?	Achieve 10^{-8} star nulling and angular resolution	Life Finder, Planet Imager	Distributed Spacecraft, Telescopes

*For competed Principal
discussions in this chapter.

Figure 2-5: Requirement for New Technologies to Answer Science Questions

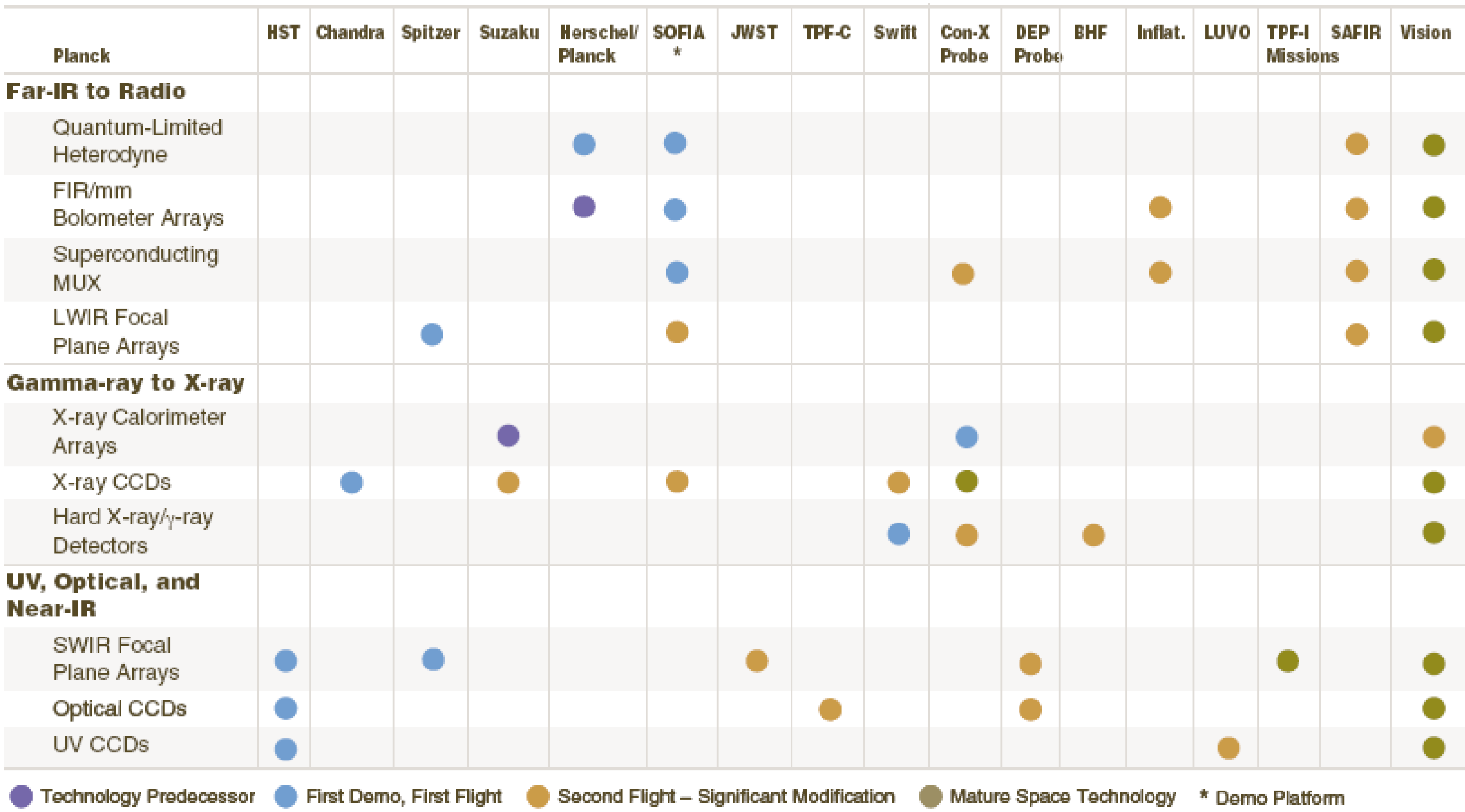

Figure 2-6: New Technology Infusion Strategy

3 Instances of Innovation Pathways

Having described the underlying systems architecture of NASA's innovation system in Chapter 2, this chapter presents the details of the empirical evidence which forms the basis for the process model articulated in Chapter 4. For each pathway, the data is presented in three forms: As an analytical chronology documenting an integrated explanatory story of the path taken; as a visual map which formally abstracts the event sequences along the dimensions of funding mechanisms, personnel participation, technical change and contextual factors; and as a transition plot of the event sequences highlighting the mechanisms that precipitated punctuations in the path evolution.

In the visual maps, the x-axis timescale is held constant for each map to highlight the spacing in history and also to clarify events that cut-across multiple cases (e.g., Bush VSE). The y-axis is cut into four meta-categories, the first three of which each correspond to a subsection of Chapter 2: funding, personnel and technical architecture.

Funding is further subdivided into branch overhead, internal R&D, NASA level R&A, suborbital and balloon programs, and flight specific funding. The particular funding mechanisms are indicated by ovals spanning their duration. The acronyms are all expanded in the background section (Ch 2.1) and the associated narratives. The color scheme should be interpreted as follows: white is a directly relevant funding source; grey denotes the funding of a relevant tangent (e.g., in the Si map, funding that is majority TES is in grey); black is an explicit branch (e.g., PAPA was not used to mature the CADR technology, but its existence helped keep the team alive.) Dotted outlines illustrate uncertainty.

The breakdown of the personnel categories also correspond to those described in the background section (Ch 2.2). The three scientist categories – administrators, observers and instrument developers – and the two engineer categories – manager and technologist are collapsed to four categories total. Administrator and Manager are merged. The arrows show the individuals' careers at NASA. The dotted lines represent time at NASA but not focused on the particular innovation. Non-NASA personnel are also included, in the category they most closely fit.

The correspondence of the technology sub-categories to the sections of Ch 2.3 is the least obvious. The division of component and architectural levels are separated as level 4 and 2 respectively per Figure 2-3. The particular components and architectures are labelled in words. Parallel lines with a black star on one illustrate multiple simultaneous development tracks which resulted in the selection of a particular concept. Thick lines illustrate multiple variations on the same effort, that can't be distinguished as clear alternative approaches. Major jumps in performance are indicated with steps circled in red, recorded at the level they were observed. In addition to component and architectural developments, major equipment acquisitions and mission-level context changes are also noted using similar symbols.

Context changes not captured in the other swim lanes are noted with labeled stars. Blue denotes a mission-level event; red, a major technical or scientific context change (e.g., the discovery of GRB afterglow fundamentally shifted the mission context); and green is a policy change. The context categories roughly correspond to shocks as defined in Chapter 4. Finally, the characteristic epochs (formalized in Chapter 4) are demarcated by shades of grey in the background, bounded by red initiation and termination lines. In order of increasing darkness, the greys are technology exploration, architectural exploration, exploitation, branching and treading water. When technology exploration is overlaid (and shorter) than an epoch of architectural exploration, it indicates that the level dropped to address a particular component level issue. Similarly, branching can be overlaid on, and extend beyond an epoch of treading water.

Where epochs are relatively straightforward to identify, shocks are often treated as random, unforeseeable, exogenous disturbances. In order to identify patterns of shock sequences and common transitions, an NxN matrix was prepared for each case (Table 3-1 through Table 3-5). Epochs are listed on both the horizontal and vertical axes. Each cell records the events that lead to a transition from the row epoch to the column epoch.

The following sections are organized by innovation pathway, with the three views presented for each.

3.1 The Continuous Adiabatic Demagnetization Refrigerator (CADR)

This pathway describes the development of a Continuous Adiabatic Demagnetization Refrigerator (CADR) within NASA's Goddard Space Flight Center's (GSFC, or Goddard) cryogenics and fluids branch. After more than a decade of concerted development efforts, the team is still awaiting a mission opportunity on which to fly the full system. Like the traditional ADR, the CADR will maintain spacecraft instruments at the required mili-Kelvin operating temperatures. Unlike the incumbent, the CADR operates continuously (eliminating the need to pause observation time while the cooling system recycles) and provides significant mass reductions for a fixed cooling power requirement. These advantages will become increasingly important as future X-ray, IR and submilimeter observatories incorporate larger and more advanced detector arrays.

Gestation period

Although explicit CADR development didn't begin until 1998, relevant groundwork was laid much earlier. The magnetocaloric effect, upon which the technology is based, was first observed in 1880, and applied to develop the first magnetic refrigerator in 1933. Goddard first became interested in cryogenics in the late 70s when it became clear that future missions (including IR, X-ray and submilimeter) would require detectors to be cooled below 1K. At the time, in a seminal position paper, Dr. Stephen Castles, the head of the then newly created Cryogenics branch, argued that ADRs were the most appropriate approach to cryogenic cooling for space applications and that Goddard should develop them as an in-house core competency.[D19] Through the 1980s, the branch let a sequence of SBIR contracts to develop the requisite extremely powerful magnets, and received on the order of $8M per year from "code-R" (the historical NASA technology directorate) to develop salt pills and, to a lesser extent, heat switches – the other critical components.[16] [I24]

[16] The different book keeping standards at the time, limit the comparability of expenditures.

The first flight ADR was developed in support of the Chandra X-ray observatory; specifically to cool the XRS instrument [I30, 46]. However after nearly $3M of further development funding, [I46] the XRS instrument was demanifested from Chandra and reincarnated as the Goddard-furnished XRS instrument on the ill-fated Japanese X-ray observatory Astro-E (which was subsequently lost due to a launch failure).[17] [D20] Nonetheless, by 1996, when CSE#2, a low temperature physicist PhD by training, joined Goddard, ADRs had become the industry standard, being developed for multiple smaller missions, one of which he was working on when the idea struck.

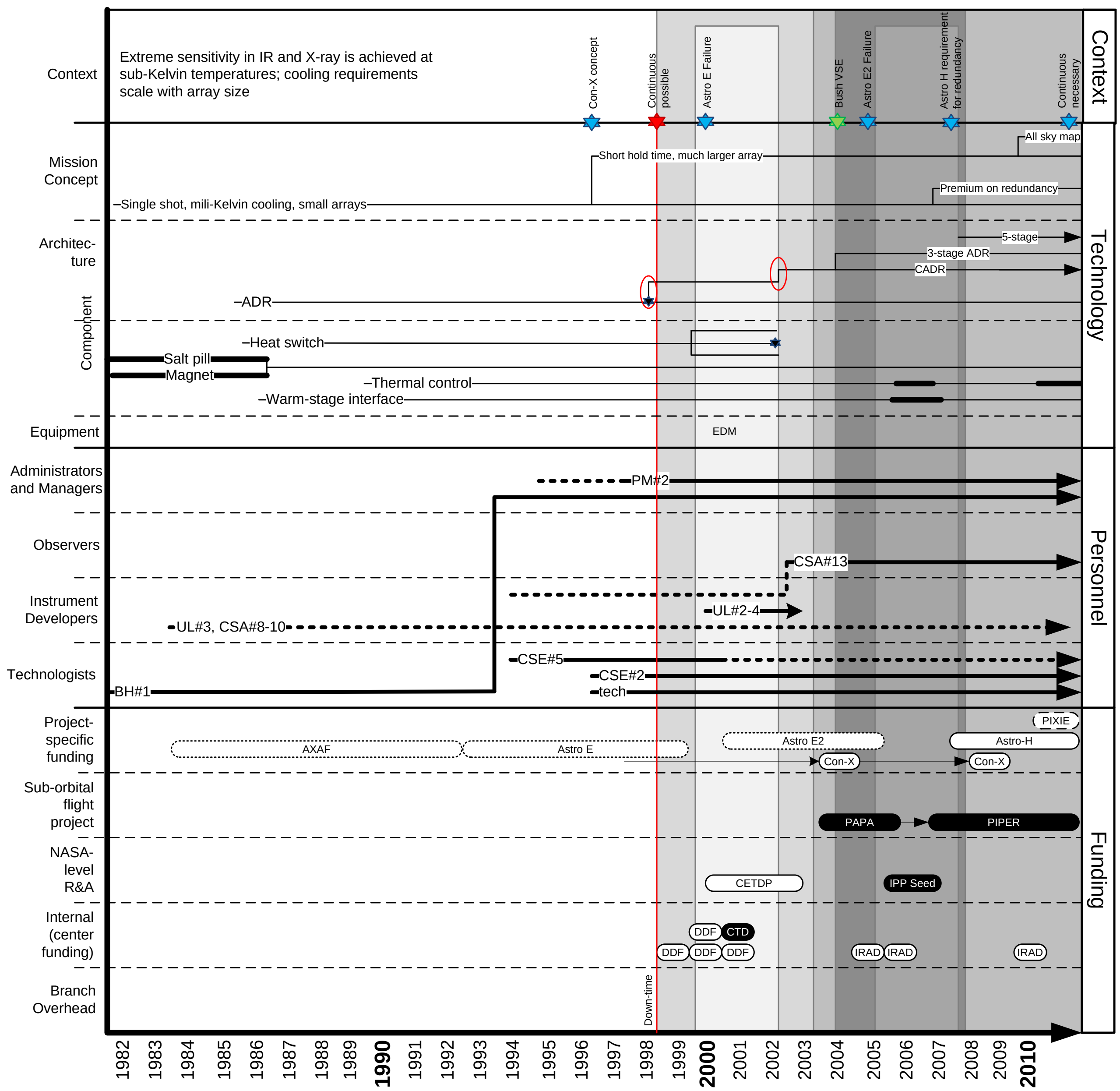

Figure 3-1: CADR Visual Map

[17] The follow-on Astro E2 vented its stored cryogens shortly after launch, so it's impossible to confirm instrument performance.

In spring 1998, with the next generation of X-ray observatory prominently on the horizon, the cryo branch was faced with the realization that the standard approach to ADR development was unsustainable – based on trends in mission hold time and duty cycle requirements, ADRs would soon be prohibitively massive for space applications – CSE#2 and his office mate CSE#5, another low temperature physicist, began brainstorming alternative approaches. [I30, 70] As recalls CSE#2 "*one day the idea just came to me:*" [I45] The idea was an architectural innovation in nature – it could be achieved with the same components already being employed on the Adiabatic Demagnetization Refrigerator (ADR) he was working on at the time. It was elegant in its simplicity; rather than operating a single stage ADR over its full range and then waiting to recycle it (i.e., remagnetize and demagnetize), in principle, with enough cascading temperature stages, the coldest stage could be kept continuously cold.

Architectural exploration

The advantages of continuity were clear – up to 25% more observation time on any given mission. [I45, I70] So after running the idea by his office mate without surfacing any show-stoppers, the two began exploring the idea in more depth. After a summer of "*playing with*" simulations and "*messing around*" a little in the lab in their spare time[18] the two were sufficiently convinced of the promise and feasibility of the idea to seek out formal development funding. [I30] In fall of 1998, they initially applied for internal R&D funding in the form of DDF (Directors Discretionary Funding) [I30, D1]. The proposal was accepted [I45]. Although modest, this $65K for the first year was sufficient to begin exploring the critical element of the concept: Heat transfer between the two coldest stages of the cascade. [I45, D2]

Technology exploration: innovating to solve an identified problem

It became quickly apparent that several major component level technical hurdles would need to be overcome in order to realize the design concept. As stated in a 1999 grant proposal:

> *Success basically hinges on developing heat switches that can conduct heat very well in the on state at low temperature, yet provide good isolation in the off state. Several switches are needed to span the temperature range from 20-30 mK up to 10 K.* [D2]

Even as the original DDF was underway, CSE#2 and 5 sought out the more substantial technology development funding they would need to mature the capability to a point where it could be picked up by flight missions. The next logical step was to apply for "level 2" funding administered by HQ (called NRA – NASA Research Announcement); an application was submitted and subsequently rejected. [I30, D5]

The rejection was not a surprise to CSE#2:

> *Our feeling at the time, was that the way these technology calls got structured, HQ knew of technologies that were up and running, ready to be funded, and they kind of craft the NRAs around those, so in the NRA might solicit cooling technologies with capabilities*

[18] ADR development is a time intensive process replete with "down time." CSE#1's branch head was happy from him to use this spare time to explore what she agreed was a promising new idea.

> *that they've heard through the grape vine might be proposed to this NRA. We didn't fit into that description, so it wasn't surprising that we weren't chosen.* [I30]

In order to continue the development effort, CSE#2, 5 and two other staff from the cryogenics branch applied for several funding paths simultaneously. [D2, 3] In 2000, they received 2 year long DDFs and pitched another level 2 grant. One of the DDFs was dedicated to heat switch development, pursuing multiple approaches to achieve the required performance. [I45] In fact three summer students were each given a different strategy to explore. One of these paths was successful – a gas-gap heat switch – for which a patent was eventually issued in 2005. [I30, D16]

Returning to the architectural level

With the second DDF, they set out to prove-out the ability to make the next temperature transition up, while preparing for the next NRA opportunity. The next NRA solicitation was released under the newly created CETDP (Cross Enterprise Technology Development Program), which was a much better fit for the CADR development, "*lo and behold, the technology descriptions included "continuous" refrigeration systems for 50 mK and below (an exact fit).*" [I30] The CETDP grant was awarded – 3 years totaling $1.9M (including civil servant labor) [D6]. According to CSE#2, no active efforts were made to encourage HQ to release a solicitation targeted at their ongoing work; he believes that the previous year's failed bid may have communicated the promising technology and need for it [I30, 45].

During those three years of CETDP funding, substantial progress was made along an increasingly clearer development trajectory. They progressed from a 2-stage prototype, to a 4-stage system operating continuously at 50 mK that could dump heat to a 4K He bath (parameters suitable for flight missions then in the concept stage). [I30, 45, D12, 17, 18] In addition to the technology-centric advances, the stability and flexibility of the CETDP funding enabled the group to develop important tacit competencies. CSE#2 hired and trained several students, whose assistance contributed significantly to the development [I45]. Further, he used part of the funding (as well as some "cobbled together" resources) to purchase an electric discharge machine (EDM). CSE#2 believes that the EDM, and the technician who became an expert in its use, may be the single greatest explanation for their current status as the world leaders in ADR technologies.

During the same period, the team also received small amounts of Commercial Technology Development (CTD) funding - $25K in each of 2001 and 2002 [D21] – to "*augment*" their other work and explore small CADR systems for lab applications. [I45]

Exploitation

At the end of the 3-year CETDP there were only two, albeit resource intensive, technical issues remaining before the CADR system could be considered "TRL 6" (i.e., ready for flight project-specific development). No structural analysis, or vibration testing, had been undertaken and thermal stability needed improvement.[19] Now in 2003, the team was confident that these remaining challenges could be overcome under an additional 3-year CETDP which seemed

[19] ADRs are normally stable while cold because nothing is changing. However, in a CADR, the temperature cycling creates fluctuations that needed to be actively controlled; a capability that needed to be developed from scratch.

imminent [I30, 45, 70]. At his end of year project review, CSE#2 had connected with the gentleman managing the CETDPs (who happened to be a fellow UIUC alum) and been given the impression that he was in a good position to receive a follow-on grant as soon as the FY2004 program funding was approved [I45].

Treading Water and Branching Out

However the funding was never approved. In fact, 75% of NASA's technology development funding was cancelled that year and reallocated to support the Constellation program. This left the team with a capability that was too mature to be suitable for the early-stage seed-funding that was still available, yet not mature enough to be taken-up by a flight project. The four years that followed are sardonically referred to by the group as the "*dark ages*" [I45, 46].

The funding drought was not confined to the CADR development (and its place in the valley of death); R&D funding was tight across the board. [I24] The cuts to intramural R&D funding coincided with the roll-out of full cost accounting. Where civil servant labor was previously paid out of generic overhead monies, under full cost accounting, time worked must be book-kept in relation to specific projects. This relatively minor administrative change had important implications for how the branch could operate. Where the branch had previously reserved 1 FTE for interesting, but not-yet-fundable concept exploration, this was no longer feasible when FTEs became "*real money*." Similarly, the early stage DDF of $75K went much farther, when it came with effectively unlimited[20] labor [I46, 74].

During this period, the CADR push stayed alive (despite suggestions that it be temporarily put on hold) due to fund-finding ingenuity, and a little begging, on the part of CSE#2 [I45, 46]. The meaning of "stayed alive" merits some clarification in this context. They – neither the branch head (BH#1) nor the champion CSE#2 – were ever concerned that the technical capability would become obsolete [I45, 46]. All of the developments thus far were also relevant to the traditional ADR design; and had distinguished the group as world leaders in the area[21] through their application, in stages, to every next flight project. Nor did they ever question whether an operational system could be developed once R&D funding was restored [I45, 46]. According to the BH#1, her suggestion that the project be temporarily *mothballed* was for purely financial reasons; her job was to keep her staff funded, and money was extremely tight [I46]. She assigned all her senior staff to write proposals and find ways to insert their expertise into the few flight projects that had money (mainly JWST, the decade's flagship IR telescope). However, she recognized that every hour worked on JWST was a threat to the ADR competency; unlike the technology, which would *keep* for several years, the tacit knowledge stored in the people who worked on it wouldn't [I45]. And, once an individual has transitioned to a new project, especially an important flight project, it was nearly impossible to staff them back to R&D.

However, while the branch head saw this as a necessary evil (from her staffing responsibility perspective), [I46] CSE#2 saw this as a potential show-stopper to be actively combated (from his championing perspective). In particular, he was worried about losing one key technician – the

[20] Many of the scientists and technologists worked "night and day" (in their free time) on the pet projects that inspired them.

[21] Their ADR design is several times smaller, more efficient, less massive and power intensive, and even costs less to manufacture than the competition.

expert in electric discharge machining – who was "*the kind of guy who would rather retire and work on his motorcycle*" [I30] than transition to another project while waiting for CADR funding to be restored. And rebuilding that kind of expertise would have taken a very long time. So, the CSE#2 found just enough funding to keep the project *alive*.

Between 2004 and 2009, the funding came in the form of two IRADs (Internal Research and Development – the later incarnation of DDF) of $135K and $100K in each of 2005 and 2006, [I30, D7, 8] yearly project support from Con-X (the next big X-ray observatory) development funds of about $25-100K per year [I30, 57], the equivalent of about $175K (+ matched funding of $275K) from an IPP seed fund partnership [D10], and on the order of $50K for the development of a 3-stage continuous ADR for an IR balloon instrument called PAPA and later PIPER [I79]. Of the four funding types during this era, the use of the IPP seed fund was the clearest. It enabled collaboration between the Goddard team and sections of the relevant industry. Of the ~$450K, $150K was contributed as matched funding by a firm specializing in low temperature read-out electronics. Together they investigated control circuits for the low temperature stability issue discussed above. Another $50K came from one of the key industrial manufacturers of space qualified cryo-coolers. This allowed the team to explore the interface with the type of mechanical refrigerator that would be used in future missions [I45].

The $50K from the balloon program, though a monetary token, was critical from the point of view of keeping the manufacturing team working [I76]. Towards the end of 2003, CSA#13, an IR Astronomer who had previous experience working closely with the cryogenics branch, approached CSE#2 to build a CADR for his balloon experiment [I79]. CSE#2 was of course more than happy to do so. Although PAPA was never completed in the three years allocated, it was later transitioned to PIPER, a follow-on five year balloon instrument program. PIPER will fly in 2012 with a 3-stage CADR cooling its detectors [I76]. To date, the engineering model has been used extensively to test PIPER's optical system. While the development of a CADR for a balloon flight does not directly contribute to the flight maturity goal (due to stark differences in operating environments), in addition to maintaining expertise within the group, it gave further credibility to the operational concept [I76].

At $25-100K per year, the Con-X funding was more a way of sanctifying a relationship with a project, and augmenting other resources, rather than an avenue to further the technology development efforts [I45, 75]. To understand the project's perspective on the relationship requires some brief background about Con-X (now reincarnated as the International X-ray Observatory IXO). The Con-X program first received technology development funding in 1998 leading up to their 2000 decadal survey bid [I75, D22]. They let an NRA that year, soliciting proposals primarily in the mirror and calorimeter area (the Cryogenics branch was part of Goddard's bid for the calorimeter instrument). The Con-X mission ranked second (to JWST) in the Astrophysics Decadal Survey for 2000. As a result, while the mission did not receive approval, a mission study was directed to Goddard which included an average of $6-10M a year in technology development funding over the last decade [I57]. That money, while subject to significant variance related to budget uncertainty in the JWST program, has allowed them to make significant progress in both detector and mirror technologies that will enable this extremely ambitious undertaking, should it rank first in the 2010 decadal survey. While continuous cooling to cryogenic temperatures is not a critical mission enabler, the advantages in terms of potential science return are compelling (as an extremely positive bonus) [I57, 68].

Thus, from the Con-X project's point of view, maintaining a relationship with CSE#2 over the years has been well worth the non-competed trickle of matching discretionary funds they have provided [I57]. Further, in return for the funding, CSE#2 has supported the project on multiple occasions by preparing progress updates (for the numerous reviews that a directed study program is subjected to) and sitting on expert review boards.

Changing Context

Also during the "dark ages" an unfortunate set of events in JAXA's Astro program created an opportunity to demonstrate the CADR component capabilities that had already been developed. Recall from above, that the first Astro spacecraft, designated "E" was destroyed due to a launch failure in 2000. The second Astro spacecraft, a direct copy designated "E2" was launched successfully in 2005, but due to an overlooked design problem, the stored cryogens were vented from the Dewar shortly after launch, resulting in a loss of cooling capability (rendering the cryogenic instruments useless) [D20]. With the next generation Astro-H already in the conceptual design phase, system redundancy became an important selling point [I30, 78] – the Japanese government would not tolerate another embarrassment. To achieve redundancy in the cooling subsystem, a 2-stage ADR design was baselined in the initial proposal (2008) that could be mated to both a 1.3K liquid He Dewar and a 1.7K JT cooler (in case the He bath failed again) [D23, I45, 78]. The two stage design was approved (despite being unproven technology) since even at 1.7K, the magnet that would be required for a single stage ADR was prohibitively large.

The 2-stage became a 3-stage ADR, capable of operating at 5K, a year later when, following i) challenges faced while space qualifying the then baselined state-of-the-art 1.7K mechanical cryo-cooler and ii) successes with mating the multi-stage ADR system to a warmer cooler (as demonstrated through the IPP seed funding). Multiple technical solutions were considered for filling the 1.7 – 5K cooling gap; however, the 3-stage ADR system prevailed as the lowest cost and risk solution [I78]. The additional qualification and production only cost the program an extra $750K [I78]. From the perspective of the project, this solution served to accomplish the desired redundancy; from the perspective of CADR development, it provided an opportunity to flight qualify many of the components that the ADR team had developed to support a continuous (multi-stage) ADR. Although the Astro-H team was reticent at first, to accept the risk associated with flying an unproven technology, necessity prevailed [I30, I78]; besides "*CSE#2 is very persuasive*" [I46].

Flight oriented development: exploitation

More than just creating a flight opportunity for some critical pieces of the new CADR technology, developing the 3-stage ADR for Astro-H gave CSE#2 and his colleagues a relevant "day job" again [I30]. Now, a significant amount of the time he wants to be spending on furthering the CADR can be justified as relevant to a chargeable project. Further, having several important pieces fly on Astro-H, gives credibility in terms of risk reduction, and continues the path of the CADR towards the goal of TRL 6 [I45, 79]. In fact, as the ramp-up towards near-term project relevance increases, CSE#2 has received an additional IRAD funding to investigate the remaining low temperature stability concerns [D9]; and once achieved, that, plus the vibration testing that will come through the Astro-H program will yield an effective TRL 6 [I30].

The importance of TRL 6 is that it allows non-flagship missions (i.e., other than IXO) to consider CADRs in their baseline. This is relevant because where IXO can do without the continuous capability [I68, 75], for ASP (a MidEX, Absolute Spectrum Polarimeter) for example, which is a scanning mission, the continuity is critical [I30, 79]. Yet, as a mid level Explorer, ASP can't baseline risky supporting technologies. There is a race therefore, to get the CADR to TRL 6, in time for ASP to happily bring it the rest of the way. Incidentally, the project scientist for ASP is CSA#13 (from PAPA and PIPER); he is confident in the technology and the team [I79]. As of writing (early 2010) the CADR is baselined in the study-phase and all the components and control schemes that will be needed for the 5-stage ASP CADR are being demonstrated under a current IRAD. The TRL 6 goal is looking obtainable. And, if IXO is ranked first in the 2010s decadal survey, there will be another flight opportunity there as well.

Continued branching out

These flight opportunities are not yet guaranteed. Thus, in recent years, as another way to buy-down future flight risk, and increase the potential for flight opportunities, the cryogenics branch has continued to pursue a number of ongoing R&D efforts which build on, and feedback into, the continuous/multi-stage ADR paradigm. Specifically, ADR stages capable of working up to 30K have been under development for some time [D24]; this represents the less efficient region of mechanical cryo-cooler operation. However, if ADRs are to become feasible in this region, major innovations in the area of superconducting magnets will be required [I76]. To this end, the group has supervised a series of SBIR contracts with two small businesses. It turns out that the same techniques driving improvements in high temperature magnets, can be used to improve the low-current leads bringing power to the ADR magnet; this technology will be employed on Astro-H. And so, the process continues.

Table 3-1: CADR Shock Sequence

Epoch	Start	Component Exploration	Architectural Exploration	Exploitation	Treading Water and Branching Out	End
Start			*IXO requirements come out in 1998 (much more cooling power/mas) *CSE#2 has architectural insight (multiple ADRs => CADRs)			
Component Exploration			*New gas-gap heat switch invented			
Architectural Exploration		*Initial prototype reveals heat switch challenge		* working prototype => clear development path		
Exploit					* Bush VSE => cut R&D funding * Con-X delayed * Other projects drawing resources	* If PIXIE is selected as MidEX, requirement for CADR
Treading Water and Branching Out	* lab applications			* Sequence of Astro failures creates emphasis on redundancy (specifically requirement for ADR to mate to both He tank and mechanical refrigerator) * Mission requreing CADR enter planning phase		
End						

3.2 The Cadmium Zinc Telluride (CZT) Detector Array

This pathway describes the development of soft gamma-ray/hard x-ray detector arrays at NASA's Goddard Space Flight Center (GSFC, or Goddard), through a collaboration between the gamma-ray spectroscopy group and the detector branch. The Cadmium Zinc Telluride (CdZnTe, or CZT) detectors which resulted from this decade-long development were first flown on the wildly successful SWIFT mission in 2004, and have since enabled a whole class of imaging applications in the hard X-ray/soft gamma-ray range. The CZT detectors filled a long recognized need for a room temperature semiconductor capable of high resolution imaging and spectroscopy in this particular waveband.

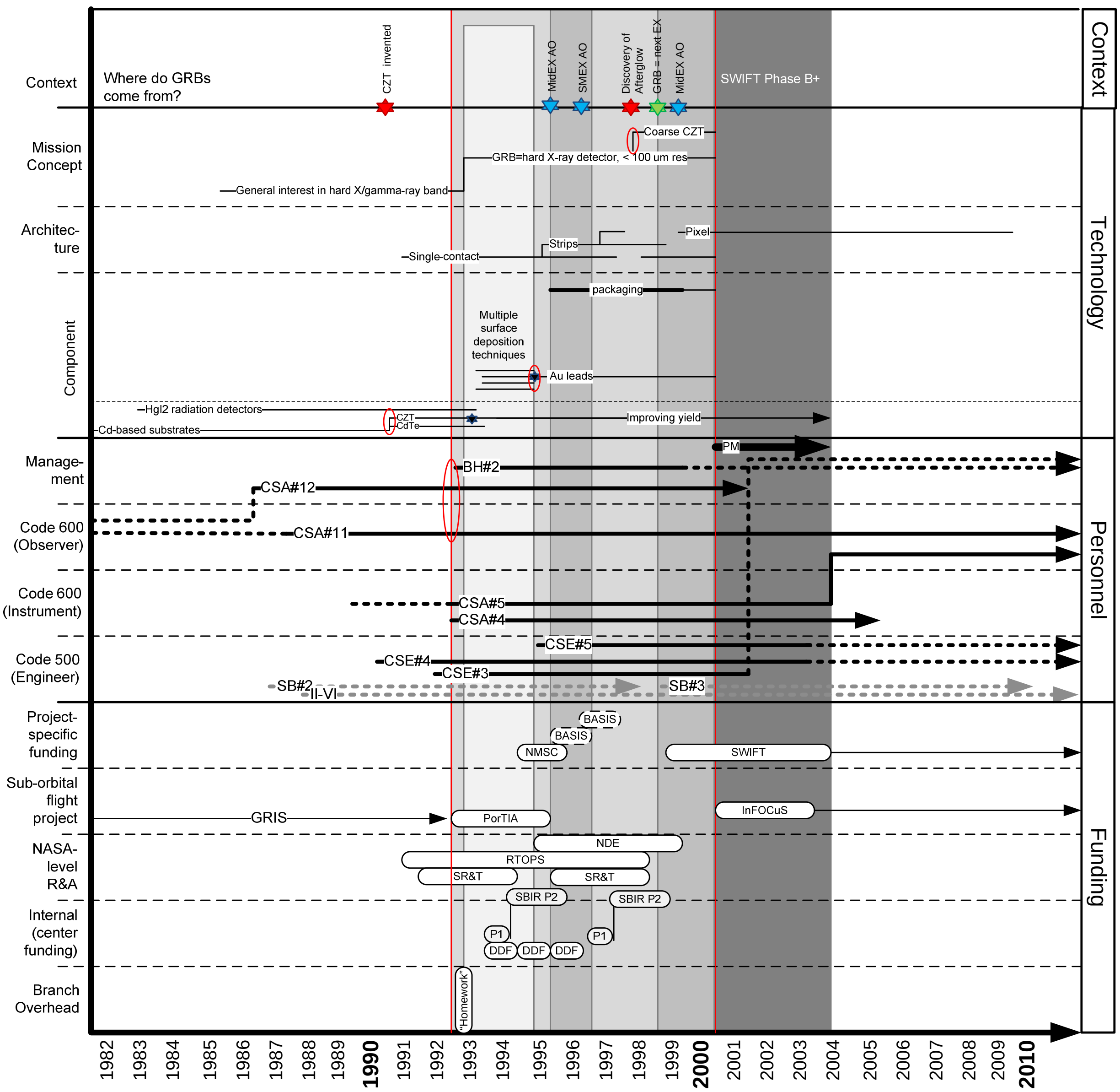

Figure 3-2: CZT Visual Map

Gestation Period (pre-1993 collaboration)

Although the Goddard gamma-ray spectroscopy group and the detector branch didn't begin collaborating on a CZT development program until 1993, relevant roots of the innovation pathway trace much farther back. In the early 1990s, there were very few instruments covering the hard x-ray / low-energy gamma-ray spectral band (5-500 keV) in existence. This gap was due more to the technical challenges associated with making detectors in this range, than a lack of scientific interest; however challenges of the former limited numbers of the latter [I73]. The Goddard scientists were hoping to fill that gap by branching out into this new energy band, which they believed held the key to understanding such diverse phenomena as:

- *element creation, explosion dynamics and event rates for supernovae*
- *the origin of gamma-ray bursts through sensitive searches for cyclotron and other lines*
- *physical conditions in the vicinity of neutron star surfaces through observations of cyclotron and other lines in x-ray pulsars*
- *physical conditions in the central engines of AGN* [D1]

At the time, the default hard X-ray spectroscopy detectors were Ge:Li (Germanium doped with Lithium). These detectors had relatively low Z (which meant that the detectors would need to be very thick to create sufficient stopping power at the relevant energy band) and needed to be cryogenically cooled (requiring the accompaniment of the mass and complexity associated with advanced cooling apparatus, effectively rendering explorer class missions out of reach) [I52, 53]. Recognizing that suitable detector technology was a prerequisite to achieving their objectives, the group had been "*on the lookout*" [I73] for a suitable detector for some time. Being "on the lookout" in this context involved regular attendance at the conferences where researchers presented new developments in the realm of high Z, room-temperature solid state detectors [I73].

By the late 1980s, the scientists began to hear about two new semiconductor compounds – Mercuric Iodide (Hg I_2) and Cadmium Telluride (CdTe) – which, following nearly two decades of research since their discovery in 1970/1971 were beginning to show promise as radiation detectors [D37, I80]. The former development had been heavily funded by the Department of Energy in the context of nuclear monitoring [D1]. The latter was under development by smaller companies interested in applications to medical imaging [I80]. Both materials offered the promise of room-temperature semiconductor gamma-ray detection, but each presented significant technical challenges which limited their practical utility. By the late 80s, when the scientists were monitoring developments, HgI_2 seemed the leading alternative [D8, D1].

Thus, in 1991, the scientists put in an SR&T (Sustaining Research and Technology, now called NASA Research Activity or NRA) proposal to study the application of HgI_2 devices at NASA. They won the SR&T bid as well as a follow-on grant to investigate the relative detector performance characteristics of different state-of-the-art detectors using balloon-based platforms (including the cryo-cooled baseline Ge:Li, room temperature HgI_2 and alloys of CdTe and Zn) [D1, I52]. For the second grant, a young astrophysicist (CSA#4) was hired as an NRC post doc, specifically to investigate room temperature gamma-ray detectors. It was during preparation of that proposal that the Director of Space Sciences (CSA#12) recommended that the scientists collaborate with Goddard's detector branch and made the introductions. The director had learned the value of this type of interdisciplinary collaboration through prior work with the detector branch head (primarily in the context of microcalorimeter development for AXAF (Advanced X-ray Astrophysics Facility) – see microcalorimeter case).

Project Initiation (A phone call and an executive decision)

In a sense, NASA's decade-long development of CdZnTe gamma-ray detector technology was initiated by a phone call from above. One afternoon in 1993, Goddard's detector branch head got an unusual call. It was from the Director of Space Science at Goddard. His request was for the detector branch to support the high energy physics group in developing room temperature gamma-ray detectors, an area of science he believed would be critically important in the future. Based on preliminary investigation by Goddard's science community, two potential candidate detector materials seemed promising: HgI_2 and CdZnTe. The director expressed no preference between them. As recalls the branch head (BH#2), "*this was not a guy you said no to.*" So, following the conversation, the branch head returned to his desk and "*did his homework*" on the options [I59].

Interestingly, the Director of Space Science doesn't remember that conversation as having been particularly unusual or important [I74]. While it wasn't common for a science director to call an engineering branch head, it was an important part of what the director saw as his job. As he viewed the situation: "*either I could give equal, uniform blanket approval to everything that came out of my organization; or I could demonstrate some preference for some things that I thought were more promising.*" He chose the latter approach realizing that "*I could be much more effective if I occasionally (not all the time) demonstrated passion (or anger)… this was one case where [BH#2] was being pulled [in other directions] so I thought it was important to let him know how important I thought this was*" [I74]. Although the director had no formal influence over the engineering directorate, he found that these "*demonstrated preferences*" served to guide the level of effort, provided by engineering to support proposals, and projects.

This case was no different. Within a couple of weeks of "*homework*", the branch head had (independently) come to the conclusion that CZT was the most viable path forward [I59]. In the time since the scientists' assessment of the state of the field, a new material growing process had been developed for undoped CdTe as well as the invention of a new semiconductor compound CdZnTe [D9, D10]. These two advances changed the relative assessment of the device options. Based on the updated information, the branch head chose CZT for two reasons. First, HgI_2 is a "*miserable material to work with*," [I59] particularly from a health and safety point of view. To work with it in their fabrication lab, many expensive changes would have had to have been made. CZT didn't pose these concerns. Second, the main group working on HgI_2 was located in Israel. While this did not preclude collaboration, it certainly would have made it more challenging. Expertises in CZT processing were domestic. Combined, these factors made the branch head's decision to pursue CZT detectors a straightforward one [I59]. Thus, a few short months into NASA's development, a major potential technological trajectory had already been pruned. HgI_2 was pursued in the broader community and was later found to have other performance-oriented challenges compared to CZT/CdTe (which have since emerged as the dominant technologies) [I48].

Goal Guided Exploration

Almost from the beginning, the science goal of finding the origin of Gamma Ray Bursts guided the technology development. GRBs are the most luminous electromagnetic events occurring in the universe, and for several decades represented a major mystery in the field of high energy astrophysics. Originally discovered by a satellite designed to monitor for covert nuclear weapons tests in space in the 1960s,[22] and declassified in 1973, these gamma-ray bursts of "cosmic origin" became a topic of much scientific speculation. No one knew what they were or where they were coming from. Even accurately positioning these short-lived (~1 second) high energy bursts, which seemed to crop up at random locations, posed a major challenge: Since, unlike light at less-energetic wave-lengths, gamma-rays cannot be bent (and focused), the detector area needs to be as large as the desired imaging area [D39]. Also, since scientists were unable to predict where the next burst would come from, this imaging area needed to be wide [D29].

This science need translated into a requirement for a wide area detector plane with sub 100 µm position resolution [D13]. The intention was to procure CZT wafers from industry; have the detector branch pattern fine contact structures, wire bond leads and package the detectors as a suitably large array, per the design of the science team, as shown in Figure 3-3. However, in order to accomplish this, an interrelated set of major technical challenges needed to be overcome at each level of manufacturing and integration.

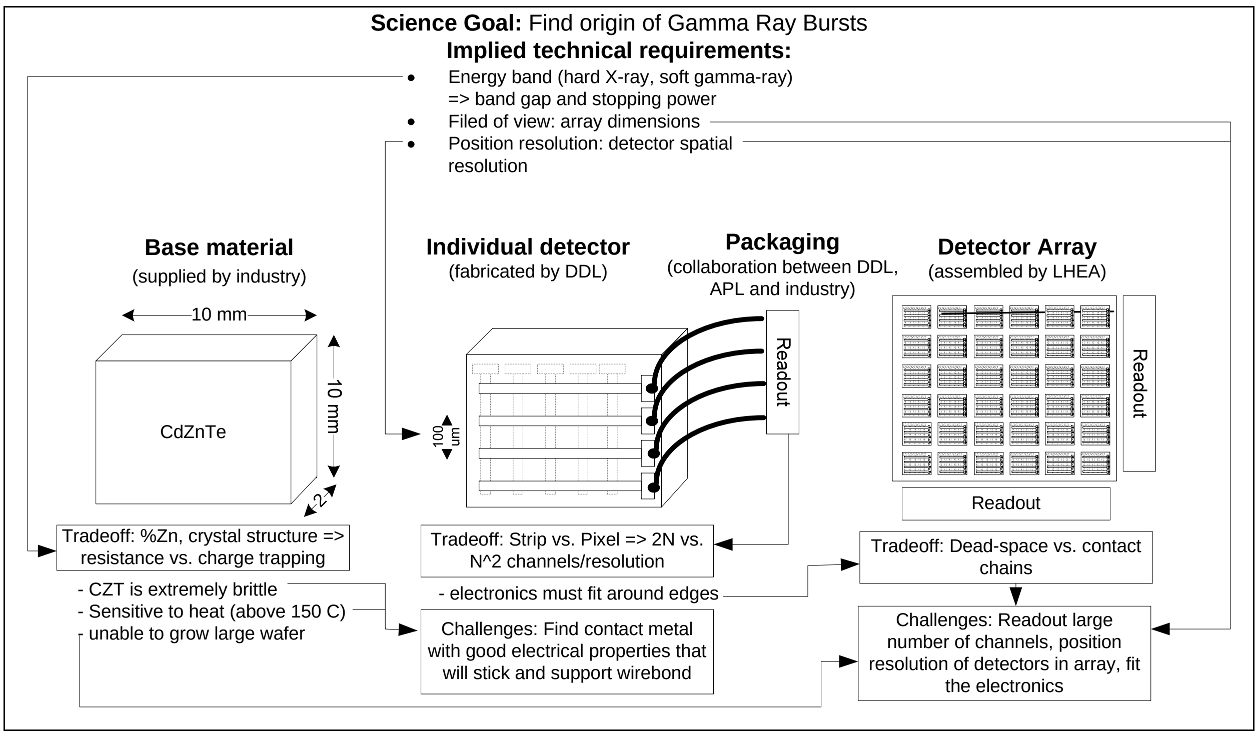

Figure 3-3: Flow of Design Tradeoffs in the CZT Innovation Pathway.
Acronyms defined: DDL (Detector Development Lab); APL (Applied Physics Lab); LHEA (Laboratory for High Energy Astrophysics).

Although the semiconductor compound CdTe was invented in 1970 as a room-temperature radiation detector, it wasn't used for that purpose until two decades later. At the time, various doping compounds were explored, but a formula for a practical detector base material eluded the investigators. In developing detector materials, a key tradeoff is between electrical resistance

[22] For a complete history, see: http://heasarc.gsfc.nasa.gov/docs/history/

(low dark current) and charge trapping (efficient charge collection); introducing impurities increases both. Although CdTe compounds of the 70s and 80s made poor detectors, they made great substrates for HgCdTe (Mercury Cadmium Telluride), a popular IR detector material [I48, I52, I55]. As a result, bolstered by heavy DoD investment in HgCdTe and the supporting development infrastructure, a small CdZnTe substrate industry had emerged. As a substrate industry, value was placed on an ability to grow large uniform crystals, and that's were development efforts had been focused [I80].

However, in late 1980s and early 1990s two related breakthroughs reignited interest in producing cadmium telluride compounds as detectors in their own right [D8, D9, I48]. A new growing process for CdTe doped with Zn was developed, as well as a method for growing undoped CdTe wafers. Each approach had its limitation. Mechanically, both CZT and CdTe were brittle and temperature sensitive, but comparatively CdTe could be grown with a larger area of single crystal grains compared to CZT. Electrically, CZT had the advantage of not polarizing and had lower dark current. However unlike in previous generations, these limitations were seen as challenges, not show stoppers. CZT was pursued primarily in North America, while Japan and Europe invested in CdTe development [I80, D40].

By 1994, when the Goddard detector branch had begun its development in earnest, the state-of-the-art was to grow 10mm square wafers and deposit a single contact of electroless gold with an eye dropper. Even at that size, the yield was less than 3% and, although the electroless gold contacts had good electrical properties, the method was too crude to produce the required spatial resolution [I80]. In addition, the contact was too thin and porous to withstand the wire bonding needed to attach leads. To address these limitations, the Goddard team began investing and experimenting along several dimensions of the problem. On the materials front, knowing that they would be constrained by the quality of materials that could be procured, the Goddard team began collaborating closely with the two leading CZT producers in an effort to broaden the supplier base. They entered into a sequence of SBIR (Small Business Innovation Research) contracts (in 94 and 95 [D32]) with the small business that had pioneered the new growth techniques in 88-92 while simultaneously investing "*heavily*" in materials with the more established substrate growing firm [I59]. In both cases, the contracts provided materials for the in-house detector development program as well as resources for the companies to improve their own processing techniques.

The decision to work with two materials growers simultaneously was based on a hunch, and sheds some important light on the way these types of decisions were taken at the time. When the relevant detector branch engineer (CSE#4) was assigned to lead the CZT development, collaborating with the inventors of the enabling techniques was only natural. However, doing his due diligence, the engineer also reached out to his HgCdTe colleagues (who were more familiar with the lay of the substrate industry) for advice [I82]. One trusted colleague suggested he visit a company called II-VI, the established substrate grower mentioned above. During the visit, the Goddard team was impressed by the professionalism of II-VI and learned that they had recently spun-off a group to compete in the emerging CZT sector; it turned out that one of the inventors of the new technique had moved over from the initiating small business. The "*feeling*" that the engineer "*got from the visit*," convinced him to diversify the Goddard investment [I82]. History has shown him to have been correct; the original small business never succeeded in the CZT detector business; in the mid 90s, the company took a new name and has since focused on medical imaging. The substrate spin-off (although it struggled itself) remains a major player today.

On the detector side, the Goddard engineers experimented with combinations of surface preparation chemicals, contact deposition techniques and contact materials [I48]. They leveraged lessons learned from past in-house detector developments for other flight projects as well as standard practices from the semiconductor industry at-large [I48, I59]. This past experience defined the strategies and materials with which they experimented, but within that space it came down to a painstaking process of trial and error. As recalls CSE#3: "*it was painful!*" "*Nothing would stick to the damn thing!*"[I31] The technologists spent hours and hours in the lab trying different approaches. As viewed by the scientists, including CSA#4: "*it wasn't quite random trial and error – he was an expert in this and had lots of tricks - but sometimes it seemed that way*" [I52]. After more than a year of systematic trial-and-error, they finally identified a combination of metals that would 1) stick to the base material, 2) make good electrical contacts, and 3) survive the wire bonding process. The branch head, BH#2, remembers the excitement the first day the technologist found a suitable contact; "*that was the big breakthrough*," now you could have a device [I59].

Once a suitable material had been found, "*depositing the microstructures was easy*" [I48] using standard fabrication procedures; however, choosing what to pattern represented a key design tradeoff that crossed disciplinary boundaries. Although fine resolution detectors were common place, large arrays of fine detectors were not. The challenge was in scaling up the number of output channels that needed to be processed by the readout electronics or ASICs (Application Specific Integrated Circuits). To achieve x-y position resolution with a pixel detector requires a lead attached to every pixel (i.e., N x N channels). This number can be reduced to 2N by employing crossed-strip contacts (as illustrated in Figure 3-3) with a single lead for every strip [I53]. The tradeoff was one of efficiency vs. maturity. While strip detectors minimized the burden on the ASIC, strip detectors were less used (so the corresponding packaging was less mature). The decision to use strip detectors meant that readout leads would have to be individually wirebonded to each of the strips and all of the electronics infrastructure would have to be placed on the side, with implications for the array architecture [I80]. These types of trade-offs were resolved in real time by an *ad hoc* interdisciplinary team. A strong working relationship was forged by at least weekly meetings and often long hours, working side-by-side in the lab [I52]. This close collaboration of scientists and engineers was relatively unusual; as recalls one of the scientists: "*it was kind of interesting; they were all surprised to see a scientist over there all the time. I don't know if I annoyed them or not [the Branch Head] would say oh no, an invader! But meeting weekly really helped*"[I52].

This work was carried out in parallel with several other flight and development projects in the DLL (Detector Development Laboratory), which enabled cross-pollination of ideas and techniques as well as book keeping financial flexibility [I48, I59]. To the extent that detector branch labor was book-kept to this development effort, the funds came from an *ad hoc* "*retainer*" that was paid by the Gamma-ray spectroscopy group to the detector group on a yearly basis; since the early 1990's they have contributed an average of $100K per year to the detector branch's operating budget [I48, I59, I73]. This retainer nominally covered the labor of approximately one person-year. However, it is generally understood that this "*came nowhere near to covering their costs [during the CZT development]*" [I48]. The branch head saw this as a joint investment in Goddard's future, and covered the remaining costs with monies from a series of on-going RTOPS (Research and Technology Objectives and Plans Summary) grants [I59]. RTOPS were a historical form of directed technology funding to NASA Centers that no longer exists. As recounted by BH#2, this level of flexibility was possible because "*it was a different time. Back then the guidelines were very loose. [Grants read] something like we entrust you to go and understand the state-of-the-art in this area*" [I59]. Initial materials purchases also came through these RTOPs accounts; later though, materials procurements were made directly by the science team. All of the science funding came from three parallel sources. The team received relatively stable yearly installments of $100K per year (won with annual proposals) from a DDF (director's discretionary fund) [D3-5] grant and a sequence of three-year, $200K per year, SR&T awards [D1, 2]. The third funding stream came as support for the groups' on-going balloon campaigns (discussed below).

A Focusing (yet unsuccessful) Sequence of Mission Opportunities

By 1995, the Goddard group was the only one in the world able to produce CZT strip detectors with sub 100 micron position resolution [D22]; however an individual detector was a far cry from the array that would be required to implement their BASIS (Burst and All Sky Imaging Survey) GRB finding mission concept [D13]. In 1995, the team won a study contract as part of the "New Mission Concepts for Astrophysics" call. In this context, and in preparation for an imminent explorer call, they began mocking-up a prototype array as a proof-of-concept. Although the 6x6 "engineering model" was only mechanical (i.e., they couldn't actually read out an image) it served several purposes. First, it convinced future reviewers that the idea was feasible. More importantly though, it identified and forced a solution to several unexpected challenges. The act of placing detectors in the custom-made, plastic, egg-carton-like support structure taught the team how to handle the detectors; how fragile they were; how to glue them in; how to position them. Out of this effort, Metrology (i.e., accurately knowing the position of the strips within the array) was identified as an important design challenge. While this effort was lead by the scientist, the detector branch participated heavily [I80].

The next year, in August 1996, a real flight opportunity presented itself in the form of a NASA call for a MidEX (medium explorer) class mission [D29]. The team pitched an extremely ambitious device concept, called BASIS (The Burst and All Sky Imaging Survey), which included a 6x6 array of 36 <100μm pitch resolution CZT strip detectors. Looking back "*we were crazy to think this was obtainable*" [I52]. The review committee agreed with that assessment, and the proposal was not selected, due to serious concerns about technical feasibility. Despite the rejection, the team persevered, continuing to improve the capability along the dimensions of material quality and detector architectures, "*very dejected*" but confident that the science was sufficiently worthwhile that a new opportunity would arise [I85].

A year later, another explorer class call was released, this time for a SMEX (small explorer) mission. They re-pitched the BASIS concept. Not surprisingly, what had been "extremely ambitious" for a MidEX, was no less risky for the smaller (cheaper) SMEX opportunity. Again they were rejected. However, the exercise of proposing was productive for several reasons. First, it maintained a mission focus to the development; and the excitement that embodied. Second, it communicated continued scientific interest in GRB positioning to the HQ funders as well as Goddard management. This latter point merits further explanation. Recall from above, how the director of space science believed that in a resource constrained environment, it was his responsibility to focus available resources on particular projects. In this context, it meant coordinating a pre-screening when an explorer call came out, and allocating engineering support resources to some proposals and not others [I74]. Although he had no formal authority over the engineering directorate, his soft influence was significant. Thus, an important part of proposing early (perhaps before the technology was mature enough) was to signal that this was an important area for future calls.

Exploration at the Architecture Level

Not having been selected for mission development, while continuing to have access to resources from the ongoing SR&T and RTOPs grants, gave the team time to continue developing the technology. Through the close collaboration with the CZT materials industry, enabled by SBIRs, a significant amount of technical information had been shared [D38]. By 1995, the original small business, now operating under a new name, was producing more sophisticated detectors as well as growing materials. The larger company had invested heavily in a semiconductor fabrication facility and was employing similar techniques to Goddard for depositing fine contact patterns [I80, I82]. However, the base material yield was still quite low. It became apparent that as future missions would require very large detector arrays (i.e., many detectors), low yields and the correspondingly elevated costs, would become a constraining factor.

While searching for a way to screen the procured materials pre-processing, the technologists serendipitously expanded the team to include an in-house expert in Non Destructive Evaluation (NDE). As recalled by the NDE expert, "*they had heard that I had an IR camera, so they came knocking on my door*" [I58]. The CZT team got more than a camera. Together with the NDE expert, they developed a set of materials screening processes that correlated observable defects (in the IR) to detector performance [I48, I58]. What this meant was that large CZT wafers could be grown, and screened so that the largest useful pieces (i.e., without defects) could be identified and harvested. This technique improved the effective yield significantly, and has since become an industry standard. In fact, the NDE expert became so intimately familiar with the device fabrication techniques that he took over as the tech lead after the original technologist moved into management. This partnership also created an additional source of funding; the NDE expert was able to secure $60-70K per year, over 4 years from a dedicated non-destructive evaluation funding pot [I58].

Packaging the detectors (i.e., connecting the contacts to readout electronics) became an area of focus during this second round of exploration as well. Most of the standard techniques required the addition of significant heat; however one of the challenging characteristics of CZT is its sensitivity to temperature. Over 150 C, its electrical properties degrade. While they were trying to solve this problem, the team became aware of a small business doing work on cold soldering, when they attended a presentation given on-site at Goddard [I82]. After the presentation, they approached the presenters and began discussing modes of collaboration. As a small business, developing an SBIR proposal seemed like the path of least resistance. This company won two rounds of SBIR contracts in 1997 and 1998 [D32] to investigate the application of the low temperature packaging techniques to CZT. As with the materials companies, Goddard shared their patterning technology with the company (since it was relevant to their packaging solution). While the technology developed through these SBIRs was not implemented directly on SWIFT, the capabilities that were developed have been leveraged on future efforts, including the CalTech-led NuSTAR mission. And, the company has since developed expertise in detector development as well as packaging and in effect transitioned the competency away from Goddard [I80, I82].

Another important collaboration in the area of packaging was formed with APL (the Applied Physics Laboratory). Having worked together before, with APL "*just down the road,*" and with more sophisticated packaging facilities, it was only natural to leverage their expertise as the packaging became more complex [I48]. Where the small company described above had expertise in soldering, APL helped more with array design and construction; particularly through the use of their automatic wire bonding machine [I82]. The work with APL was managed through a standard fee for service contracting vehicle. To this end, the team also worked closely with, and learned a lot from, scientists and engineers in Europe who had been developing a similarly large array of CdTe for INTEGRAL. Although SWIFT and INTEGRAL were conceived around the same time, the European team managed to get approval several years earlier. As remembers one of the Goddard technologists: "*We learned a lot from each other. We even went over to France to see what they were doing. Information sharing was very free*" [I48].

Balloon program

A parallel balloon program, which was run out of the same gamma-ray spectroscopy group, provided several opportunities to "fly" CZT detectors. While the scientific focus of the balloon programs imposed drastically different detector requirements (making the engineering cross-over somewhat limited), the context did provide some level of excitement for the team, some stable funding (balloon campaigns are funded out of the same program as other NASA Research Activities (NRA)) and motivation to experiment with different types of patterns and packaging to suit the various objectives [I86].

Well before the CZT development had been initiated, the gamma-ray spectroscopy group had been maintaining a fairly active balloon program called GRIS (the Gamma Ray Imaging Spectrometer). During its two decade history, GRIS made nine flights, carrying cryogenically cooled germanium detectors, measuring the high energy background. As discussed above, the NRC post doc who was initially brought on to compare various detector approaches, did so with a piggy-back instrument on GRIS [D33].

PorTIA, the Piggyback Room Temperature Instrument for Astronomy, flew an inch square (single pixel) CZT detector three times in 1995. These flights served to characterize the detector background; an important precursor for future CZT instruments. From a detector perspective, although the PorTIA detectors weren't new per se, "*we were learning how to fabricate, test, and package CZT detectors in general so this provided us another challenge*" [I86, D41].

In 2001, a follow-on balloon-born gamma-ray program called InFOCuS (International Focusing Optics Collaboration for μCrab Sensitivity) was initiated. InFOCuS pairs newly developed focusing X-ray optics with a CZT pixel detector focal plane (because of the international collaboration, CdTe detectors were also flown, as will be discussed later) [D42]. From the detector developers perspective, being involved in the InFOCuS program drove them to develop pixel detectors (since the focal plane is relatively small, minimizing channels with strip detectors was not a driver as it was for BASIS) and enabled a close working relationship with CdTe developers around the world including with respect to the array architecture collaboration discussed above [I80].

Changing context leads to feasible design.

Shortly after the second BASIS rejection, the technical constraints of the problem where fundamentally changed by a scientific discovery. In 1998, the European BeppoSAX mission observed a GRB "afterglow" in the X-ray and later UV/optical ranges. It was important enough in the field, that the Goddard lead scientist remembers it vividly:

> *I can tell you exactly where I was when I learned about the BeppoSAX result. I was at a conference in Japan. [gamma ray bursts was one part of it]. Pirro was at that conference, and he was the project scientist for BeppoSAX. On March 1st, he came into the conference with a plot in his hand… I can still remember him walking into the room because he was a little bit late and we were about to start… he said 'we have a new result! We have a new result! He's waving this plot and we all gathered around his paper and it was the afterglow discovery from February 28th. So we knew about it the day afterwards; that's not normal in Astrophysics, but it is part of the culture of the GRB community.* [I73]

The result was significant for several reasons. First, the major driver of the early developments was the need for a wide field of view with extremely high spatial resolution, in a wave band with difficult optical properties. Now, only the presence and approximate position of the short-lived gamma-ray bursts needed to be detected; the accurate positioning could be done by a separate X-ray telescope, with a narrower field of view, slewed to the approximate location after the observation. This drastically reduced the technical constraints on the CZT arrays, to within the capabilities of the day.

The whole community recognized that there was an opportunity here. Within NASA, there was a sense that the Agency had pretty much decided that the next step in gamma-ray bursts, was an explorer class question and that they weren't going to devote a major mission for it. In fact, in that year's Administrator's address to the American Astronomical Society annual meeting, Dan Goldin mentioned (as recalls the scientist):

> *a) gamma-ray bursts; b) how exciting the BeppoSAX finding was; c) how NASA really wanted to make the next step following on the Italian lead; and d) that this was the kind of problem that he thought was perfectly suited for the explore program* [I73].

That was an unusual stance for an administrator to take, and it had a significant impact on the next explorer call [I73]. The reason it's unusual is because the explorer program is a competitive program serving all fields, and so usually NASA is very hands off about saying what they want to do – they just leave it up to the competitors and the peer review process to select what science to fly. The community interpreted Goldin's statements as "*essentially saying that NASA was* *going to select a GRB mission in the next explorer round"* [I73]. Not surprisingly, five GRB mission concepts were submitted in response to the next call.

SWIFT Mission Development

The call came in 1999 with the announcement of a next MidEX (mid size Explorer). The team proposed the much simpler (from a detector point of view) SWIFT concept, which incorporated the large area, coarse resolution CZT array and slewing X-ray telescope described above [D31]. The detectors were non-patterned CZT, supplied by industry (the substrate company). The proposal was successful [D35] and resulted in the launch of SWIFT in 2004.

Although none of the strip deposition and wirebonding techniques that had been developed to support the BASIS concept ended up being used directly on SWIFT, the lessons learned through the development process and other complementary capabilities played an important role in SWIFTs success [I48, I73, I81]. For example, it is unlikely that the CZT materials company would have had the capacity to produce the required 32000 detectors for SWIFT's detector array without NASA's earlier investment [D34, I80]; neither would the NASA team have had the experience to specify "contractable" requirements nor been able to evaluate material quality on the flight procurement without the 10 years of prior experience.

However, since the detector design for SWIFT was so much simpler than what had been developed for BASIS, and the materials growing companies had improved their detectors fabrication practices through close collaboration with Goddard's detector branch, suitable detectors could now be procured from industry directly. This effectively eliminated the need for detector branch involvement in the flight development. In fact, the scientists decided to handle the procurement themselves. As a result, all the engineers moved on to management or other projects, and Goddard has since lost its lead in CZT detector fabrication [I80].

CdZnTe after SWIFT in the US

The science team has moved on from CZT detectors as well; though in a different way than the engineers. SWIFT is still flying; returning data; making discoveries [D43]. It is considered a wildly successful mission and the data and analysis continues to keep the science team busy [I73]. As a result, NASA's gamma-ray spectroscopy group has not pitched a follow-on mission that would leverage all the CZT strip and pixel detector development. New detectors have been procured for the InFOCuS balloon program, but funding for balloon flights has been intermittent at best [I80].

The next American hard X-ray mission will be NuSTAR. It is a NASA Small Explorer mission led by CalTech, managed by JPL, and implemented by an international team of scientists and engineers that is scheduled for launch in 2011 [D44]. Goddard is participating in a technology consulting capacity [I80, I82]. It will fly a high resolution CZT pixel array behind focusing optics (so the detector plane is small). The detectors are being provided by the CZT material grower that got its start developing substrates for HgCdTe, packaged by the company that applied its cold soldering techniques to CZT under a Goddard COTRed SBIR.

Hard X-ray Direct Semiconductor Detectors Internationally

As described above, teams in Europe and Japan chose to address the electrical stability concerns of CdTe rather than the manufacturing challenges of CdZnTe that occupied the Americans. Both strategies have yielded successful detector technology, as evidenced by SWIFT in the US, INTEGRAL in Europe and Astro-E2 in Japan. Through the 1990s, JAXA set up a similar close relationship with their local material growers and detector manufacturers as did NASA. Interestingly though, where the success of SWIFT in some sense ended the development efforts at NASA, a series of failures in JAXA's Astro program have kept development efforts alive. In fact, if the International X-ray Observatory (IXO) moves forward, the Hard X-ray back plane will be a Japanese contribution, leveraging CdTe technology. This is not to suggest that history has proven CdTe to be the superior technical approach – that Japan will likely contribute the hard X-ray telescope is driven as much by politics as any other factor. Nor does it suggest that failures of Astro-E/E2 have enabled to the success of CdTe technology; although the prolonged sequence of similar mission opportunities has provided stable funding and context for continued development. However, if the failures allowed development to continue and become a core competency, while the success of SWIFT terminated the in-house detector program, and if CdTe's status as a core competency is motivating JAXA's claim to the IXO hard X-ray telescope, then the importance of mission sequencing may merit further exploration.

Table 3-2: CZT Shock Sequence

Epoch	Start	Component Exploration	Architectural Exploration	Exploitation	Treading Water and Branching Out	End
Start			* External tech breakthrough in radiation detector development; CSA#11 identifies opportunity * Science director calls BH#1 and tells him support CSA#11			
Component Exploration			* Patterning and attachement techniques invented			
Architectural Exploration		* Selected architecture requires technical breakthroughs		* Explorer mission opportunity * GRB Afterglow discovered * Explorer mission opportunity		
Exploit			* Not selected in Explorer competition * Proposal revealed outstanding technical challenges			* SWIFT approved for formal development
Treading Water and Branching Out						
End						

3.3 The X-ray Polarimeter

This case describes the development of a photoelectric X-ray polarimeter, within NASA Goddard Space Flight Center's (Goddard's) Laboratory for High Energy Astrophysics (LHEA). After more than a decade of development, it will fly aboard the Gravity and Extreme magnetism SMEX (GEMS), to be launched in 2014. The X-ray polarimeter detector system, as the name suggests, will measure the polarization of astronomical sources emitting in the X-ray band, like black holes and neutron stars. Although efforts to measure X-ray polarization have been pursued since the dawn of X-ray Astrophysics, due to extreme technical challenges associated with achieving sufficient sensitivity, this information has been previously unobtainable. The data GEMS will return promises to constrain multiple open theoretical debates.

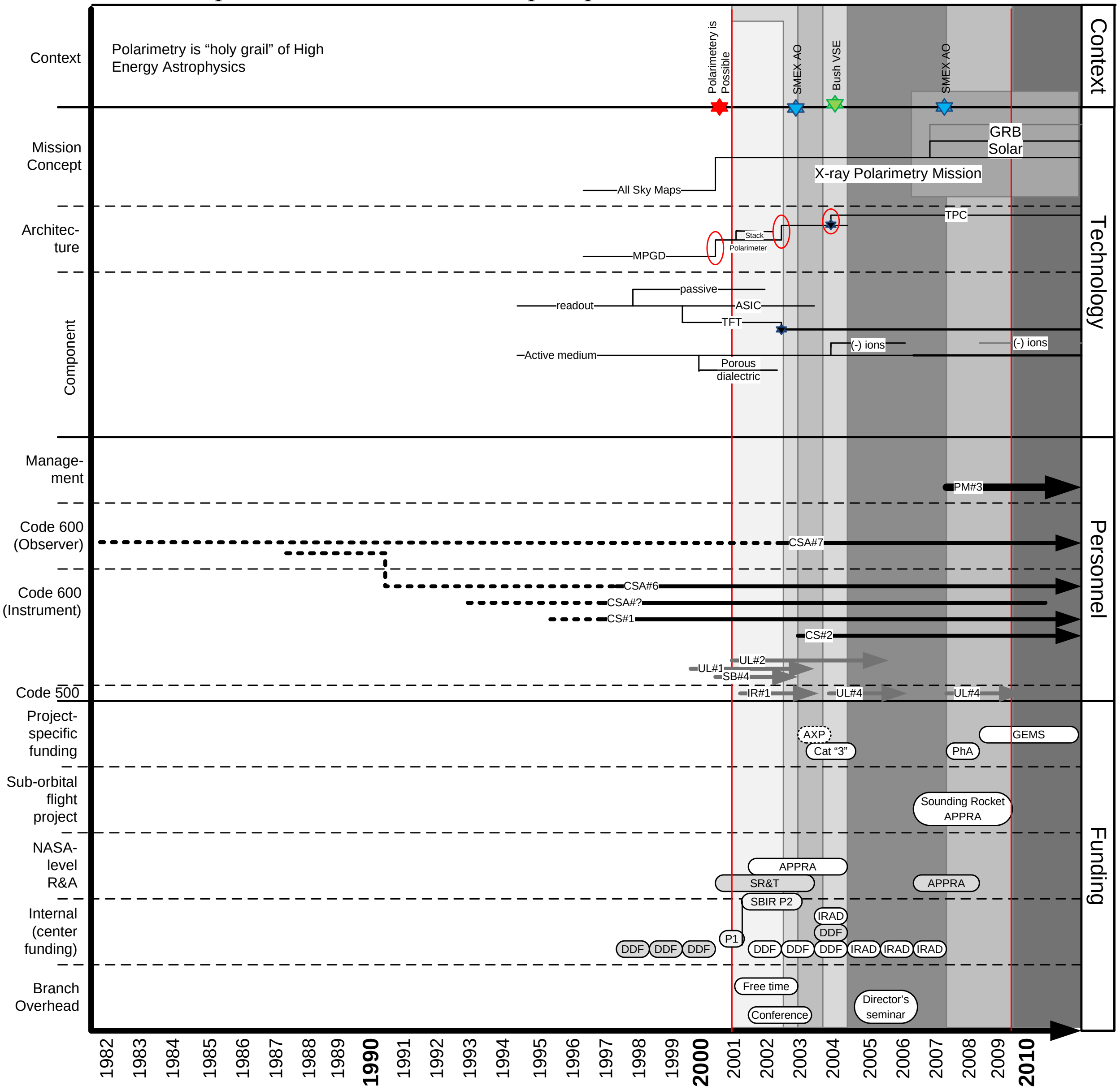

Figure 3-4: X-ray Polarimeter Visual Map

Gestation Period

In the years leading up to 2001, when Goddard's LHEA (Laboratory for High Energy Astrophysics) initiated explicit efforts towards the development of an X-ray Polarimeter, the groundwork for several foundational elements of the future program had already been laid.

By the late 1990s, three of the five individuals who would become core members of the GEMS polarimeter team had united and begun working together for two tangentially related purposes: Leveraging new developments in gas proportional counter technology to develop micropattern detector (MPGD) arrays for (1) a SMEX mission called LOBSTER and (2) for the Next Generation High Energy Gamma-ray (NGHEG) telescope. The two applications both valued large format detector arrays with low power draw, which made proportional gas counters a logical choice. The new developments in array architectures (MPGD's are very finely spaced arrays of proportional counters) and readout strategies promised finer resolution on these large format arrays [D11, D12].

The three technologists brought a varied skill set to the team. The first (heretofore referred to as CSA#6, begun his career as an astrophysicist on the XTE (X-ray Timing Experiment), which used Gas Proportional Counters. Along the way he had become involved with calibration and testing of the instruments, so when LOBSTER began development, it was only natural that he get involved [I60]. The second (CS#2) was an expert in gas proportional counters. He'd been trained as a high energy physicist and spent his career as a *concepts guy*, advancing, and finding new applications for the technology. He'd come to Goddard mid career to work on MOXE (Monitoring X-ray Experiment) and specifically to work on gas counters for that mission. MPGDs for LOBSTER being an exciting new application for the technology, he joined the team [I62]. That's when CSA#6 began working with CS#2. CS#2 met the third core member[23] about a year later, when, struggling with his detector set-up, he was introduced to a *guy who was known for making things work*. This "guy," an astrophysicist had been working on a similar problem in the domain of gamma-ray telescopes [I62].

The concept for micropattern gas detectors, which they were leveraging for LOBSTER and NGHEG, was invented in Europe in the 90s. The idea was a simple one – array proportional counters with very fine spacing – but the implementation was very challenging. At the time multiple strategies for reading out the devices were being explored (c.f. D30, D31). The three Goddard scientists began experimenting in their own right; fabricating devices by hand using a technique called UV laser ablation. After a few years of reading out the detectors with crossed-strips, they sought to read-out their gas counters in a pixelized way because "*[they] thought it would be really important for gamma-ray telescopes in the future"* [I62]. In order to get funding for the effort though, "*[they] were looking for a more immediate application of interest to justify the development*" [I40] (NGHEG was a little too far off since the current gamma-ray telescope GLAST was still in development at the time).

The application area of X-ray polarimetry presented a potentially good justification, both because it was a logical application of the technology, but also because polarimeters were the then head of the high energy physics group's "*favorite pet rock*." "*Every year he would ask us to figure out polarization and every year we would come up with nothing,*" remembers CS#2 [I62]. However, when CS#2 and CSA#6 sat down and tried to work out how pixelized read-outs of a gas counter would lead to polarimetry, their back-of-the-envelope calculations suggested it wouldn't work. Nonetheless, they decided the work was important enough to pursue [I77]. This early R&D work was conducted in part in their "free time" and in part funded by a combination of DDF (director's discretionary fund) [D1, D2] and APRA (level 2 funding) for applications to NGHEG [D11, D12]. As recall the scientists, although there were other larger funding avenues

[23] Who passed away during this study, and was therefore unavailable to be interviewed.

"*in those days, DDF was our most reliable source of funding. You could write a one page proposal and get $75K funding… a year!*" [I62]

A brief explanation of the technology (depicted graphically in Figure 3-5) is required in order to understand the implications of each component innovation described below. Both gas counters, and the polarimeters that are based on them, have four main components: a drift electrode, a gas medium and the corresponding casing, a signal amplifier and readout electronics. In historical gas counters, photons enter the detector from the top and are slowed down by the gas medium so that their incidence can be measured by passive wire leads at the bottom. The main design parameters in this system are the diffusion distance and the readout strategy. With respect to the diffusion distance, there is a fundamental tradeoff between quantum efficiency and modulation (essentially, a longer distance allows you to stop higher energy photons, but the stopping process creates random error that limits the measurement resolution). Thus early designs were only effective over an impractically small energy band. With respect to the readout strategy, where early detectors essentially measured the presence of a photon as a binary yes/no, later designs augmented the readout information in two ways. They added a signal amplification stage (GEM) and replaced the passive leads with active readouts with Cartesian position resolution. This is important because the angle of drift is the basis for constructing polarization information.

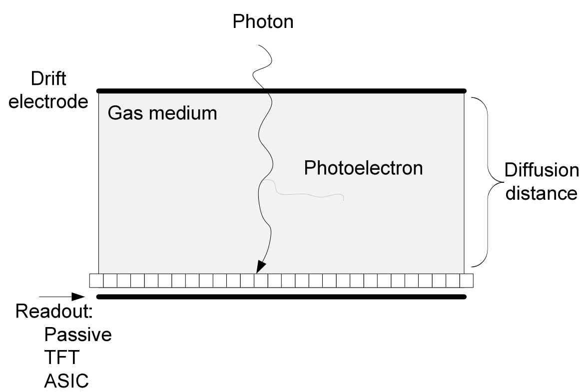

Figure 3-5: Cartoon of Polarimeter Architecture

During the NGHEG and LOBSTER efforts the team sought to improve both of the above described parameters. In terms of the readout electronics, they complimented their in-house exploration of "micro-well" detectors (which is a more sophisticated passive readout that allowed closer spaced pixels and therefore finer resolution devices) with a collaboration with a University Lab (UL#1) to develop TFT (thin film transistor) array readouts [D8, D11]. The goal with the TFT development was to embed an active anode in the gas counters that was optically thin and wide-area. The early TFT work showed enough promise that the Goddard and University team proposed an expanded study, awarded as an SR&T grant (Science Research and Technology, a NASA-level funding mechanisms providing in this case ~$220K per year for three years from 2001-2003). According to CS#2, although the part-time grad students worked hard, "*they never really got anything working*" [I77]. He believes that this was because he never managed to secure the critical mass of funding that would have been required to make a serious attempt.

The exploration of alternative active mediums was also initiated with DDF funding [D2]. The team found a small business (SB#1) with expertise applying porous dielectrics, and began collaborating with them. Instead of using gas (as is typical in *gas* counters, but which is technically difficult to manufacture) they investigated the use of thin window, porous dielectrics deposited on the active readout, as a cheap alternative to gas counters and CCDs in X-ray astrophysics. Following the year long DDF, the small business submitted an SBIR proposal to continue the work. Both phase I and a follow-on phase II contracts were awarded, providing funding for continued work through 2003 [D41]. This technological branch never yielded fruit.

While the Goddard team was focused on the application of MPGDs to All Sky Monitors and Gamma-ray telescopes, in Italy, a University group (UL#2) was committed to developing a practical x-ray polarimeter. A critical result from this group would serve to re-focus the efforts of the Goddard team around the application to polarimetery.

Focusing Event (Costa et al. 2001 Result)

Goddard's interest in using micropattern gas detectors for polarimetry was re-ignited in 2001, by "*fantastic results*" from a group in UL#2 [D29]. As recalls CS#2:

> *It's kinda funny. That group has been working on polarimeters for a while and had been doing it with strip-readout (one dimentional read out) and they'd never really gotten anywhere. In their papers it always said that if they'd had a pixilated readout it would be great. And I had never paid any attention … Then they submitted this Nature paper and somebody here – who I knew – was asked to review the paper… and because of my expertise they gave it to me to take a look… Assuming it was yet another nothing result, I took it home and forgot about it until it was bed time... when I remembered I was supposed to read it. That was a mistake because it got me so excited that I couldn't sleep! They had fantastic results reading out these detectors in a pixelized fashion. It meant we could make polarimetry work!* [I62]

The UL#2 result proved that sensitive (up to a 100 times more sensitive than previously achieved) polarimetery was achievable using gas detectors read-out in a piexelized way. This breakthrough was significant enough that the work was published in Nature. However, substantial expansion of the active readout area would be required before the concept could be considered practical for X-ray astrophysics. The Goddard team believed that the expertise they'd been developing under the guise of MPGDs for LOBSTER and NGHEG could improve the UL#2 devices substantially. Thus, within weeks of learning of the UL#2 result, the Goddard team began writing proposals to pursue X-ray polarimeter development [I60, D14]. Although they continued with the LOBSTER work, and the other projects for which they had responsibilities, their hearts had shifted to polarimetery.

Making it work (2001 – Black et al. 2003)

Although UL#2's result was the demonstration of a physical property, their insight was an architectural one. Where CS#2 and CSA#6 had concluded that MPGDs would not yield polarimetry based on their back of the envelop calculations, UL#2 had constructed a laboratory experiment that proved a way that they could. However their polarimeter was not a practical instrument because it could only work in a narrow band. The Goddard team believed that they could make the UL#2 design practical (broadband) if the detectors could be made optically thin and stacked (see Figure 3-6) [D14]. The idea was that higher energy photons would be read-out by detectors deeper in the stack. In this way each detector could be tuned for a particular narrow band, with broadband characteristics achieved in combination. This concept leveraged much of the team's earlier work, including the thin-film fabrication techniques explored with UL#1.

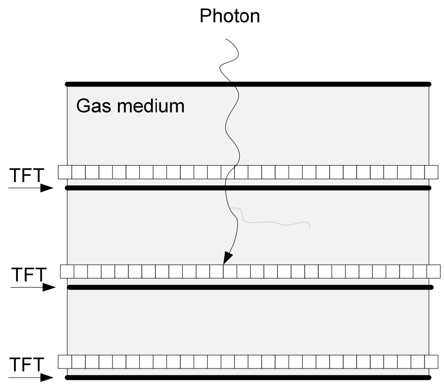

Figure 3-6: Stacked, Optically Thin, Polarimeter Architecture

Despite the excitement of the team, finding resources to support their development effort proved challenging at first. The stacked concept was initially outlined in a 2001 APRA proposal; a proposal that was not funded. The NASA funding managers were skeptical of any promise of polarimetry; X-ray polarimeters had been promised several times before, yet to date positive polarization measurements have only been made of one Astronomical source (the Crab Nebula, the brightest one). The funding managers were not convinced that this development effort would prove any more fruitful. So, the team returned to the ever faithful DDF, from which they received a year's funding to explore polarimeters based on the UL#2 concept [D4].

During that first year, they made a breakthrough that would vindicate them, enabled by a partially-planned chance encounter at the registration desk of a conference in Leicester, England. That's where CS#2 met IR#1, an industrial researcher from a large multinational's research park. Although CS#2 had registered for the conference before learning that IR#1 would be there, when he noticed the topic of IR#1s talk in the schedule, he made a point of meeting him [I62]. For years, CS#2 had been trying to read out his MPGDs with TFTs (Thin Film Transistors); now in IR#1, he'd finally met someone who had a functional TFT array for him to play with. CS#2 pitched his idea to IR#1 and, excited by the prospect, IR#1 invited CS#2 and his team to visit his research facility and "*give it a try*." So, CS#2 and the third core member got to work mocking up a gas counter polarimeter set up to be read out by the TFT. "*We finally got two detectors that actually worked, kluged together this thing, literally out of sheet metal and went down to [the research center] and set it down on one of their TFT arrays and sure enough it just worked! It was great!*" [I62] The trip was actually paid for initially out of pocket by CS#2 and his colleague (that's how important this was to them), although they were eventually reimbursed by a DDF the following year (which had been submitted in advance of the trip).

Armed with this extremely promising, albeit nascent, proof-of-concept, the team sought an opportunity to mature the lab experiment into a prototype instrument. They re-proposed to that year's APRA call; again it was rejected, but this time for different reasons; a response to an Announcement of Opportunity (AO) superseded the APRA funding.

Focusing Event (SMEX AO 2003 –> Category 3 Technology Funding)

In 2003 an AO for a SMEX mission was released [D36], presenting a focus for the team's efforts. As is typically done when a similarly important flight opportunity presents itself, the LHEA came together to strategize about what they should pitch [I61]. It was determined that an X-ray Polarimeter mission was the way to go. It was an important area of science and they were being "*laggard*;" "*it was the right mission for the time*" [I61] especially given the recent progress of the Goddard polarimetry team. A proposal was put together – called AXP (the Advanced X-ray Polarimeter) – based on the preliminary results obtained during the TFT experiments. Both UL#2 and IR#1 were listed as collaborators on the proposal. The design paired existing optics with the TFT read-out, amplified by a Gas Electron Multiplier (GEM).

In addition to partnering with IR#1 and UL#2, while preparing to pitch AXP, CSA#7 an experienced astrophysicists was brought on to lead up the effort on the science side. As she remembers it, it seemed an interesting project and the only other more logical person to fill the project scientist role was already overcommitted, so she volunteered [I61]. CS#1 remembers it slightly differently: "[the branch head] *pulled me aside one day and said [CS#2], I just wanted to give you a heads up that I'm assigning [CSA#7] as the project scientist. I interpreted that to be his way of telling me that he thought it was time to bring in a grown-up to keep us children on track"* [I62]. CS#2 went on to explain how CSA#7 added important credibility to the team. People were still a little skeptical about polarimetry, but CSA#7 had a reputation for only making well-supported statements; "*so they were more willing to take us seriously after she came on board*" [I62].

Despite the CSA#7 credibility factor, the AXP SMEX was not selected for further development because the selection committee was not willing to accept the risk posed by such an immature primary instrument. They did however rank the proposal as a "category 3" instrument and allocated $300K to mature the technology further. "Category 3" is a rating reserved for instruments that show enormous scientific potential but are technically too immature (and therefore risky) to be flown at that time. The idea is to give such instruments some time and resources with which to mature the concept for potential re-proposal at a later date. [I66]

Maturing the Capability, or just exploring some more…

The $300K over two years from the HQ Explorer Office was matched by a Goddard internal investment of an additional $200K [I61, D40]. This enabled the team to bring on an additional contractor – the fifth core member, CS#1, who brought polarimetry experience and a passion for working out instrument implementation issues – this provided an important complement to the creativity in concept generation, characteristic of the other contractor, as they moved forward with instrument development. For almost a year after the technology development award, the team focused on maturing the capability (nearly) as advertised. However, even during that first year, the TFT idea was superseded by a CMOS ASICS under development by the Italians, which had stronger heritage. Given the tight budget, they couldn't afford to pursue both options [I77].

The development path changed drastically in 2004, when CS#2 got distracted by a new diffusion suppression concept that could address what had been believed to be a fundamental limit of the device. This insight would lead to another round of breadboarding and a new architectural concept.

Making it work 2

At the time, it was believed that polarimeter sensitivity was fundamentally limited by a tradeoff between quantum efficiency and modulation (due to electron diffusion in the drift region). The insight was that a Time Projection Chamber (TPC) could create a virtual pixel detector, using time to derive a second coordinate from a strip readout. While TPCs had been in use for at least a decade, the novelty was in the virtual pixel encoding. This strategy effectively eliminated the diffusion tradeoff. As shown in Figure 3-7, since in a TPC, the photon enters from the side (rather than the top) diffusion distance can be increased without increasing the spacing between the two electrodes (the main source of the random error).

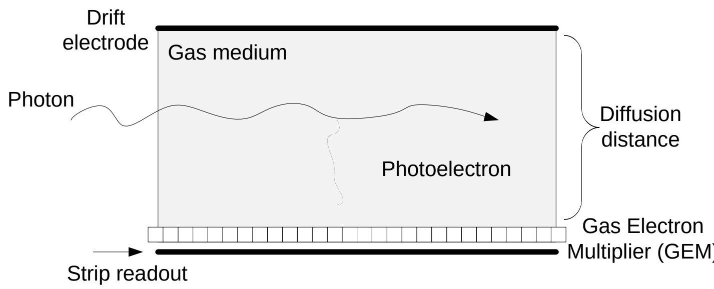

Figure 3-7: Time Projection Chamber Polarimeter Architecture

The inspiration for the TPC concept struck while CS#2 was visiting a researcher at a neighboring university. The visit was prompted by, of all things, the content of the Nuclear Instruments and Methods journal formatting template. As CS#2 was in the process of submitting the TFT paper, he noticed that the content of the template – suppressing diffusion with negative ions – might serve his purposes exactly [I62]. Since the author of that paper was located at a university down the road, CS#2 called him up and asked for a demo. The professor obliged. As recalled by CS#2: "*as I sat there watching the signals come in in slow motion, that's when I had the idea for the encoding scheme that would make pixelized images*" [I62]. Although the particular gas that was being used had too high a Z to be useful in an X-ray polarimeter (it would absorb too many of the photons), the Goddard team would explore the feasibility of suppressing diffusion with other gases at several future points in the development.

The idea of creating a virtual pixel readout in a TPC set-up was sufficiently unconventional that CS#2 kept it to himself initially because he didn't want to embarrass himself if it turned out to be as crazy as it seemed. "*I put myself in a quiet corner to figure out why it wouldn't work*" [I62] and consulted a few of his closer colleagues. Several simulations (and weeks) later, without having uncovered any violations of physical laws, the team decided to pursue the idea. So they "*built one up out of pieces off the side of the road, as [CS#2] liked to say. And it worked! It was great, it was super sensitive. It didn't do imaging, but it was super. It was a polarimeter, which is what we were trying to do*" [I24]. The whole effort was funded by a single DDF and what remained from the technology development funding. The challenge was now to get this revolutionary concept into a workable prototype that could win an opportunity as a next SMEX.

Treading water

Unlike in 2003 when a SMEX opportunity followed the TFT breakthrough almost immediately, now, following the TPC demonstration, there was no equivalent flight opportunity available.

This lack of immediate flight opportunity was not necessarily a bad thing. In many ways, the 2003 SMEX had come too quickly; they received technology development funding precisely because the instrument was not yet mature enough to be approved for project development. In the same way, in 2005, the TPC-based concept was still very much a breadboard. However, moving from breadboard concept to mature enough for project development is a difficult proposition. This phase of development is often called the "valley of death" to evoke an image of a funding drought between research and development. As will become clear, the team did indeed face some degree of funding drought, but they also faced a gap between the competing objectives of always looking for a new better way, and making sure the existing system will always work.

From a funding perspective, where in the early years, DDF provided a relatively stable source of R&D funding; now in the mid 2000s, this was no longer the case. Once contractor salaries were removed, even the development funds from HQ really didn't get them very far: "*we were working off bread crumbs really*" [I24]. During the period from 2004 through 2008 the project stayed alive through creative scrounging for resources. IRADs (Internal R&D – the mechanism which superseded DDF) were awarded every year, totaling $500K over the period as well as supporting an average of 2 FTEs (civil servant salaries) per year. Once the AXP development funds ran out, the contractors wrote APRA (Astronomy and Physics Research Activity – HQ level funding) proposals to sustain themselves and found "day jobs[24]" to help free up resources to keep the project alive. Yet, by late 2005 the contractors were questioning the sanity of sticking with a program that couldn't pay them: "*frankly, we thought we were done. I was actively looking for a job, [the other contractor] had some job offers, but we decided to write two last proposals for the TPC stuff*" [I62, I24, I77]. Both APRAs were awarded in 2006, keeping the team alive [D35]. One was for a sounding rocket test of the technology that would become the GEMS instrument. The second was for a gamma-ray burst detector; a different design but sufficiently related to usefully push the concept forward. The team similarly submitted proposals for a solar polarimeter, one of which was funded in 2007, although they haven't yet had time to work on it [I77]. In both cases, the strategy was one of diversification. Now that they had a promising system, they wanted to re-frame it in as many ways as possible to maximize the likelihood of finding a flight opportunity. In a sense they were too successful in this respect. Where in 2005 they were scraping together resources to cover base salaries, two years later they had won more work than they could do.

Formal Project

In 2007 a new SMEX AO was released and the AXP proposal was recast as GEMS (the Gravity and Extreme Magnetism SMEX) and re-proposed [D37]. GEMS replaced the AXP concept with a TPC polarimeter. This time, with several more years of development and proof-of-concept under their belt, they were selected for phase A concept development in 2008, and in 2009, as the second SMEX mission.

Once the project entered phase B, everything changed for the core polarimetry team both in terms of work practices and resource availability. Now they had to deal with formal project management, and the schedule and budgeting that entailed. They were also joined by a team of engineers. Where until that point, they had continued to follow new technological insights as

[24] E.g., CS#1 began working on a CubeSat that was tangentially related to her area of expertise.

they arose; improving (and changing) the design as they went; never sticking with one design long enough to get a robust prototype; now as a project, producing an instrument that would be guaranteed to work was supposed to be the top priority. As with most projects of this type, the early teaming was characterized by a clash of cultures.

From the perspective of the scientists: "*they had this idea that we'll come up with requirements and they'll go off and build it and come back and present it to us. But that's not how it is. We're not far enough along with the stuff to really know exactly what you need to build… there's going to be some iteration because it's might not work as well as we'd hoped*" [I24]. From the perspective of the project management team: "*they're extremely intelligent, very passionate, they really care, but they just don't get that there's a trade-off; we're on a capped mission, when we run out of money, we run out of money." "And that's just the money, with the schedule… they keep saying just a little longer, a little longer… they don't understand that they're holding up the whole project and there's a train moving behind them. If we don't finish on time they get nothing – no science*" [I63].

Over the course of the project relationship, understanding has evolved. One key example of the scientists' shift to a project mindset is the decision to use normal gas with an ASIC developed for CERN, rather than the negative ion gas with which they'd been experimenting. Compared to the sensitivity of the negative ion solution described above, "*it's less sensitive but we're confident that it will work (which is not the case for the alternative) and that's important for a project like this*" [I77]. The design has now stabilized and efforts are truly focused on instrument development. The negative ion effort does continue however, under the auspices of the GRB polarimeter trajectory.

From a funding perspective, although winning the SMEX created a much more stable development environment than had previously been enjoyed, it also imposed many more constraints about how that money could be spent. As the second SMEX, GEMS was slated for launch in late 2014. And, following project approval, the $105M cost cap would be strictly enforced. Given that labor represents a significant project expense, the full project team could not be ramped-up until approximately 3 years prior to the planned launch. Yet, the need to maintain technical expertise and momentum in the polarimeter development presented a conflicting constraint. The GEMS project team persuaded the SMEX program office to provide an advance on the project money - $5M over 18 months – which would allow the polarimetry team to proceed fully staffed, while maintaining a part-time skeleton project team to manage their efforts. However, it became quickly apparent that some supplemental funds would be needed to mature the polarimeter sufficiently for easy insertion once the project "turned-on." This later conundrum has been mitigated as follows: The GEMS project has adhered rigidly to the budget provided by the project office, but some of the scientists may have applied for other technology development funding to pursue other applications for similar technologies. That the work is also directly relevant to the GEMS polarimeter is a pleasant side-benefit. Thanks to these years of pre-work, the team is confident that the polarimeter will be a success and the mission will yield important new discoveries.

Table 3-3: Polarimeter Shock Sequence

Epoch	Start	Component Exploration	Architectural Exploration	Exploitation	Treading Water and Branching Out	End
Start		* Implications of external breakthrough recognized by CS#1 * Insight about TFT				
Component Exploration						
Architectural Exploration				* CS#1 meets (stalks) IR#1 at conference => equipment to build idea * demonstrate practical polarimeter * Explorer Opportunity * demonstrate feasibility of TPC		
Exploit			* CS#1 finds UL#4 * collaboration sparks architctural insight (TPC)		* Drought: ran out of money expecting next explorer call	* GEMS approved for formal development
Treading Water and Branching Out	Solar polarimeter pathway launched			* Explorer Opportunity		
End						

3.4 The Ion-implanted Silicon Thermometer Microcalorimeter

This case describes the development of microcalorimeter detector arrays for applications in X-ray astronomy. The key insight, which initiated the multi-decade (X-ray) microcalorimeter innovation pathway, is generally attributed to a particular Goddard (infrared) astrophysicist; however multiple institutions and individuals have played key roles over the last 30 years. The development of microcalorimeters has been motivated by the need for a high resolution X-ray spectrometer that has high intrinsic quantum efficiency for use in X-ray astronomy. The high quantum efficiency is important because of the very low X-ray fluxes characteristic of most celestial X-ray sources. Although microcalorimeters were first demonstrated in 1984, and selected for flight aboard the Advanced X-ray Astronomical Facility (AXAF) in 1986; due to a series of unfortunate circumstances (AXAF-S split, Astro-E launch failure, Astro-E2 cryogenic system failure, discussed in the previous pathways) the detectors have yet to return data from a space-based platform. This milestone will hopefully be achieved in early-2014 aboard Astro-H. During this 30 year history, three fundamentally different classes of microcalorimeters have emerged – one based on semiconductor thermometers, one based on superconducting transition edge sensors and one based on magnetic resonance. Figure 3-9 illustrates how the best-in-the-world detector performance has improved over time. We consider the three generations here as separate, albeit entangled, innovation pathways. Only the first two histories are presented in detail (as the third is still quite immature).

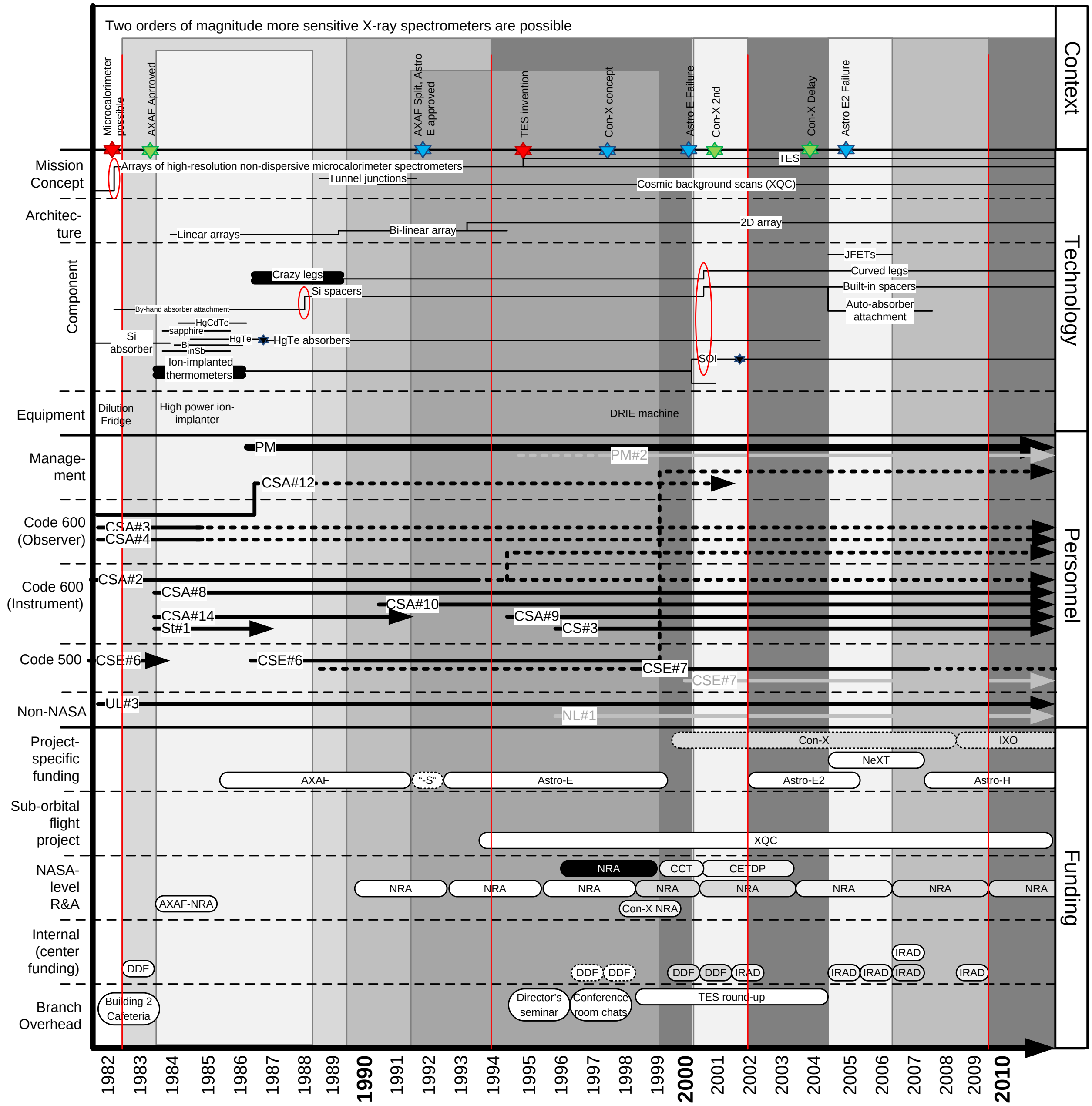

Figure 3-8: Si Microcalorimeter Visual Map

Gestation period

Although microcalorimeter development explicitly started in 1982, under the context of a call for ideas for the then future AXAF (Advanced X-ray Astronomical Facility – later renamed Chandra), much of the key groundwork was laid several years before that.

The IR astronomer (heretofore referred to as CSA#2) who would eventually propose the X-ray microcalorimeter idea joined Goddard in 1979 for a seemingly unrelated purpose; as an NRC post doc, with a grant to build an infrared spectrometer. At the time, bolometer-type thermal detectors[25] were the device of choice for his application, and he knew that he would need a lot of them. However, with the state of bolometer manufacturing in its infancy as it was – "*people were doing things like soldering wires to pieces of germanium and things like that… and how well things worked depended in detail on how the soldering went*" [I42] – and with university lab "boutique" shops being the primary suppliers, individual detector prices were quite high; much higher than the CSA#2 could afford for his spectrometer project. Faced with a conundrum, the resourceful young scientist decided that he could manufacture the detectors himself.

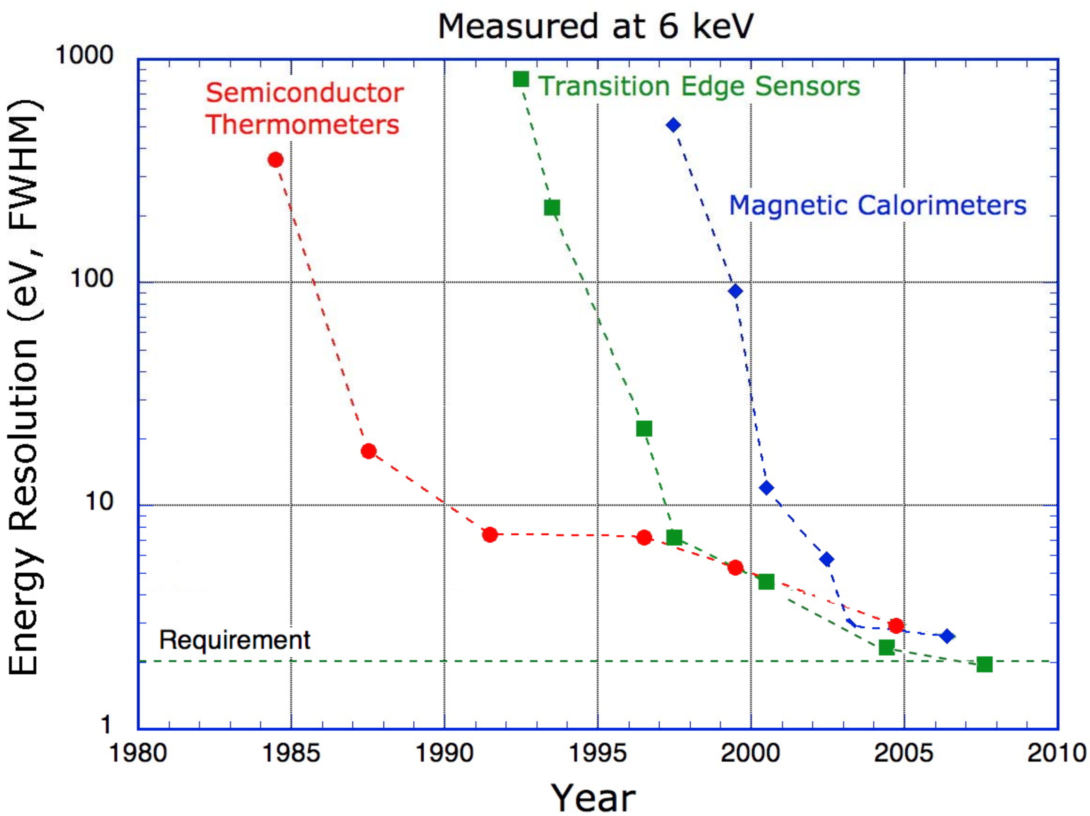

Figure 3-9: Performance Improvements of Mircorcalorimeter Devices over Time.
Each point represents the best-in-the-world, individual detector resolution, measured at 6 keV. This figure was created by the Goddard Calorimeter group.

Although he had no previous device experience, he had become confident in his "*tinkering*" ability during his grad school days,[26] and he recognized that one of Goddard's in-house competences was a relatively well equipped semiconductor processing facility. So, he walked over to talk to "*the guys in the processing lab*," [I42] told them that he was interested in making

[25] A Bolometer is a device for measuring the power of incident electromagnetic radiation. They consist of an absorptive element, connected by a weak link to a heat sink. They work by measuring temperature changes in the absorber.

[26] CSA#2 explained how a graduate degree in Physics is like an apprenticeship. The student gets real hardware experience by being involved in the full project development cycle at least once. This is different than what aerospace engineers seem (to him) to get, who touch hardware for the first time in the highly regulated flight development environment when they join NASA.

detectors and that he thought they could do it. They would prove him right; and the close collaboration that was in part forged by this initial bolometer experimentation has persisted through future detector developments (e.g., CZT described above).

In the late 1970s, most bolometer designs employed Germanium thermometers, but CSA#2 was convinced that silicon (Si) could serve the purpose as well. This was important because Si was the material with which the processing lab had experience. CSA#2 did some investigating and found a publication written by a scientist at Texas Instruments (TI), detailing a bolomemter that used silicon thermometers. Excited, he called him (the TI guy) up and asked for details. The strategy seemed promising, so CSA#2 asked if the TI guy had any scraps left over from the development. He had, and was willing to share them. CSA#2 took them to the lab, and with the help of the technicians, made his first detector. "*Pretty soon*," they were making reproducible detectors with good noise performance that was, more importantly, predictable by theory! More than just a good manufacturing characteristic, reproducibly predictable by theory was a critical capability because it meant that volume and operating temperatures could now be tuned to meet particular design specifications and by implication science requirements.

The theory that was being reproduced, and that predicted the relevant scaling laws, was itself quite new [D28]. It came out of some other tangentially related (theoretical) *tinkering*, connected more by the physical proximity of the two tinkerers than by any other attribute of their work. At the time, another Goddard scientist in the observational cosmology group (CSA#3) was working on trying to understand the noise performance of his system for COBE (the results of which would later win him a Nobel Prize). It was through that effort that he developed what would become the widely accepted noise theory for bolometer-type devices. The main relevant contribution of this non-equilibrium thermodynamic noise model was "*that he presented things in such a way that you could see the scaling laws*" [I42]. CSA#3 has often mused, that the resultant paper is by far his most cited work; far more impact (as measured by citation counts) than the work it enabled, for which he received his Nobel prize [I51].

Armed with reproducible bolometers and a theory for how they scaled with different tunable parameters, CSA#2 began thinking about the physical limits of sensitivity that one could achieve with the devices. It became clear that low temperatures held the most promise for increasing sensitivity. So, he invested in a dilution refrigerator for the lab, using funding from the flexible Director's Discretionary Fund (DDF). At the time, the Goddard group was one of the few seriously exploring low temperatures in the world [I42]. However, in hindsight "*it turned out to be an absolutely key thing to have that [dilution fridge]*" [I42] because it gave them the tools with which to experiment. This experimentation occurred largely in CSA#2's free time (he was still working full time on the spectrometer). As he explains: "*that was in the days before full cost accounting and plus that was at the time when I was spending huge amounts of time in the lab, I was here all the time […] I would come late at night after the wife and son had gone to sleep, I'd come back to the lab*" [I42].

In 1981, having established his group as leaders in the domain of IR bolometers, CSA#2 got a call from NIST. The NIST scientist on the line was thinking about using thermal detectors to detect laser pulses, and was wondering if CSA#2 knew the minimum pulse energy that one could detect with a bolometer. Although he didn't have the answer off hand, CSA#2 was confident that he could work it out. So he did. He reported his result to the NIST scientist, and put the calculations in his desk drawer in case it became relevant at some later date [I42]. This was significant because a bolometer that also measures energy (not just power) is essentially a calorimeter.

Initiating collaboration

That "later date" came about 1 year later, when a group of Goddard X-ray astronomers – who happened to sit down the hall from CSA#2, and knew him well from frequent lunch breaks in a shared cafeteria – stopped by to mention that an opportunity to propose a major instrument for AXAF would be announced in the next few years.[27] They knew that they were going to propose a spectrometer, but they needed a way to distinguish themselves if they were to have a chance of winning. At the time the standard detectors for x-ray spectrometers were silicon diodes, but they weren't nearly good enough for the next generation Grand Observatory. Multiple ideas were batted around the lunch table [I74]. Eventually the X-ray astronomers had come to CSA#2 – an IR astrophysicist[28] - because they'd heard that he'd had some success with a smaller bandgap semi-conductor for IR spectrometers which, if it worked for X-rays, could improve the X-ray resolution as desired.

After they explained the basic principles of x-ray spectrometers to CSA#2 (he had no previous experience in that wavelength), CSA#2 thought for a minute and concluded that InSb (Indium Antimonide – the smaller bandgap semiconductor) wasn't appropriate, but that bolometers were worth exploring. To that end, he pulled out the calculations that he'd originally performed for the NIST scientist; and from that he extrapolated a fundamental limit of 1eV at 6 keV if they could get the devices to 0.1 K [I42, I74]. That was very good; much better than the X-ray astrophysicists had hoped for. So CSA#2 got to work.

Proving theoretical feasibility of the concept

The initial proof-of-concept work was conducted by a core team of CSA#2, a visiting Astrophysics researcher from the University of Wisconsin (UL#3) with a reputation as a "*spectacularly talented experimentalist*" [I88], CSA#3 assisting primarily with the theoretical limits work, and CSA#14 - an Astrophysicist from the Laboratory for High Energy Astrophysics (LHEA) who brought knowledge of the AXAF context.

CSA#2 and UL#3 were not acquainted prior to the collaboration. They were introduced by the X-ray group Branch Head, who had a history of collaboration with the University of Wisconsin, where UL#3 was junior faculty. UL#3 was recruited because the effort needed an X-ray knowledgeable experimentalist – and all the potential internal candidates had either recently been promoted to management positions, or were occupied with another big project. So the Branch Head asked this junior professor to do it. Deciding to participate was a big decision for UL#3. As

---

[27] It had been ranked first in the NRC Astrophysics Decadal Survey for the 1980's (released in 1982) among major new programs. The call was for "an AXAF operated as a permanent national observatory in space, to provide x-ray pictures of the Universe comparable in depth and detail with those of the most advanced optical and radio telescopes. Continuing the remarkable development of x-ray technology applied to astronomy during the 1970's, this facility will combine greatly improved angular and spectral resolution with a sensitivity up to one hundred times greater than that of any previous x-ray mission."

[28] With the Astrophysicists discipline X-ray and IR Astronomy are different specialties with essentially no overlap, requiring different expertise.

he remembers, he eventually said yes for a variety of reasons, none of which involved a particular desire to be involved in AXAF:

> *I said, well let me think about it… and I went home and did the calculations… because it kind of seemed like a crazy idea… but the calculations sort of came out... I was in this position of being a fairly new assistant professor and having this sounding rocket program going… and sounding rockets in X-rays were kind of on their last legs because now there were all these satellites up… and it's very hard to compete against a satellite with a sounding rocket… but I was in the business of educating graduate students so I wanted to stay in sounding rockets if I could (and sounding rockets are about the only way for students to actually learn about building hardware) so I looked at it as hey, if this thing works, it's never been on a satellite, so we could put it on a sounding rocket and do some new stuff… keeping my sounding rocket program alive for at least five years (of course 27 years later, my sounding rockets are still the only things that have gotten astronomical data with these detectors)…* [I93]

The collaboration worked out better than either could have expected. As recalls CSA#2 "*From the first time I met [UL#3]it was an absolutely perfect collaboration because he was so very smart and knew a lot and had done a lot of things, I mean just perfect teaming… I ended up working with [him] for many years… we probably talk several times a week. Usually arranged marriages, you don't think of as working all that well, but it came off awfully well."* The feeling was mutual, in UL#3's words "[CSA#2] *is probably the quickest guy I've ever met… just always dripping with ideas."* [I93]

They got straight to work: "*we [CSA#2 and UL#3] basically spent the whole summer in the lab in 1983"* [I42] funded by a DDF. By the end of that summer, they had obtained the first microcalorimeter spectrum (published in 1984 as the famous [D29 and D30]).[29] Although the performance of that first detector wasn't even as good as the best silicon diode at the time (achieving an energy resolution of ~250 eV FWHM at 6 keV as shown in Figure 3-9), and nowhere near what would be needed on AXAF, it proved the concept [D1]. That was enough to get detector development underway in earnest. The next step was to explore the theorized low temperature limits. Over the next year, they worked with the dilution refrigerator (that had previously been purchased for CSA#2's lab) and began collaborating with the cryogenics branch, who were developing sub-Kelvin refrigerators (ADRs – discussed in detail in the CADR innovation pathway).

Selection for a Flagship-oriented Technology Development Effort

In 1984, an NRA (NASA Research Announcement) was released as expected, requesting instrument proposals to fly on AXAF. Goddard's X-ray astrophysics group, lead by CSA#12, pitched to fly XRS (the X-Ray Spectrometer) based on a microcalorimeter detector plane. Given

---

[29] In fact, "*the first experimental development coincidentally began simultaneously on both sides of the Atlantic in 1982. In Milan, Ettore Fiorini had been working on detecting neutrinoless double beta decay and, intrigued by a suggestion in a preprint by Guenakh Mitselmakher that the betas might be detected thermally, went to Niinikoski to investigate the practicality of this idea. They devised an approach that was developed into the first successful physics experiment using thermal spectrometers*" [D25]

the extreme novelty of the concept at the time, they proposed to continue developing the incumbent technology in parallel.

> *At the present time, Si(Li) technology represents the most generally useful proven approach to X-ray spectroscopy over the energy range 0.3-10 keV. […] There is no denying that the CCD and calorimeter (discussed in the following section) technologies have great promise, and so we assume that the best approach to developing their enormous potential for AXAF is by their selection for this definition. Until that potential is objectively demonstrated, however, we feel that it is prudent to also include in the definition studies those already-proven technologies which are guaranteed to produce important new results whether or not the newer technologies approach their potential capabilities. (p 29)*
>
> *The simple calculation below indicates, however, that it should be relatively straightforward (at least in principle) to fabricate a device which has FWHM ~1eV, or almost two orders of magnitude better than a state-of-the-art Si(Li) detector. (p.32)* [D1]

Despite its relative immaturity, XRS was selected as a category 3 instrument. The category 3 designation means that their scientific merit was rated extremely high, but the technology was deemed too immature to be incorporated into a flight project as is. The classification as category 3 translated into a year's worth of technology funding and the opportunity to prove that they could be mission ready. At the year-end review, the committee was impressed by their progress, but did not yet consider them mission ready, so they gave them a year extension on the technology development [I42 and I74]. As CSA#2 remembers it: "*They asked me what the remaining hard problems were; so I listed them. The next day, they gave us a list of things to do by next year. It was all the hard problems I'd listed!*" A year later in 1986, hard problems mostly solved, they were selected as the primary spectroscopy instrument for the mission.

A little background…

Each individual microcalorimeter detector has for main components, as shown in Figure 3-10: an absorber, thermometer(s), a weak thermal link, and a heat sink. A microcalorimeter works by measuring temperature changes due to an incident X-ray photon. The incident X-ray is "thermalized" by the absorber, producing a temperature change that is measured by the thermometer. The weak link then provides a conductance path between the absorber and the heat sink, allowing heat to be dissipated (so the absorber can return to its initial temperature). In designing efficient microcalorimeters, the goal is thus to find 1) an absorber material that is both a) opaque to x-rays, and yet has a low heat capacity (so that a single absorbed photon yields a measurable temperature change) and b) thermalizes well;[30] 2) a thermometer that is sensitive; and 3) a thermal link that is weak enough that the time for the base temperature to be restored is the slowest time constant in the system (compared with thermalization and diffusion times), but not so weak that the device is too slow to handle the incident X-ray flux.

Flight instrument detector planes consist of arrays of tightly packed detectors, with the number and size of the individual detector pixels determining the field of view and spatial resolution. At

[30] it must reproducibly and efficiently distribute the energy of the initial photon across a thermal distribution of phonons

the array architecture level, key challenges relate to ensuring uniformity across the detectors, minimizing "dead zones" (i.e., packaging the detectors so that there are no gaps where incident photons can pass undetected), thermal isolation and reading out the multiple channels. While some of these challenges may seem like generic array packaging issues, characteristics of microcalorimeters create unique challenges for fabrication as will be discussed in the sections that follow.

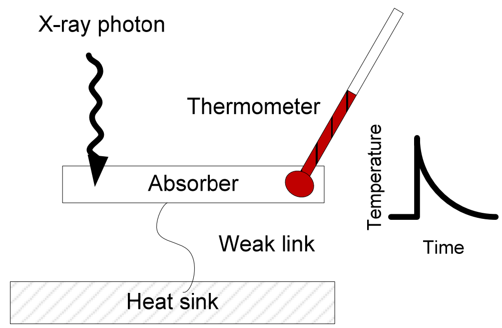

Figure 3-10: Cartoon of an Individual Microcalorimeter (adapted from phonon group figure)

Returning to the development…

The team addressed the detector and array level challenges essentially simultaneously, using mock arrays to test different detector architectures and configurations, and exploring attachment and arraying strategies in the process [I72, I68].

Initial Technology and Architectural Exploration: Making it work (1984 through late 1980s)

As the mission prospects for a microcalorimeter instrument materialized (in 1984), and a steady stream of funding from HQ seemed ensured, the team began expanding. CSA#8, who would become a key member of Goddard's microcalorimeter group, and CSA#14, joined the team as NRC post docs; in addition St#1, then a recently graduated undergrad taking a "break" before starting grad school, came to work on the project for a few years. They worked hard because *"it was exciting in the sense that it was all pioneering"* [I68]. Although much of the groundwork had already been laid, there were still many open research questions for the young researchers to tackle. As recalls CSA#8, "*when I got involved, it was like now let's think about how to actually make these things"* [I68]. "Making it work" proceeded along several component and architectural dimensions simultaneously, with each decision impacting several other factors. The below discussion is divided in terms of development streams; components first, architectural considerations later. All the streams were proceeding in parallel, but for clarity, they are described from start to finish independently, within the period under question (e.g., initial exploration).

Thermometers: Achieving the desired resistance in a repeatable way

By 1984, the first key decision, to use ion-implanted Si thermometer, had already been made based on a combination of material constraints and practical considerations. As explained in the 1984 NRA proposal:

> *We have baselined ion-implanted Si for this proposal because it has high enough Debye temperature to give a predicted FWHM which can be as low as 1 eV, and the*

> *implantation techniques required for the formation of the thermistor are well understood. […] Another possible material is diamond, which has a very high value of Debye temperature, but the required X-ray absorption […] demands a very thick detector, and the implantation of a thermistor is problematic…* (p. 33) [D1]

From the perspective of the science team, the thermometer question was settled in 1984. However, having chosen to ion-implant Si with (P and B), the team was confronted with a critical practical challenge which would occupy the Detector Development Lab (DDL) technologists for many years to come. The extant theory states that implantation dose determines the shape of the resistance-temperature curve of the base material. Thus, given a certain operating temperature, one should be able to control the resistance (and consequently measurement sensitivity) by means of the implantation density (also known as dose). However, in practice, resistance of the Si is extremely sensitive to the implantation dose, and the fabrication equipment at the time did not permit a high level of repeatability. As a result, the thermometer yield was extremely low [I83, 84].

To make useable thermometers at all, the DDL technologists employed a brute force strategy: They would put many wafers into the ion implanter and implant the full range of doses, and screen them later. In this way, the hope was that at least one of the wafers would have the desired electrical properties. As recalls the microcalorimeter fabrication lead: "*we were using this old ion implanter… we revised the process as much as we could within the hardware that we had, but it was still very variable from run to run*" [I83]. With the old tool, even implanting over the full range wouldn't guarantee a single useful wafer. "*What we ended up doing was buying this new ion implanter for a few million dollars. Its repeatability was much much better. Now you could at least get some wafer in the range that had the right dose*" [I83]. The strategy was far from efficient but it did guarantee a non-zero yield.

The decision to buy a new few million dollar ion implanter was a big one at the time. To put it into perspective, the cost of the ion implanter was equivalent to the two year budget of the entire microcalorimeter development effort at the time, and it is one of the biggest pieces of equipment that the fabrication laboratory has ever purchased. However the investment was made at a time when the lab was expanding – Goddard was developing microelectronics and detector fabrication as a core competency – and the microcalorimeter development effort had the support of more than one branch. The alignment of multiple (science needs) was key:

> *We got together, both my management and the management from quality assurance, got together and convinced the center that we needed [the ion implanter and a better clean room] With the combination of those, institutional money got put on it. That sort of argument goes around every year… there's multi-project support money… various parts of the institutions make their budgets and claims and you know lobby for this and you get your scientists to say, oh yeah, we need this to enable the next missions. In those days, there seemed to be more money than there is today* [I83].

Despite the equipment acquisition the ion-implanting strategy remained extremely inefficient (in terms of yield) and was subject to multiple sources of measurement noise; namely, "1/f noise" due to variability in the depth across the implanted dopant and Johnson noise produced by ohmic heating.

One parallel technology path that was pursued for a short time involved an effort to eliminate the Johnson noise term (which is endemic to resistive measurements). CSA#8 began investigating kinetic inductance thermometers, which measured changes in inductance (instead of changes in resistance) and therefore didn't dissipate any energy. As explained in the 1993 research proposal: "*The energy resolution of such a device could be an order of magnitude better than a calorimeter with a resistive thermometer with the same total heat capacity, and it should in principle be possible…*" [D2, p. 1]. However, flight project pressures in the late 1990s forced the team to focus on maturing the method that was already proven. "*Based on the limited success of kinetic-inductance sensors in the past several years, we suggest shifting the resources toward more promising alternatives [TES].*" *(*Reviewers comments re: NRA 95-OSS-17, D5) The Kinetic Inductance approach was never revisited, even when exploration resources became available later, because by that time, the TES approach had indeed demonstrated more promise [I84].

<u>*Absorber Material: Finding a material that would thermalize X-rays*</u>

The choice of ion-implanted silicon had implications for absorber design as well. The relatively low sensitivity of silicon thermometers imposed a requirement for a low heat capacity absorber. Coupled with the need for high stopping power (at the energies of interest) and good thermalization (consistent temperature response to incident photons), the candidate material choices were quite limited [I84, I72]. The earliest microcalorimeters employed silicon absorbers, partially because of the team's prior experience with the material, but also because of silicon's relatively low heat capacity. However, as test devices were constructed, they discovered a large additional noise term not predicted by theory, and it was immediately clear that a separate absorber would be needed [I88]. They believed that this was due to incomplete (and variable) thermalization of the incident photon. So, they began searching for higher efficiency thermalizing absorbers "*by attaching small samples of materials with low heat capacity and small or nonexistent band-gaps (e.g., metals, semimetals, narrow-gap semiconductors, and superconductors) to working detectors*" [D33].

As recalls CSA#8 "*it was a real challenge to find a material that would thermalize x-rays*" [I68]. They explored a wide range of semiconductor materials, and in this respect, the interdisciplinary nature of the team proved advantageous. Recall that the concept originator (CSA#2) was an IR Astronomer by training. "*The key thing about this whole process was this confluence of the right backgrounds […] [CSA#2] had a tremendous amount of knowledge of solid state materials that could be used as thermal sensors, so that's where it started*" [I68]. The alternatives they turned to were the materials that one would expect to find lying around an infrared lab [I88]. In addition, they drew on expertise from colleagues in the international community as well as in other groups at Goddard. "*We weren't really going through yellow pages to find material scientists, but we were talking to our colleagues in the field […] discussing things and trying things out*" [I68]. They tested InSb, Bi, HgCdTe, HgTe, and sapphire, eventually settling on a mercury telluride (HgTe) absorber [D33]. HgTe enabled an order of magnitude better energy resolution than the original Si (see 17 eV breakthrough in 1987 – Figure 3-9). As explains CSA#10:

> *HgTe turns out to be pretty much the only absorber you can use with the Si thermistors, and it's a compromise. It's got high Z, high-atomic-number constituents, it's a semi-metal so it doesn't have a gap […] it's Debye temperature (lattice specific heat) isn't particularly wonderful, but it's acceptable… But it's hard to grow which means it needs to be grown separately from the detector and be attached with epoxy one detector at a time…* [I72]

CSA#2 framed the materials choice slightly differently. In his mind, there may well be better materials, but HgTe works well enough that it was no longer worth looking for a better alternative:

> *We don't have any illusions that what is being used now is in any sense optimal. It's something that works. And materials problems are something that are typically very difficult so no one wants to actually go off on the "great quest" unless you really have to do it.* [I88]

Although the absorber material question was effectively put to rest after the 17 eV HgTe result, some parallel investigation of superconducting absorbers continued into the late 1980s. The idea was to leverage the fact that certain superconducting materials have i) very good stopping power and ii) very very low heat capacity below their transition temperature, so, in terms of stopping power per unit of heat capacity (the relevant figure of merit) they're spectacular. This property is particularly relevant at higher energies, and that's where it was first investigated seriously [I72]. The problem (in the 1-10keV range) was that superconductors don't thermalize X-rays quickly enough to be practical [I89].

This issue of thermalization rate led to another development branch. Also in the late 80s, one member of the team began working on superconducting tunnel junctions. Tunnel junctions are another form of non-dispersive spectrometers. Instead of measuring energy as heat, they measure energy as charge. The main potential advantage of this approach was the extremely high photon counts that it theoretically enables. However, in practice, collecting every electron-hole pair created by the incident X-ray proved at least as difficult as the complete thermalization problem. And, as with the superconducting absorbers, and kinetic inductance thermometers, the technology branch was pruned before its feasibility had a chance to be established due to program pressures to focus [I84]. In fact, other researchers in the community did pursue this path; although the concept never did yield fruit.

Absorber Attachment: Achieving array-wide uniformity

The choice of mercury telluride (HgTe) for the absorber material meant that each absorber needed to be attached individually post-fabrication... by hand. This literally entailed "*physically glue[ing] each of the absorbers onto the pixel with epoxy… and if the epoxy ran, it created a non-uniform attachment point… and unpredictable thermal performance*" [I84] with a range that extended outside acceptable limits for instrument performance. The necessity of by-hand attachment was "*painstaking*" and limited the scalability of array sizes to the number of individual absorbers that could be realistically attached by hand. For AXAF the array was only 6 x 6 (i.e., 36 pixels), but for future missions, the science demanded a much larger field of view.

The out-of-limits thermal variability posed a near-term and more serious problem. The *unpredictable* degraded energy resolution couldn't just be calibrated out, and thus necessitated post-production screening.

The "*big breakthrough*" was a clever, albeit crude, "*engineering solution*" that came from the scientists [I84]. They solved the epoxy running problem with a silicon spacer – glued between the thermometer and absorber, it served to manage the surface tension of the epoxy and maintain a constant bond joint from pixel to pixel. The attachments still needed to be made by hand, but at least now the interface was thermally consistent. This enabled good enough "array resolution" to meet AXAF's requirements.

Thermal Isolation: The weak link and crazy legs

The microcalorimeter cartoon in Figure 3-10 represents the "weak link" as a thin squiggly line between the absorber/thermometer and a heat sink. In the real system, this thermal isolation was achieved by etching all but very thin silicon beams which maintained a "weak" attachment between the thermometer and the array structure. The key design parameter for the weak link is a time constant - the thermal decay time constant needs to be slow enough for complete thermalization, but faster than the flux of incident X-rays. Achieving the necessary thermal isolation in practice presented one of the most difficult fabrication challenges [I83].

In the late 1980s, Goddard's DDL (and the semiconductor industry at large) was primarily using wet chemistry techniques for micromachining silicon. This limited the patterns that could be etched. Specifically, lines tended to be straight, with right angles, following the crystal structure of the material [I83]. From the perspective of thermal isolation, this presented a problem because "*if you had straight legs with regular cross-sections, the surfaces look like mirrors to the phonons… they don't scatter… makes it a very good thermal connection*" [I83]. The normal strategies of making the legs longer or thinner didn't solve the problem, because they were still "regular cross-sections." So, the technologists tried various approaches, including what they called "*crazy legs*" (i.e., physically wiggling the legs "*so there's no line of sight*"), or leaving extra nubs of silicon [I83]. Eventually, the technologists became aware of a silicon texturing technique developed by a small company in California, whom they had met at a conference. As recalls CSE#6 "*that gave us something we could engineer*" [I83]. What he meant by that statement, was the texturing approach gave them a set of parameters that could be optimally tuned to solve the problem at hand.

Arraying Microcalorimeters:

Nonetheless, the challenge with all three of these approaches was that it added bulk to the support structure with implications for the array architecture. An important detector plane design parameter is the fill factor (the percentage of the field of view that is active). The goal is to position the individual pixels as close together as possible. This becomes challenging, when the detectors need to be thermally isolated, or read-out with detailed spatial resolution. Ideally, the support structure can be hidden behind the active part of the detector, but with the requirement for bulky, wiggly legs this proved practically infeasible with the existing technology.

As a result, the first generation of microcalorimeter arrays were one dimensional, with twelve pixels positioned side-by-side in a row [I83, I88]. This left room for the beams to protrude from either side. However, the science requirements demanded a wider detector plane, so the team explored multiple strategies for producing the equivalent of a two dimensional array. Among other approaches, they tried vertically stacking multiple 1D arrays so that they would appear 2D from above. Eventually, however, they realized that they could meet the AXAF science requirements with a so-called bi-linear array. This solution involved positioning two 1D arrays in opposite directions, so that all the support structure protrudes outside the field of view.

Shifting context: AXAF Split

During the mid to late 1980s, the team enjoyed a fair amount of autonomy, charged to solve some extremely difficult technical challenges. However, the level of political exposure that produced this shelter can also cause major shifts in the project development context. In the case of AXAF, the development context was abruptly changed in 1992 when, faced with budget constraints, the project was split in two. The first mission, dubbed "AXAF-I" would fly the imaging (I) instruments; the second mission, "AXAF-S" would be the spectrometer (S) (and host the microcalorimeter array). Although the AXAF-S mission idea persisted for another year, no one was terribly surprised when in 1993 AXAF-S was "*sacrificed*" in the budget. However, the language in the budget that year left the door open for XRS to be recast as an international collaboration [I74]. Shortly thereafter NASA signed an agreement with ISAS (the Japanese agency responsible for astrophysics before JAXA was formed) to this effect. Now, instead of flying AXAF-S as a stand-alone mission, NASA would provide the XRS payload to the Japanese as the primary instrument aboard the 5th Japanese X-ray Astronomy satellite, Astro-E [D26].

Ironically, the main impact of the transition from AXAF-S to Astro-E was to raise the intensity of the schedule pressure. As explains CSA#8, "*there was now a very certain launch date because where on the US-side, the AXAF program was being delayed year for year for a long time and then got split off into AXAF... The partners in Japan had a good reputation for once they chose a launch date, they really made it, so if anything, going to Astro-E just made people really realize ok, we have to make things work now*" [I84].

In addition, the shift in mission management to Japan actually gave more control to the science team. This statement may initially seem counterintuitive since the added layer of international collaboration typically rigidizes relationships; however, in practice, two elements combined to produce the stated result. First, the prevailing culture at ISAS pays deference to scientific judgment (compared to at NASA, where project managers typically come from engineering backgrounds). As a result, the Japanese managers were relatively more receptive to scientific rationales, and trusting of science-team derived schedules. Second since the instrument was formally managed by a Japanese project manager, the in house Goddard team was "*weaker*" than it might have otherwise been. Coupled with the fact that the original XRS project scientist was also Goddard's Director of Space Sciences at the time, the balance of power was shifted to science compared to other projects. The result was that instrument technology development continued much longer into the formal project than is typical [I74].

Evolving a Project-specific Development into a Research "Group" (early- to mid-90s)

Through the mid 1990s with the development of the flight detector array for Astro-E well underway, the team experienced a "changing of the guard" while also expanding, in terms of both personnel and technical approaches. This was partially an organic evolution and partially an active decision on the part of Goddard management to ensure a solid team of "in-house" experimentalists (to avoid a repeat of the need to sub-contract with a University Professor re: AXAF) [I88].

Through the first decade of microcalorimeter development CSA#2 had stayed intimately involved with the project. However in the early 90s two things happened which forced him to reduce his involvement; and consequently make room for the "new blood." First, the COBE satellite had been launched and was returning data; and the analysis work that entailed was keeping the IR Astronomers busy (recall that COBE was the application context for which the original thermodynamic noise theory was developed). Second, the head of CSA#2's group was promoted and as a result, the larger entity that he now managed was moved to a new building. The result was much less frequent interactions. As CSA#2 explains: "s*urprisingly, or maybe not surprisingly, a quarter mile of distance puts a fairly high potential barrier…"* [I88] This physical move corresponded with the first PI (CSA#12) taking over as Director of Space Sciences and handing off the project to CSA#8 (one of the post docs who had been hired in '84). While CSA#2 and 8 were close colleagues, their relationship was different than what existed between CSA#12 and the rest of the "*old guard*." The combination of factors resulted in CSA#2 spending less time in the X-ray lab. "*It was probably all for the best that I wasn't there as much because it created a little more space for the new generation… and they were good. I tried to keep my fingers in the stuff… but I wasn't working day to day"* [I88].

It was also during this period that the Goddard microcalorimeter group was formed. It's difficult to pinpoint precisely when the group became a group, but "*by 1995 we're certainly a group in that we have multiple [dedicated] people in our team, they're working on different things – we started to split up and specialize"* [I84]. In addition to the characteristics of more than one person, with different specialties, the microcalorimeter group began applying for and sharing resources to pursue new developments not explicitly related to the project that had originally brought them together [D2-24].

The Goddard team also became an acknowledged leader in their field. In the 1996 NRA, the evaluator noted that "*the proposing team [is] an experienced group with a proven track record…"* [D5] The team believes that being known made Goddard an attractive place to be: "*it helps define a group if you're known and there are people seeking you out in terms of job applications and post docs"* [I88]. During this period, two new full time scientists were hired as well as a compliment of engineers and technicians. The first scientist formally joined the group in 1992 as a post doc, although she had already been part of the "*inner microcalorimeter circle"* for several years. She had become connected with UL#3 and CSA#2 through her dissertation work on Compton Scattering. In fact, it was her need for higher energy microcalorimeters that had led the team to explore superconducting absorbers in the 1980s [I89]. She also brought a connection to Stanford University, which would become important in the next generation microcalorimeters. The second scientist had also taken a circuitous path to Goddard, and his past connections would also provide a link to the third generation of microcalorimeter approach (of lesser focus in this research).

The team branched out in a number of parallel technical development trajectories as well. They continued development of semiconductor thermometers for both hard X-ray spectroscopy as well as soft X-ray (suborbital sounding rocket, XQC) cosmic background scans. In addition, just as it seemed that a practical performance plateau had been reached on the semiconductor thermometer technology trajectory, the team became aware of a promising new approach based on superconducting transition edge sensors (TES).

In order to maintain the flow of the narrative, the semiconductor thermister pathway is told in its entirety first, with the TES story following.

Silicon Thermisters in the late '90s: prepping for flight

By the late 1990s, most of the team's silicon thermister microcalorimeter efforts were devoted to developing flight hardware for Astro-E. "*My recollection is that we were struggling to build something up until the last moment; trying to see what's the best we could fly*" [I90]. Their R&D efforts, on the other hand, had shifted almost completely to the new class of TES calorimeters.

> *We feel that calorimeters using transition edge sensors (TES) offer the most promise of meeting these requirements [for future missions], so we propose to concentrate our development efforts on such detectors* [D9].

However, at the turn of the millennium, two unrelated events – a launch failure and a decadal survey ranking – fundamentally changed the microcalorimeter development context, and gave the incumbent technology new life.

Launch Failure of Astro-E (2000)

In 2000, Astro-E should have been the first flight demonstration of the microcalorimeter spectrometer, and the last use of the silicon thermister microcalorimeters. However, the ill fated Astro-E spacecraft was destroyed 42 seconds after launch, due to a booster failure. The team spent less than a day "*mourning the loss"* [I69] before they began franticly scraping together a proposal for a redo. Fortunately, and unfortunately, proposals for the next SMEX call were due a week later. Miraculously, they made the deadline and received funding as a "mission of opportunity" (MoO). MoOs are designed to support payload's being developed for missions lead by other countries. In this case, the XRS team was bidding to be a NASA furnished instrument on an ISAS follow-on to Astro-E (if there was to be one). The Japanese simultaneously worked to convince their government that they needed to try again. Both were successful – there would be an Astro-E2 and NASA would provide the spectrometer; and that spectrometer would employ silicon thermister microcalorimeters [I68, I69, D26].

Constellation-X Ranked Second (2001)

Early in 2001, the 2000s decadal survey "Astronomy and Astrophysics in the New Millennium" was released and Constellation-X was ranked second (among major space-based initiatives), after the James Webb Space Telescope (JWST). Based on these recommendations, NASA initiated JWST as a formal program, and directed some money to Goddard for Con-X "pre-Phase A" technology development and mission studies. While the directed study provided a strong motivation to continue TES development, it imposed much less certainty than a formal program would have (leaving further room for a future resurgence of silicon thermister R&D efforts).

If Astro-E hadn't ended up in the ocean, things might have been different. If Con-X had been selected first in the 2000s Decadal Survey, things might have been different. But, Astro-E was destroyed on launch and Con-X was ranked second, so instead of demarking a clean shift from a silicon thermisters legacy to a focused, funded TES microcalorimeter future, the new millennium saw parallel, albeit mutually dependent, development paths for the two technologies.

A New Opportunity for Semiconductor Thermometer Exploration

The SMEX proposal that the team submitted in February of 2000 defined identical detectors to those that had been delivered for Astro-E. Given that the design had incorporated all of the latest development, there was no need to update [I84]. However, the project approval didn't happen overnight. It wasn't until 2002 that all of the funding had been approved, and the international agreement with Japan had been put in place. During the intervening time, "*we were busy trying to improve the state-of-the-art of these detectors*" [I84]. The team was funded by the same sequence of ROSES/NRA grants that had sustained them through the 90s. During the 2000-2002 window, they made three major improvements, bolstered by progress in related domains and a new team member with complementary experience.

Now more than a decade-and-a-half since the beginning of the microcalorimeter development, the fabrication lead CSE#6, had been promoted into a branch head position and been encouraged to take on a wider range of projects [I83]. His career had progressed as is common for engineering professionals, out of the technical into the managerial. His replacement, CSE#7 was already quite an experienced member of the DDL, and had been previously involved in the development of bolometers (the basis for CSA#2's original microcalorimeter idea). She immediately brought to bear some insights from that parallel development, as well as novel semiconductor fabrication techniques that had evolved in the industry at large [I87].

Thermometers: Eliminating 1/f noise and improving yield

Recall that the first generation of microcalormiter thermometers were made by ion-implanting silicon wafers as uniformly as possible. And, that although the desired resistance R(T) could be achieved by implanting the full range of doses and post-screening the resultant wafers (which was inefficient in terms of yield), for any given resistance, variability in the depth distribution of the dopant created a so-called "1/f noise" term (reducing the sensitivity of the measurements).

During the second window of exploration, several attempts to address the noise challenge were made. One was conducted as "the deep doping DDF." The goal was to achieve the desired depth distribution uniformity by incorporating the doping into the growing process itself. They worked on it for a year, but were never successful. At the time, the process of growing silicon wasn't well enough controlled to "*count the right number of atoms going in as they did the growth*" (which would have been required) [I87].

The approach that eventually became standard involved the adoption of a new wafer material; it solved both the 1/f noise and yield issues simultaneously. The material was called Silicon on Insulator (SOI) and consists of a thin layer (1μm) of Silicon separated from a bulk layer of Silcon by a thin oxide layer [I87]. The idea came from a group in France, as will be discussed in more detail below (in the section on mechanisms of adoption) [I88]. The oxide layer served as a diffusion stop. Now, when the implanted silicon was annealed, the dopant diffused uniformly within the limit provided by the oxide layer. This eliminated the 1/f noise term almost completely, effectively "*putting us on a different curve*" [I88] with respect to noise reduction.[31]

With the new material, the yield jumped from ~3% to nearly 100% almost overnight [I84, I87]. They were still implanting over the full range, but now, nearly all the devices that you would expect to work did, so the screening was easy [I87]. Combined with the fabrication improvements discussed below, Astro-E2's energy resolution was twice as good as the array flown on Astro-E (6 eV compared to 12 eV) [I88, D43].

<u>*Detector attachment: crazy legs revisited*</u>

In the same way that SOI wafers had a built in diffusion stop, the oxide layer also served as an etch stop. Coupled with the semiconductor industry-wide transition from wet etching (and the corresponding need for straight lines with right corners) to DRIE "deep reactive ion etching" (which enable complete pattern flexibility), many of the challenges faced during the Astro-E development were now easily overcome.

From a fabrication perspective, as recalls CSE#7:

> *When I took it [Astro-E2 detector fab] over, the design looked very much like it would have looked for the bulk silicon devices. There were straight legs, square corners, all sorts of things that you have to have for bulk silicon devices… and I said, ok, but we have lots of things we can do with SOI and with dry etch techniques that we've learned with microbolometers… we can make curved structures, we can make strain relievers… we can do all kinds of things. We can fix problems that had come up during Astro-E…* [I87]

From the perspective of the science team, each step in the transition for straight lines with square corners had a purpose, and performance needed to be carefully characterized before the move to crazy legs could be made. As explained by CSA#10:

> *The set of test devices made to first characterize the SOI technique for x-ray devices was also designed to help us quantify thermal conductance in the very thin silicon... the series of devices with straight legs of different widths and lengths was to help us characterize the scaling of thermal conductance…The move to bent legs also had to be accompanied by mechanical analysis (to make sure the resonant frequencies of the structures were acceptably high.)*

In the end, the new techniques allowed for significantly more compact support structures and essentially eliminated the array "fill factor" problem discussed above. The weak link could now be hidden completely behind the absorber material, with the detectors packed as tightly as desired in a real 2D array.

---

[31] The team had previously been stuck in a "*square root regime*" (i.e., if you wanted to make the devices quieter, they needed to be bigger, which, in a small area meant thicker. But, when you increased the volume, you increased heat capacity, which reduced potential energy resolution (nearly defeating the purpose of the volume increase) [I84, I88].)

Absorber attachment: a slightly less painstaking process

Of course the need to attach each absorber by hand persisted, and continued to limit feasible array size. However, even though the new techniques couldn't eliminate the manual post processing step, they could at least incorporate the need for a spacer into the silicon design (see Figure 3-11).

> *Once they told me how they have to glue the absorber onto the calorimeter and they told me they had to put like a little silicon chunk in there to stand it off from the thing, I said well wait a minute… I've got a way to do that much better, make your little spacer out of this material I've been working with.* [I87]

At first the scientists were skeptical about making a design change that didn't directly improve performance. In fact, they carried the two development paths for some time. It may have partially been a conscious, or subconscious, boundary contest about where scientists or technologists should have relative design authority. Either way, once the scientists had the opportunity to verify the performance of the embedded spacers "*they found that the SU8 spacers were much better than having the silicon spacer, they adopted that as well.*" "*There were numerous changes that were adopted from things that I'd learned working on other projects.*" [I87]

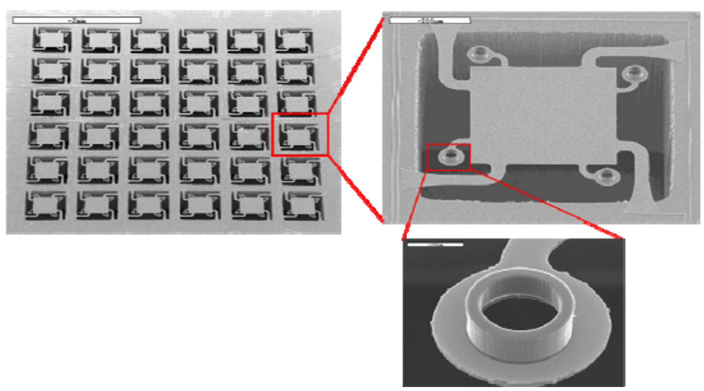

Figure 3-11: Close-up of Absorber Spacers, Etched in SOI

Mechanisms of adoption, on the process side

Between Astro-E and Astro-E2, the team achieved a resolution improvement of a factor of 2. These improvements were due to process improvements in the way they manufactured the device (not what they manufactured – i.e., they were no longer in a "*technique limited regime*") and were enabled by innovations in the related semiconductor industry, imported to the microcalorimeter development as part of the tool kit of a new addition to the team. In fairness, Goddard (and the previous technologist) had already begun to make the transition to DRIE etching techniques, but they were using them to produce better versions of the same long straight legs, rather than exploring the full range of new microstructures that could now be achieved.

The idea to use SOI wafers for microcalorimeters was initially suggested to the X-ray group by CSE#7, and came indirectly from CSA#2 and the parallel bolometer work. He had become aware of the relevant material and technique as a "silver lining" to a failed bolometer instrument proposal bid. During the revue, CSA#2 noticed that the winning detector concept, developed by a group in France, was achieving better performance than should have been possible. He knew their volume, and given Goddard's scaling laws, they should have been too noisy. Yet, the results clearly showed that they weren't. "*They were doing better than they had any right to… that led me to believe that they had a fundamental advantage.*" [I88] Upon deeper inspection, the "fundamental advantage" was a far better implantation technique; they were using the diffusion in SOI technique described above.

It's worth noting that the technique wasn't a secret because of any active IP (intellectual property) protection on the part of the French company. It just so happened that "*the head engineer was a Frenchman who didn't speak any English, so he wasn't going to present anything, and the higher ups who spoke English didn't have the detailed expertise to know the most important part of the design… so the information just wasn't coming out.*" [I88]

The proposal competition enabled CSE#2 to see the results first hand, and he brought the technique back to Goddard. As he puts it: "*I've always been a big fan of copying anything that looks like it's going to work; because we don't need to, or want to, invent these parts. If you can get it from somewhere else you should get it from somewhere else.*"[I88] Funded by a DDF "the ultra thin bolometer array DDF," CSE#7 (the Astro-E2 fabrication lead, who at the time was CSE#2's lead technologist on the IR microbolometer project) began investigating the application SOI, and in particular, the diffusion techniques. The original intent was to make bolometers on a nitride membrane. As recalls CSE#7: *"I was supposed to back-etch the silicon and leave the free-standing nitride structure. But I didn't do what I was told to do… I removed the oxide and nitride and left the free-standing silicon…and I said look at this stuff this is beautiful material"* [I87] The ultra thin, flexible yet strong, single crystal, completely planer structure that remained was exactly what they needed. CSE#2 agreed. "*That little DDF many years ago enabled all of this. All of the bolometer arrays and the first microcalorimeter arrays that were done with ion implanted SOI membrane.*" [I87]

Throughout her time on the microcalorimeter project, CSE#7 continued to work on the microbolometer development as well. Architecturally, IR microbolometers and X-ray microcalorimeters are very similar, so from a detector fabrication perspective, there is a lot of synergy. However, since they serve distinct science communities, cross-pollination of solutions is not always achieved as often (or as soon) as one might hope. As fabrication got more sophisticated, particularly as both microbolometers and calorimeters transitioned to TES, a group called the "TES roundup" was established to facilitate knowledge transfer. The idea was to bring all the scientists and technologists together on a monthly basis to discuss common challenges. During the early stages of the development, attendance at these meetings was high, but as the crunch of flight projects hit, the meetings occur much less frequently and attendance is quite poor. Told from the perspective of the technologist:

> *[W]hen these things were in their early stage of development, those meetings were a very useful tool to get the scientists in the same room, talking to each other… Now the only way they talk to each other is if they have a problem and you convince them… 'maybe the X-ray group has seen something similar, you could call them?'* [I87]

Implicit in the above quote is an expertise hierarchy between the science and technology communities at Goddard. While the scientists certainly have a lot of respect for the technologists' expertise in fabrication, that expertise is somehow seen as a lower art. Technologists make design suggestions, and Scientists make decisions. Technologists make the devices, Scientists tests and verify them to make sure they meet the science goals (although this is not the case for all instrument developments). In addition to the above quotes, the below illustrate this relationship:

> Technologist: *"We're the machine shop. The X-ray group are experts in their field, world leaders and all that, and so they have a very clear idea of what they want and how to get it"*
>
> Scientist: "*I became aware that there was concern on the detector systems branch, that they weren't getting as much credit … so we started making sure that they were always on our papers."*
>
> Technologist: *"We're technical partners in this. They're the scientists and we're the technology people."*
>
> Scientist: "*It's a long standing collaboration [with the DDL]that's been very fruitful… we're working together on figuring out how to make these things work, not either saying these are the science requirements, now you design it, nor this is the design, now you make it… they're part of the group, they're part of the effort"*
>
> Technologist: "*The way things work here at Goddard, even for technology development you really need a scientific justification for it, so you always want to have a scientist on board with your proposals if you expect them to win. We've got a lot of great ideas, but if it doesn't have some kind of bearing on the science community then it's not going to win.*"

Con-X Delays: A Third Opportunity for Semiconductor Microcalorimeter Exploration (2003/4)

Although Astro-E2 (like Astro-E before) had been intended as the last flight of silicon thermistor microcalorimeter technology, funding uncertainties in the Con-X program gave the semiconductor development program a (second) new life. In 2003/4, NASA-level funding "reallocation" prompted the team to hitch their technology program to a nearer term opportunity, while diversifying their development efforts. This logic was articulated in the 2004 proposal excerpt below:

> *When we wrote the predecessor to this proposal in 2001, our primary goal was to implement TES microcalorimeters in the format required for Constellation-X. While we have made impressive progress towards that goal, **the goal itself is shifting**. With the extension of the Constellation-X time line due to reallocation of the NASA budget, changes to the reference design are already being discussed. **The delay to the Constellation-X program** (first launch now projected for 2017, while once we were working towards 2010) **make NASA participation in an intermediate term x-ray spectroscopy mission, such as the Japanese New X-ray Telescope (NeXT)** [Tawara et al., 2003], **a strategic investment.** We propose to continue to develop arrays of x-ray*

> *microcalorimeters, no longer towards a single, specific spaceflight implementation, but rather towards several different strategic optimizations. These include the original Constellation-X baseline, arrays of the larger but slower pixels needed for NeXT, optimization of arrays with more imaging elements as needed for some reconfigurations of Constellation-X or vision missions such as Generation-X and MAXIM, and optimization for the soft x-ray band. The end product of our research will be a robust technology that can be readily optimized to address a variety of questions about the origin and evolution of the complex structures of the Universe.* (emphasis added) [D14]

Compared to Astro-E and E2, the NeXT science objectives would require a much larger detector plane. This forced the team to begin addressing the "by-hand" scalability limitation of the HgTe absorber attachment (even though this was no longer a long term concern with the TES approach). Increasing the number of detectors also resurfaced some of the heat dissipation issues discussed earlier. A number of technology development efforts were initiated in preparation for a NeXT bid.

<u>*Automated Absorber Attachment*</u>

Recall that the state-of-the-art in absorber attachment involved carefully dabbing four dots of epoxy, under a microscope, onto four attachment points on each detector and then carefully positioning the HgTe absorber. For the 6x6 Astro-E/E2 detector arrays, this process was tedious. For the 16x16 array that would be needed for NeXT, it verged on impractical. For the 1000+ pixel array planned for Constellation-X, the idea seemed insane. So, the DDL began investigating methods to automate the attachment process. Within a year, they settled on a stamping approach. "[W]*e found a way to automate it with stamping. We'd stamp into the glue and then stamp into the absorbers; let it cure; pick them up; and then the absorbers were stuck to the detector.*" [I87] Now instead of the number of post fabrication steps scaling with the number of pixels in the array, the extra steps were fixed at one. However, this technology was never used, due to another context change (described below). Nonetheless, the capability to attach large numbers of "things" to arrays of "things" is a valuable, generally applicable, tool that will certainly be used in the future [I87].

<u>*Low temperature, thermally isolated electronics*</u>

A fundamental challenge of implementing low temperature detectors is that the detector needs to be kept cold (in this case <70 mK), while being readout by electronics that both produce heat and function poorly at cryogenic temperatures. Specifically, as the number of pixels in an array grows, so do the number of readout channels, and by proxy, the magnitude of the heat produced by the electronics. The cooling challenge can be addressed both by improving the cooling power of the cryogenic system (see CADR case) and improving thermal isolation of the electronics.

Both semiconducting bolometers and microcalorimeters use Junction Field Effect Transistors (JFET's) as the first elements of the detector readout chain. The challenge is explained clearly in D20:

> *For the Japanese New X-ray Telescope (NeXT), the pixel count will be increased by more than a factor of 10, and yet the array and readout circuitry must still fit within roughly the same envelope as XRS. Since the JFET's operate at temperatures substantially*

warmer than the detectors, complicated mechanical methods of thermal isolation (tensioned Kevlar and hand-soldered 0.5 mil wire) have been necessary to operate the transistors without loading the detectors with heat or dumping too much power to cryogenic systems. The existing transistor circuit packages will be too large to scale up to 400 detectors, so we need a way to make JFET assemblies that also incorporate thermal isolation as well as overall miniaturization and large-scale monolithic fabrication. By making devices with inherent thermal isolation, we stand to gain large savings in volume and power dissipation, thereby making much larger arrays of semiconductor bolometer and x-ray microcalorimeters possible, and with high uniformity and reliability.

Funded by three years of DDF (2001-2003 and 2005), low-noise JFET module-on-a-chip's were fabricated on SOI wafers. However, the project was abandoned before it reached maturity due the following context change.

Another Failure

The second re-birth of silicon thermistors was further crystallized by another failure. In 2005, the second opportunity to obtain high resolution high throughput X-ray spectroscopy also ended in disappointment. Although Astro-E2 launched successfully, the cooling system failed 2 months into the on-orbit check-out, rendering the cryogenic microcalorimeter system useless [D44]. They just couldn't catch a break. And this time a follow-on opportunity was not immediately available.

Playing it a little more conservatively: Astro-H instead of NeXT

When the 2004 ROSES proposal (D14) was written, NeXT was evoked as a potential near term application; an R&D justification for diversifying the group's technology development efforts to mitigate against the uncertainty associated with delays in the Con-X program. After the 2005 Astro-E2 failure, the nature of NeXT as a flight opportunity changed in several ways. First, it seemed more certain that NeXT would fly in some form, since both governments would want to capitalize on all the previous investment. At the same time, mission success became even more critical; the third time had to be the charm. Second, it reduced the threshold of what would constitute sufficiently new science. All the arguments for why Astro-E and E2 could answer important science questions still applied, and now NeXT could answer them for the first time, rather than needing to build on those discoveries. Combined, this meant that for the post-Astro-E2 failure NeXT, hitting a low price point was more important than advancing the frontier of the possible. In practice, this translated into reducing the active pixel count to 32 and reusing hardware that was originally developed for Astro-E2.

From the perspective of the scientists, this was a strategic decision:

When we wrote our NeXT proposal… there was always all kinds of strategy type things going on, about how much can we afford to ask for and still be competitive… and the thinking was that we should try to keep our budget to about half of the total budget that was available so that NASA will be able to pick more than one Mission of Opportunity (MoO). If we tried to grab all of the money, that might be too much of an all or nothing.

> *So we were trying to keep our proposal to a certain level… once we did that it precluded the larger array.* [I84]

From the perspective of the technologists, it was demoralizing to say the least. As expressed with varying levels of political correctness[32]:

> *They won the proposal, but then it appeared that they had to back off on the number of pixels that they were going to be able to fly, so all those new technologies that we developed weren't going to be used.*
>
> *I think that our branch would have been happier, after all the work that we did, if we'd have been included in the flight proposal. If you go back through the development work, and the strong arguments that lead to them winning the proposal, much of that work was done here. And then to have us be the only ones at Goddard to be cut out of it, at Goddard?… I think that was a little bit short sited. I know why they did it, and I know that they were instructed to do so by higher ups, and I know that their main driver was to win the darn proposal, but…*
>
> *It was a big hit to the Branch to lose this valuable flight work. But then TIRS came along and the Branch focused our resources on that, so we managed to stay employed!*

The strategy was successful. Three years after the Astro-E2 failure, in 2008, the calorimeter group won their second MoO to develop SXS (Soft X-ray Spectrometer) to fly aboard Japan's 6th X-ray astronomy mission – Astro-H. The launch is scheduled for 2014, 30 years after this innovation pathway was initiated.

Concluding comments

Assuming that Astro-H returns data successfully, and now that TES is fairly mature and is markedly better for larger field of view applications, it is unlikely that semiconductor-based microcalorimeters will be used again. Nonetheless, it is interesting to note the extent of the performance improvement that occurred between the AXAF baseline and what will eventually fly on Astro-H. Given that TES were invented, and projected to be vastly superior as early as 1996, had political and technical context changes not perpetuated the old technical approach it is unlikely that these process-enabled performance improvements would have been realized. Today, the performance differences between Thermistors and TES (at the individual detector level) are negligible. The key difference is one of scalability – TES absorbers can be fabricated monolithically, and the readout electronics are lower power with better isolation. However, had the tradeoff between improving the semiconductor approach vs investing heavily in the TES approach been articulated in the above terms, the conclusion to transition would likely have been different. This is not to say that they made the wrong decision, rather to point out the extent of performance improvements that were achieved through process refinement, after the first flight instrument.

[32] These quotes are not attributed at the request of the informants.

Table 3-4: Semiconductor Microcalorimeter Shock Sequence

Epoch	Start	Component Exploration	Architectural Exploration	Exploitation	Treading Water and Branching Out	End
Start			* X-ray observatory ranked 1st in Decadal * X-ray folks look to CSA#2 for help on next generation spectrometer * CSA#2 has architectural insight (bolometer as calorimeter)			
Component Exploration			* Good enough result shifts focus to architecture	* Astro E2 selected as MoO * Astro E2 fails on-orbit * NeXT recast as Astro H with new priorities		
Architectural Exploration		* Concept demo reveals technical challenges (absorber material etc..) * NeXT spec drives new component requirements		* Mission pressures force focus on array optimization		
Exploitation					* AXAF slit creates uncertainty in "S" portion	* Astro E2 approved for formal development * Astro H approved for formal development
Treading Water and Branching Out	*Turnover in group *TES insight					* AXAF-S reincarnated as part of Japanes Astro E
End		* Astro E launch failure * Con-X ranked 2nd	* Con-X delays * NeXT opportunity on horizon			

3.5 The Transition Edge Sensor (TES) Microcalorimeter

Gestation (pre-1996)

In a sense, all of the previous decade of semiconductor microcalorimeter development played an important role in the gestation of the TES microcalorimeters. By the mid 1990s, the Goddard calorimeter group had emerged as a recognized leader in the field, and as a result, had continued to attract top talent from, and established strong connections with, most of the leading institutions and groups around the country. In addition, an enabling infrastructure of fabrication and test facilities was already well developed and staffed within the DDL. However, while this network and these complimentary assets were important enablers of the TES development, in many respects, the new TES technology, which was invented outside of the Goddard microcalorimeter community, was competence destroying.

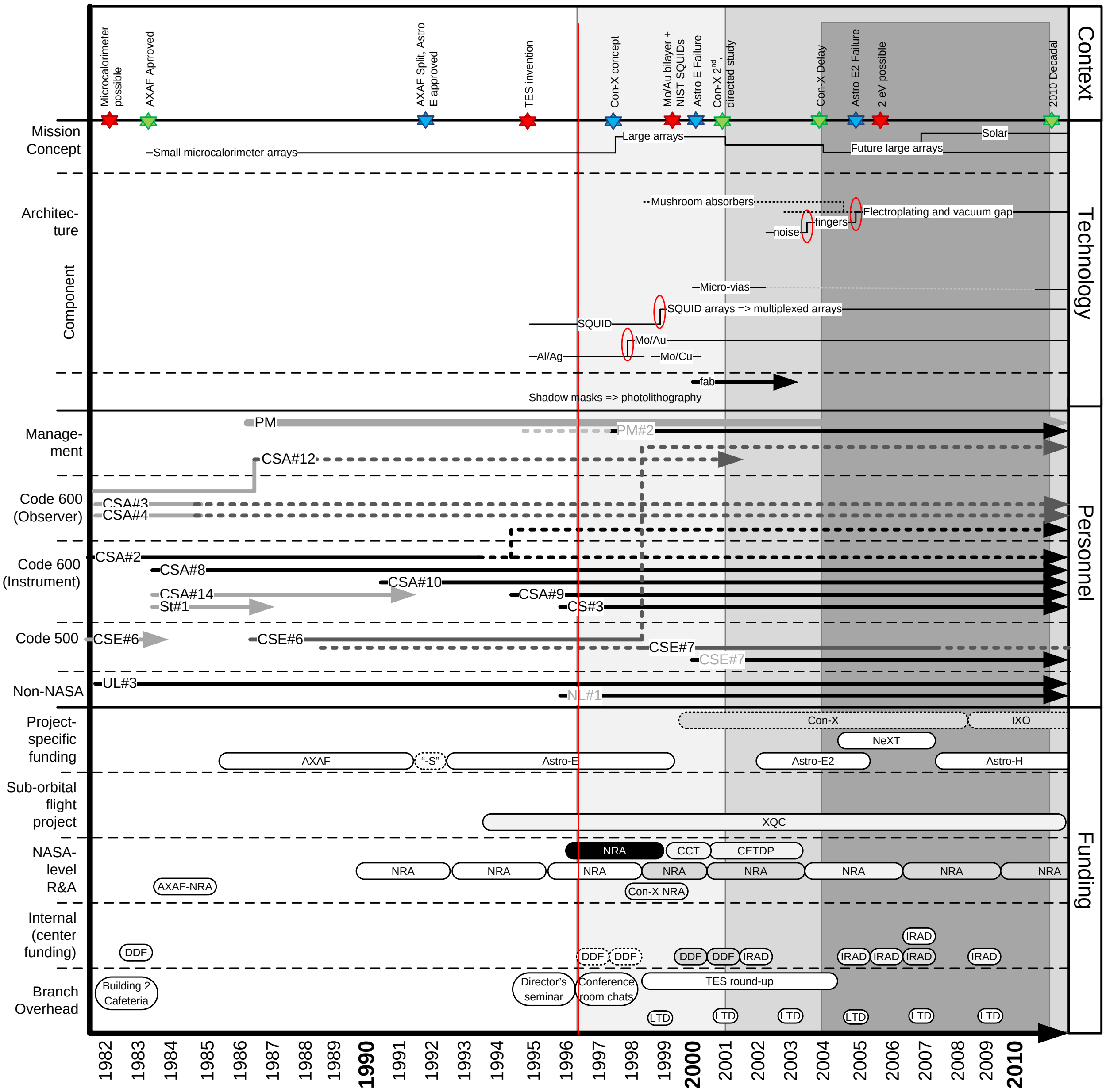

Figure 3-12: TES Microcalorimeter Visual Map

Goddard was made aware of the promise of the new TES technology during a colleagues' seminar, given at Goddard in 1995; the speaker was a professor that CSA#10 had known during grad school. "*Everything is connections. Because he knew me and [CSA#2] he visited Goddard and wanted to catch us up on the latest*" [I72] He shared recent results from one of his grad students' work on TES that looked very promising. At least in part because of their shared pasts, CSA#10 took the claims seriously on their face and began investigating the new area. The results did indeed look promising. Although the theoretical limits of TES sensitivity[33] were similar to those of the ion-implanted silicon thermometers, the new approach had two key practical

[33] As part of his dissertation, the grad student extended the 3M calorimeter theory to include superconducting devices and predicted a theoretical limit of ~1eV energy resolution [D38].

implications that made the prospects of reaching the professed theoretical limits more reasonable.[34]

Superconducting and semiconducting microcalorimeters both perform non-dispersive spectroscopy by measuring temperature changes in an absorber, due to the thermalization of X-ray photons, as changes in resistance. The main difference is that superconducting thermometers permit a higher heat capacity budget since they are extremely sensitive thermometers in their transition regions. Firstly, this meant that the low heat capacity constraint that had driven the selection of HgTe as the absorber material, no longer applied. Now normal metals could be considered, which thermalize deposited energy quickly and efficiently. This also had huge implications for manufacturability and consequently scalability of the array field of view. Where each HgTe absorber needed to be attached by hand, now absorbers could conceivably by deposited as part of the fabrication process. The second key implication became more relevant later in the development. Where ion-implanted thermisters were readout using JFETs (which had heat dissipation issues and didn't multiplex well [D20]), TES devices used SQUID (superconducting quantum interference device) readouts, for which these challenges were less of an issue. "*So we started getting into the field.*" [I72] In fact, TES was already called out as an important area for future investigation in the 1996 X-ray calorimeter NRA [D5].

By 1996, the grad student mentioned above (heretofore referred to as NL#1) had graduated and joined the National Institute of Standards and Technology (NIST), to apply the TES technology, that he had invented, to semiconductor fabrication. However, thanks to the presentation at Goddard, one of his early projects involved a collaboration with CSA#2, to develop arrays of "pop-up"[35] TES microbolometers [D46]. The work was funded, for three years, culminating in a failed proposal for the ESA-led Plank mission [I90]. Although they didn't win the bid, that three years of work laid the foundation for much of what would follow. Firstly, the proposal allowed NL#1's to hire his first post doc (who would later become CSE#8). CSE#8 spent 3 years, from 1997-2000 at NIST developing SQUID multiplexers and TES bolometers. When his post doc ended, CSA#2 suggested that CSE#8 join Goddard and found him a job in the DDL. The move guaranteed that a strong relationship would be maintained with NIST[36], as well as CSE#8's tacit knowledge of their processes and technologies. Second, it justified Goddard's investment in the deposition equipment (in 1997, ~$100K) that would enable future in-house fabrication. Third, it likely lowered the barriers to entry for the X-ray group, since the DDL had already begun ramping up the infrastructure to support the IR group (and their TES bolometer efforts). Finally, it matured the SQUID multiplexing technology to a point where the readout didn't limit device performance. In the words of CSA#10: "*That was a breakthrough that was a necessary development for TESs to get where they've gone. If SQUID arrays had not already been invented, TESs would have needed them to be invented.*" [I89]

---

[34] *At the time, there was a bigger difference in the resolution expected for TESs vs. Silicon… they hadn't quite converged as much as they have now* [72].

[35] The term "pop-up detector" is used to describe a strategy for tightly packing arrays, which involves folding bolometer the legs (weak link) so that they are hidden behind the pixel area.

[36] Incidentally, CSE#8's educational background was a PhD in thin-film physics from Brown.

Early developments (96-99)

Although Goddard began exploring TES bolometers and microcalorimeters in 1996/7 it took several years before the X-ray group became competitive in this new realm. Their leadership in silicon thermistor microcalorimeters didn't translate directly to the new TES devices. The team didn't emerge as "*real players"* until 1999/2000. *Up until that point, we were just considered people who were dabbling in TESs. Suddenly we were on the map."* [I89] The 1999 NRA reflectively attributes the early challenges to a combination of competing priorities, a steep learning curve and an enabling breakthrough:

> *In the last three years, while completing the XRS instrument and flying XQC, the Goddard calorimeter group has been acquiring expertise in TES technology specifically and superconducting electronics in general. This has been achieved both through acquiring personnel and through the lessons of practical experience. Hard lessons learned with the Al/Ag system led to the development of Mo/Au TES bilayers.* [D9]

One of the "acquired personnel" was CS#3 who, at the time, was working on tunnel junctions as part of a post doc in Europe. He knew CSA#9 (the second scientist to join the microcalorimeter group in the mid 90s) from related work during CSA#9's post doc. When it became clear that the team needed someone with prior cryogenic superconductor expertise and a propensity for fabrication work, CSA#9 suggested CS#3; called him up and invited him to apply to Goddard. CS#3 accepted the invitation, and following a open competition, begun work on bilayers in 1996 [I94].

<u>*A New Materials System: Switching to Molybdenum in the Bilayer*</u>

The Early NIST TES devices used an Aluminum-Silver (Al/Ag) system, and produced pretty good results. However, despite complete knowledge sharing (CSA#10 even spent a "sabbatical" at NIST as a visiting scientist) and significant effort, neither Al/Ag nor Al/Au TESs were ever made to work at Goddard. NIST's fabrication "*was complicated and required some specialized shadow masks to do the films in-situ without any etching involved."* [I91] Goddard wanted to leverage the more sophisticated photolithography techniques employed in the DDL, but the devices kept corroding [I94]. At the time, CSA#10 and CS#3 believed that the problems with the shadow masks may have been due to differences in humidity in Colorado vs. Greenbelt. They later realized that aluminum bases systems were incompatible with the chemistry of photolithography, and began looking for a new material system. Even if the shadow masks approach had been made to work at Goddard, the strategy isn't suitable for close packed arrays with precisely defined features, so they would have had to switch to photolithography eventually anyway [I89. I91].

By 1998, the Goddard team had abandoned the NIST material system in favor of Mo/Au TES bilayers. "*Once we made that jump, then it was almost like it was a matter of time before we'd start getting good results."* [I89] This was right around the time that other TES groups started doing the same. There is some disagreement among the Goddard team about where the idea came from. CSA#10 remembers sitting down with CS#3, who had

> *pulled together a bunch of alloy tables and recommended that we use a Molybdenum-based system… I don't think that we were ever copying. I think that this is actually a case*

where if you sit down and look at the chemistry and you know you need a superconductor with a T_c in a certain range, and a good robustness to diffusion and intermetalics… I think you end up with a molly-based system. There are some people who think that it may have filtered into the discussion through some side channels that I wasn't aware of. [I89]

According to CS#3, the choice of Molybdenum was strongly influenced by his past experience working with "*Tantalum from when I was doing tunnel junctions.*" That led him to consider "*refractory metal (it's much more robust)*" and of those "*the closest T_c I found was actually molybdenum…*" [I94]. He began testing the molybdenum with the equipment within the DDL and by '98 had produced Goddard's first Mo/Au bilayer.

Based on similar rationale, NIST chose Mo/Cu (copper) and Goddard chose Mo/Au (gold), which they both still use today. There are pros and cons of each. Mo/Cu oxidizes when exposed to air, and the Mo/Au can create an intermetalic that is self-superconducting at certain temperatures. The former can be mitigated through surface treatment. The latter hasn't caused problems in the relevant operating range. Goddard's detectors are currently considered superior in the context of X-ray Astrophysics, but the distinguishing characteristics are not directly attributable to the choice of Mo/Cu.

<u>Learning to Use Superconducting Electronics</u>

It took another couple of years between Goddard's switch to Mo/Au and their first good test result. While working with any new material requires some amount of getting-up-to-speed time, the real hurdle involved the transition to superconducting electronics. The first TES paper had stated that the devices would need to be read out SQUIDs, and so that's what they did. However, unbeknown to them until around 1999, the commercial devices they had been using weren't capable of measuring a good resolution even if they had had a good detector. It turns out that the conventional "flux-lock-loop" available at the time just wasn't, and never would be, able to handle the fast slew rate characteristic of the fast TES transitions [I89].

In 1999 Goddard began purchasing custom TESs directly from NIST. Once they had the NIST SQUID series arrays, they converted the old dilution refrigerator (the one that CSA#2 bought in 1979, that had done all of the silicon thermister bolometer and calorimeter work) to work with SQUIDs. "*When we got that up and running we had a good test platform for our TES calorimeters… In 1999 we made even better devices and had the set-up to actually test them… I remember it was in the spring 2000 because we were doing the testing right after Astro-E ended up in the Ocean…it was kind of consolation*" [I89]

These "December '99 Devices," tested in 2000 and published as [D36], put the Goddard calorimeter group on the TES map. The results also legitimized the funding stream within Goddard. Where the '96 NRA reviewers indicated that "*the scope of the program as described in the proposal does not warrant such a large increase…*" in reference to the TES branching out, and the '99 NRA was only funded for two years because it was deemed overly optimistic and unfocused, by 2001 (after the strong results) the TES program was seen as worthwhile in its own right, and has been separately funded at ~$1-3M on a yearly basis since [D2-24].

A Next Big Mission on the Horizon (97-2000)

In discussing the TES microcalorimeter innovation pathway, it is impossible to divorce the technology from the future program concepts that were being floated around Goddard at the time of its inception. The (still) future X-ray Observatory, currently known as the International X-ray Observatory (IXO) evolved as a merger of the U.S. Constellation-X (Con-X) and the European X-Ray Evolving Universe Spectroscopy (XEUS) mission in 2008 [D45]. Con-X's initial call for instrument concepts was formally released in 1998 in preparation for the 2000s decadal survey, but the initial formulation took place informally a couple of years before that; around the time that the TES concept was initially being explored.

As CSA#10 remembers it *"[a colleague] who at the time was trying to get a mission concept together, said we need something with a spectral resolution about 2eV and collection area of […] can your calorimeter do that?"* [I89] Her initial response was a non-committal "*maybe? Theoretically it could…"* [I89] Recall that the theoretical limit worked out in 1982 by CSA#2 et al. was 1eV; however at the time, the best single-pixel resolution that had been achieved with any calorimeter was ~7eV. CSA#10 recalls that the discussion evolved through a series of informal gatherings, discussions in the conference room about what was possible and what it would take to get there. They concluded that 2eV was attainable, but would require a major investment in the new technology. "*Although we were already looking at the TESs as something that we should be developing, we hadn't gotten very far with them in those early days."* [I89]

When it came time to submit a proposal for the Con-X spectrometer in 1998, Goddard submitted "A Comprehensive Approach to Developing a 2 eV Calorimeter Spectrometer for Constellation-X." [D6] As described in the proposal:

> *The scale of this array, coupled with [the technical requirements] represents a substantial extension of the current state of the art in x-ray detectors. To accomplish this goal with a maximum likelihood of success, we have assembled a team composed of three leading groups in x-ray microcalorimeters. […] We propose to follow parallel paths of investigation and to develop several technologies critical for the support of all or most of these paths.*

Two points in the above quote merit further discussion. First, the qualitative density of the microcalorimeter network is noteworthy. The three groups on the proposal are Goddard/University of Wisconsin-Madison (UL#3), the National Institute of Standards and Technology (NIST) (where the Stanford grad student moved upon graduation NL#1), and SAO/LBNL/Brown University (CSA#9's and CSE#8's alma mater). Second, the difference in assumed credibility between the group that proposed microcalorimeters in 1984 for AXAF, and the one that proposed TES for Con-X in 1998, is remarkable. Where the 1984 proposal emphasizes the mathematical rigor of the calculations and suggests that maintaining a backup plan is a wise choice, the 1998 proposal focuses on the proven track record of key individuals. While parallel paths are laid out, they are presented as an appropriate research risk diversification strategy, not a back up (in case we fail to improve the state-of-the-art by a factor of 3).

The ambitious program of technology development was approved in support of Constellation-X.

> *At the time of the selection of technologies to develop for Constellation-X, however, resistor-based calorimeters were the clear front-runners, with ion-implanted Si, neutron transmutation doped (NTD) Ge, and superconducting transition-edge calorimeters each having attained resolution of about 7 eV at 6 keV. The nearest contender, the electron tunneling normal-insulating-superconductor (NIS) calorimeters developed at NIST, had achieved at best 22 eV at 6 keV. The long term potential of the other calorimeter technologies ought not to be overlooked, however. The paramagnetic calorimeters, in particular, may deserve a later second look.* [D35]

The technologies were selected for funding, by an independent peer review at headquarters. In the years leading up to the 2000s decadal survey, multiple technologies were carried, with the goal of making as much progress as possible before a down-selection was required. Although the funding was supposed to be for three years, "*we [the project team] funded them for a period of time, probably never at the level that we said we would because we never got that level of funding [from HQ] because it's always gotten cut…"* Within a year *"some of the areas kind of dropped out, and some of the areas were kind of getting funding from elsewhere, or weren't as critical."* [I57]

Nonetheless, the three groups – Goddard/UW, Lawrence and NIST – formed an integrated product team and created their own project stability. They outlined a clear roadmap for achieving the required 2eV energy resolution [D35].

Constellation-X: The Impact of Not Being First (2001)

However, Con-X was not ranked first, as they had hoped. Early in 2001, the 2000s decadal survey "Astronomy and Astrophysics in the New Millennium" was released and Constellation-X was ranked second (among major space-based initiatives), after the James Webb Space Telescope (JWST). Based on these recommendations, NASA initiated JWST as a formal program, and directed money to Goddard for Con-X "pre-Phase A" technology development and mission studies. During the decade of the 2000s, Con-X was funded between \$4M and \$12M per year, most often in the \$6M-\$10M range. Of that 70-80% has been dedicated to technology development, but most of that went to the mirror group. While this may seem like a substantial investment, it hasn't felt that way for the people involved. Con-X is an extremely ambitious mission, with breakthrough performance improvements required across multiple technology areas. In addition to the technology development resources having been spread thinly, significant budgetary uncertainly has severely limited project management's ability to plan. As described by the project manager:

> *We've had fits and starts a) because of the money being erratic (for several years we've had midyear budget cuts that are really drastic… and then b) you have to reduce your staff… and it's hard to build it back up again. An example is that we had this really good mechanical team working on the mirror* [mounting… But he] *got called off because we're low priority…* […, but] *technology requires the right kind of person, and it just takes time to get that person up to speed. It's not like we're just designing [*something that we've done before*]… so we got another mechanical engineer, and he wasn't a good fit… then we got another one and they worked together, but still not a good fit.. and then*

> *we finally got someone else… and we were progressing, but really slowly… eventually we realized that we just had to get the right person, but it was really like 2 or 3 years that we made no appreciable progress on the mechanical mounting aspect.* [I57]

The established stability of the calorimeter group served to mitigate the effects of budgetary uncertainty on the TES development. While the effort was initially funded through the Con-X project, it also received significant and steady resources from APRA (Astronomy and Physics Research and Analysis – the ROSES program element dedicated to Astrophysics technology development), in three year tranches. During the decade of the 2000s, the calorimeter group never got more than 50% of their TES development funding from the Con-X project "*mostly it was more like 10%*" but their status as "*recognized experts*" developing a "*key enabling technology for the next grand x-ray observatory*" certainly factored into the attractiveness of their sequence of APRA proposals. Although the Con-X project team had no formal influence over the peer-reviewed APRA funding allocation, there was some level of informal coordination:

> *The people at headquarters [who make the funding decisions] full well knew that this was the prime candidate technology. We would kind of exchange with HQ at the beginning of the fiscal year… they would tell me how much funding was going to applicable technologies so that we could take that into consideration in our funding levels, and obviously we told them what we were funding… everybody understood all around that the funding for the calorimeters would come from multiple sources.* [I57]

This limited soft-power influence over Con-X-specific technology development funding allocation created an unusual management problem for the Project Manager. Despite being responsible for maturing the technologies that would be required to make her mission feasible, PM#2 only had direct influence over a small amount of the actual budget devoted to that end. The true responsibility for R&D management fell to the scientists for all intents and purposes. In the case of the calorimeter, she believes that scientists managed their efforts like managers (unlike some of the other technologies). What this meant for her:

> *With the calorimeter, in this particular instance, they seem to be very focused toward developing the technology toward IXO, well in tuned with what the needs are… I mean frankly, some of the other areas of technology, even when we were funding them, they wouldn't write their APRAs applicable to Con-X. Our TES guys proposed totally in line with what Con-X needed, and then they pooled their resources, and were strategic… A lot of it stems back to how the people propose.*
>
> *That's the route, in NASA, I think HQ knows that, that's part of what APRA is there for… to develop the technology for these upcoming missions that may or may not have the money to support that technology!* [I57]

On the part of the microcalorimeter team, they didn't just seem strategic; in fact, they set out a long range, broad spectrum, plan in 1998 and have, for the most part, followed it since. In addition to the Con-X support, the team received ~$1M per year from APRA as well as variable increments from IRAD ranging from ~50K to ~$400K, with support for 1-5.5 FTEs [D17-24]. In terms of technology strategy, one of the lessons learned from the silicon thermometer microcalorimeter development was the importance of arraying issues. Although solving practical problems wasn't scientifically "sexy," the credibility that the group had previously established allowed them to take the long term view expressed in the following. The new lead scientist (CSE#10) decided that:

> *demonstrating arrays, and reading out arrays, would really show that the technology is ready. With what's happening on IXO right now, I feel a bit vindicated. There was a period of time when we were slogging… we weren't making splashy announcements about breakthrough resolutions because we were solving arraying issues... but it was worth it because we are getting very good resolution (among the best in the world) in arrays… and so having a mind to 'this needs to be demonstrated for a mission'… having that mindset in an early phase even when that missions is really far off, I think that mindset is really important to making technologies ready for missions and I think this has benefited us and that's why the US technology, the NIST multiplexers and the Goddard arrays is the baseline technology for the IXO calorimeter right now. We pushed them to a higher demonstration level.* [I69]

TES Exploration: Inventing to Meet a Specification (the 2000s)

An important byproduct of the TES development being formulated in the context of a next big mission, which was itself, a "next mission" in a string of Grand X-ray Observatories, was the specificity of the early requirements. As recalls CSE#8 (the technologist who had post-doced at NIST):

> *As early as 1998 or 9, they had a pixel design that they were working towards… because the astronomers need [a particular functionality]… here's the [requirement], so we worked for a decade basically with that target in mind… we slowly brought [resolution] down to the Con-X and then IXO spec… and the science guys were bringing the readout [capability] up so that it could interface with this pixel…* [I90]

This specificity was a mixed blessing. On the one hand, it constrained their ability to experiment; on the other hand, it provided an unusual level of stability both in terms of funding and personnel.

> *It was sort of frustrating at times because you wanted to try something really different, but you had to make the same stupid pixel over and over again that you knew wasn't going to work… but the mission forced us into that mold and now we have a very successful design… [because of that] we were able to get funding and maintain the size of team necessary to solve the problems.* [I90]

From a technical point of view, the first half of 2000s decade was a time of getting settled. The relevant equipment had been procured in the late 90s (where the silicon thermisters require a high energy implant, TES needs a very different piece of equipment to deposit thin metal films) but it took until 2004/5 before the process really stabilized. CSA#2 calls this era "*technique limited"* meaning that while the scientists may have a pretty good idea of what needed to be done in theory, practical implementation issues on the fabrication-side constrained what could be done. Of course, the act of solving practical issues (and the corresponding fundamental understanding that was entailed) had an important impact on what to try next.

<u>*Basic Production Issues: Galvanic Corrosion and Leads That Won't Stick (Stabilized ~ 2004)*</u>
During the early stages of detector development, scientists and technologists focus on different aspects of the concept of "demonstrated." As explained by CSE#8, a fabrication guy, "*they just want devices that work and if you make one good one, that's great, they'll just test that one for a long time."* [I90] However, from his technologist perspective, getting one to work means little; developing a repeatable process capable of producing the 1000s of identical detectors has to be the goal:

> *We have these huge wafers and the device is only 100 µm on a side – from a 4 inch wafer you can get lots and lots of devices… and you only need one to work, and they'll test it and publish the results and everything is great, but when you need 3000, production issues are difficult problems, and they don't care if it's solved during the development phase…* [I90]

The above quote probably overstates the difference in perspective. It's not that scientists aren't concerned with repeatability; they absolutely are. The difference is in what they perceive as a fundamental problem. Not being intimately involved in the fabrication, it is impossible to differentiate between what is reported as a "*process mistake rather than a more fundamental problem.*" [e94] The implication of this distinction is partially illustrate below, and revisited with respect to electroplating in the next section.

In this case, the early production challenges largely involved materials processing questions. The TES itself is a bilayer (thin layers of Molybdenum and Gold). When interfacing two thin metal films, a voltage is produced between them and that often causes one of the metals to corrode. Since putting unlike metals together is an intrinsic characteristic of TESs, it couldn't be avoided, so overcoming this challenge was a matter of "*learning the chemistry"* and not exposing the device to non-neutral PH solutions while yielding that particular step. In fact, some of the corrosion "accidents" inspired the vacuum gap absorber solution described below.

The other basic production challenge involved attaching leads to gold (similar to CZT case). Gold is "*one of those blessing and a curse type material."* [I90] It has great electrical properties and doesn't oxidize, making it stable over long periods of time, but not much sticks to it, and most of the materials that do "*are kind of gooey"* and hard to work with [I90]. They eventually found that if the lead was attached to both the Mo layer and the Au layer, the interface was robust. Nonetheless, the mitigating the "curse" side of the gold continues to an area of R&D.

A related challenge to the method of lead attachments was the related question of where to attach them. As array size increased, wiring leads to each individual pixel became impractical: "*there just wasn't enough real estate"* [I94]. So the team began investigating efficient attachment schemes. One promising approach was superconducting through wafer micro-vias. The effort was funded by DDF for two and a half years before Con-X funding cuts caused all non-essential development paths to be put on hold. Seven years later, this strategy is being revisited in the context of microbolometers [I94].

Noise Reduction #1: Strange Geometries and Metal Fingers (Breakthrough ~ 2003)

While the technologists were focused on getting the process under control, the science team was trying to understand the performance of the successful devices that had been produced. In the same way that unexplained noise terms had plagued the original semiconductor microcalorimeters, so too was it an issue for the TES development. Echoing the 1/f catch-22, the team wanted to operate lower in the superconducting transition to improve sensitivity, but found that a mysterious noise term (which got worse lower in the transition) negated any gains. There was considerable disagreement in the TES community – which largely included Goddard, NIST and a Dutch group working on calorimeters for Xeus – about the underlying mechanism [I89].

This community discussion was structured by a series of focused conferences. The group would meet on a bi-yearly basis at the "Low Temperature Detectors (LTD)" conference. In 2001, it was decided that a smaller informal "TES" meeting would also be held in the off years. NIST hosted the first one, and this unexplained noise was a major topic of conversation. The American groups believed that it was a voltage noise, while the European's claimed that their (same) noise was due to thermal fluctuations. It's possible that the conflicting explanations stemmed from differences in their approaches. Where the American's both used Molybdenum-based materials systems, the Dutch used Titanium/Gold, which they operated at a higher resistance. As CSA#10 explains "*people tend to get settled into what first works.*" She believes that molly-based systems are more robust, but acknowledges that there are some operational advantages to higher resistance systems. Regardless, the same fix, improved the underlying noise problem in both devices [I89].

By the time CSE#8 joined Goddard in 2000, everyone knew that square TESs were noisy, but no one knew what would work. "E*veryone was trying different configurations… I'm going to make mine circular; I'm going to make mine a diamond;* brute empiricism [I90]. Just as CSE#8 was leaving, NIST wrote a theory paper, arguing that putting metal on top of the device would suppress the order parameter. The paper suggested a pattern of metal stripes just along the edge. In fact, they also have an all-encompassing patent [D37] "*and probably completely un-enforceable*" [I90] for putting metal stripes on TESs. "*It's interesting how ideas can spread like wildfire;*" [I89] when word got out that NIST was actually getting noise reduction with these metal stripes, everyone started trying it. At Goddard, the IR group started doing it first (CSE#8 was initially working on bolometers), then the X-ray group followed. By LTD 2003, "*we all showed up with stripes on our devices. Even the Dutch, but they said oh, that helps us cool the interior… we spent that conference arguing about why the stripes were working.*" [I89] But the underlying mechanisms didn't really matter, empirically it worked, and some configurations seemed to work better than others.

As CSE#8 remembers it, all CSA#2 saw at the meeting was "*garbage, garbage and more garbage, except for [the Dutch] talk on perpendicular stripes.*" [I90] Their insight was that the lowest noise configuration was when the stripes were placed perpendicular to the current flow; but that insight was incomplete. As presented, full length perpendicular stripes could cause a normal metal series resistance and either make the detector unstable or significantly reduce its sensitivity. So, CSE#7 (who was working on TES bolometers in parallel with the silicon thermisters) drew "*this snaking thing, where the perpendicular stripe didn't go all the way* *across."* [I90] It worked well, and that's the design that the community has settled on. In CSE#8's view:

> *It was [a] case of an active community that were all really focused on the problem [working together] ….groups outside of Goddard proposed solutions and [we built on them] then in our group internal to Goddard, it was science and engineering working together to come up with a solution…when you're fighting this invisible monster kind of thing, you need a lot of ideas on the table…* [I90]

And so the invisible noise monster was slain by multiple empirical swords, in some wavebands anyway.

Noise Reduction #2: Electroplating and Vacuum Gaps Absorbers (Breakthrough ~2005)
The "fingers," or "stripes," took care of the noise term in the IR bands (satisfying the bolometer team), but while it certainly helped with the X-ray microcalorimeters, they were still plagued by inconsistently noisy devices. They were getting wide variations in resolution, and even the best devices weren't good enough (~4 eV compared to the required 2). They believed that the problem was in the material science of the Bismuth absorber. The single biggest breakthrough in the TES development came with the transition to electroplating (as a deposition technique) and separating the semi-metal absorber from the superconducting TES by a "vacuum gap." This insight solved several interrelated challenges [I89].

First, recall that one of the key advantages of the TES is that it's extremely sensitive thermometer allows the use of normal metal absorbers (and their correspondingly high heat capacity, but good thermalization). However, putting a normal metal in direct contact with a superconductor kills its superconductivity. They hoped that by using Bismuth (a semi-metal with very low electron density) they could avoid this problem by putting down an absorber layer that wouldn't affect the TES; but, "*we just found that if we didn't get interaction it was accidental…[for example] if there was some contamination, it provided a nice little barrier layer, but that was accidental and we couldn't count on it [happening]."* [I89]

The solution emerged organically from a series of group discussions, making it is difficult to reconstruct exactly who thought of what, but the discussions were internal to Goddard. Because of the contamination accidents, they started brainstorming ways to limit the contact. At the time they were already using a "mushroom absorber" (big top on a thin stem) to hide the electronics, so they started thinking about how far they could extend the cantilever. This line of reasoning lead to the notion that additional supports could be built outside the TES to support an extreme cantilever. Around the same time, they were also considering spacer materials (barriers like the accidental contamination). In the end, the two concepts merged when they realized that empty space was the best barrier of all; hence the vacuum gap [I89].

Second, Bismuth is a very rough material; rough to the point that the Bismuth layer affected the deposition of the normal metal layer. To fix this problem, CSE#8 suggested that they try depositing the Bismuth using electrolysis instead of evaporation, as they'd been doing. He believed that electrolysis was a better approach in general because (1) evaporating wastes a lot of material (so it would be cheaper) and (2) the material quality should in theory be better too [I89].

However, developing a Bismuth electroplating technique proved quite challenging. Although electroplating is a fairly standard and well established semiconductor fabrication technique, no one does it with Bismuth - "*bismuth is weird*." CSE#8 hired a post doc (who stayed on as CSE#9), with experience in electroplating to "*develop the recipe.*" [I90]

During the post doc, CSE#9 worked with some colleagues in a basic research – thin-film physics – lab at John's Hopkins. He also explored some different configurations of the absorber, as another angle of attack for reducing noise. Many of the ideas were farfetched. For example, he found a way to "*get bismuth to just below the melting point to see if he could re-crystallize it into a big chunk… that and things like that were doomed… so he had a lot of garbage ideas going on and the electroplating thing… and that one turned out really well.*" [I90] It didn't take long from idea to proof that electroplating is actually a much better, much more reliable way of making high quality absorbers. However in hindsight that was only partially true.

Initially electroplating did work extremely well. As CSA#10 remembers it: "*We had a series of great runs with the electroplated absorbers before we started seeing the limit of that technique's reproducibility. For a while there, it seemed like we couldn't make a bad device.*" And, when problems started to show up, the technologists initially attributed them to process error [e94], so the science team discounted their significance appropriately. In the end, it turned out that the challenges weren't human error *per se*, but neither did they affect (and negate) the importance of the electroplating breakthrough. In the words of CSE#8:

> *I think TES has been blessed several times with [things working] the first time and you didn't know why, but you were able to capitalize on the big success and then it would be several years later and several failed devices later before you figured out why that first one worked so well. That was true of electroplating… [CSE#9] made some really nice films during his post doc, but he wasn't exactly monitoring everything he could have…*
>
> *We've lost the recipe several times since then and gone through periods of making bad devices…*
>
> *Electroplating is like the chemical nightmare, there's all these different buffers and things in the solution, there's four or five different components…they're actually helping the bismuth precipitate in uniform ways, [whereas] sometimes it makes clumps and sometimes it makes dust… we eventually learned that you were balancing the chemical environment with the buffers.. .things that a chemist might just know … but you put a bunch of physicists and electrical engineers on it and it takes a while…* [I90]

The above quotes capture two capture important aspects of maturity. Looking back, the technologists emphasize the challenges of scaling up production; which occupied them for several years after the big success. The scientists on the other hand, focus on the proof of feasibility which changed the shape of a development trajectory.

In the late 1990s, NIST had been the only group that had working devices. When Goddard came online in 2000, the leadership was less clear. NIST was focused on lower energies (where the photon can be absorbed by the TES directly) and higher energies (where a huge absorber was necessary). Since Goddard had been focused on Con-X/IXO requirements from the beginning, arraying issues had led them to develop detectors optimized for X-ray astronomy. The vacuum gap breakthrough was motivated by the needs of this particular waveband; its invention served as a signal that 1-10 keV was Goddard territory.

With electroplating and vacuum gap absorbers, Goddard was now making detectors with sub 3 eV energy resolution at 6 keV. Ten years into the development efforts, the Con-X goal was within reach.

Scaling up SQUID Readouts, an ongoing process

One area that NIST continued to dominate was in the SQUID arrays, and later multiplexers. This was partially a strategic decision on the part of NASA HQ, as will be discussed more below. Recall that the original TES paper had identified SQUIDs as the way to read them out. No one really considered using anything else, but different multiplexing strategies, including "time division," "frequency division" and "resonators," were explored and eventually combined [I90].

The original SQUID series arrays were developed in CSE#8's group at NIST during the 90s. Series array SQUID amplifiers can handle much higher slew rates than single SQUIDs because the larger scale of their output (from the series combination of 100 SQUIDs) permits reduction of the strength of the inductive coupling of the input. Multiplexing became important as increasing sensor array sizes required larger numbers of readout channels. The group created a multiplexer by putting a single input SQUID into each TES bias circuit, and then switching those SQUIDs on and off so that only one at a time communicated with a shared series-array SQUID. Around the same time other groups were exploring "frequency division" since they needed to maintain an AC bias, and others still were developing multiplexing via resonators. Each strategy had relative advantages and disadvantages. Current devices have evolved as a time, frequency, resonator hybrid, combining the best of all alternatives. As array architectures have become more ambitious, the readout burden has increased exponentially, and the readouts have evolved incredibly in the last 15 years to keep pace [I90].

Goddard's involvement in this aspect of the evolution has been primarily as a customer. Despite submitting several proposals starting as early as 2000, neither the Goddard TES bolometer nor microcalorimeter teams have ever been funded to develop an in-house SQUID capability. Several rationales for this strategic HQ decision have been offered. The Jet Propulsion Lab (another NASA-center) was already very strong in Niobium tri-layer fabrication, "*so it could be that they [HQ] felt like they'd paid for that [capability] already.*" [I90] Another explanation traces back to the relative priority given to science R&D versus cross-cutting capabilities. Within NASA's selection framework, improving a particular mission performance (e.g., achieving higher energy resolution in an IXO-like detector) is valued higher than improving broadly relevant enabling infrastructure (e.g., breakthroughs in channel reduction are a necessary precursor for the large format arrays planned for next generation IR and X-ray observatories). Fortunately, in the case of SQUIDs, there was a national lab and university community that was more than happy to focus on developing SQUIDs.

It's difficult to assess, even *ex post*, whether the decision not to develop SQUIDs in-house resulted in cost savings. Through the 2000s, HQ via Goddard has paid the NIST group millions of dollars to produce SQUIDs. Of course, had they developed their own capability, "*HQ would be paying us millions of dollars too I suppose… to buy all kinds of dedicated equipment, staff it, maintain it on a yearly basis, all the materials… it would have been a substantial expansion of our group.*" [I90] Ignoring the question of whether Goddard or NIST was in a better position to make progress on the technology for the moment, from a purely "make vs. buy" comparison, it's worth noting that although Goddard chose to "buy" it was from a government lab serving essentially the same tiny market (there are very few customers in need of readouts for thousand element superconducting TES arrays). This co-specialization arrangement with NIST has worked well because of the strong relationship that continues to exist. In turn, the NIST group "*have developed relationships with the Berkley [frequency] and JPL [resonator] groups to make sure that everybody is benefiting from the breakthroughs in technology. There isn't any cut-throat competition issue where you hide what you're doing to make your instrument slightly better…*" [I90]

Theory and Empiricism: Branching Out and Stepping Back (2006-present)

Once an R&D program is folded into a project, understanding the physics of one's devices becomes a luxury; and TES had spent its entire history as part of a project. The project didn't care why the metal comb-like fingers suppressed noise; just that they did. But there is a limit to how much one can proceed without the other. As CSA#10 describes it: "*once we got devices that worked well enough, we definitely focused on doing integration and demonstration… but we never let go of trying to understand some of the more theoretical aspects… we would also theoretically try to engage outside scientists… e.g., talk to folks at the University of Maryland (down the street) and get them to work on our problems.*" [I89] In the mean time, the brute-force experimentation was guided by intuition. For example, the noise suppressing shapes and structures that they tried, were selected based on potential mechanisms "*… well maybe it's due to flux flow noise… and if that's true, then how do we create nucleation sites for the flux flow to only move in certain areas? … so it's never a "this looks pretty, let's try that…" it's always more of a if this is what's going on, then making something like this could affect it.*" CSA#10 still considers this brute empiricism because the trials weren't designed to rule out competing explanations (i.e., "*if it does improve it, it doesn't prove anything since there are multiple affects going on*). If it did work, they just incorporated it and moved on [I89].

It wasn't until years later that they seized an opportunity to re-engage with the theory and finally start to understand some of the fundamentals of the devices they'd been building for almost a decade. Around the 2006/7 timeframe, one of the scientists in the calorimeter group began looking for resources to apply the calorimeter concept to the domain of Solar Physics. The original intention was to advance the emerging magnetic calorimeter technology, and at the same time tap into a new potential patron mission-area. All his proposals were rejected, because the technology was considered too immature. So, he tried again; this time proposing Solar X-ray TESs. This time, they won a 3 year APRA grant 2007-10 which has been renewed in 2010 for another three years [I89].

Solar X-rays require much smaller, much faster TESs compared to Astrophysics. Being forced to play in a very different length scale has "*helped us put together effects that we'd noticed before – that these devices are acting coherently over very long length scales – as you shrink the devices down, you could clearly see that we would get transition temperature of the device to scale with* *the separation of the leads… even with the bilayer being exactly the same."* [I89] While they had intuitively known for many years that all of the superconducting-normal metal interfaces should have some impact, this insight allowed them to identify so-called weak-link superconductivity [I89]. They now had a theoretical explanation for why the lateral proximity effect is as important in describing the behavior, as the vertical proximity of the layers. From a practical point of view, a good theory saves you money. Being able to do the device optimization on paper saves a lot of expensive experimentation (in both time and materials).

In this case, branching out into a second line of research allowed the team to step back while still moving forward:

> *[It] kept our more fundamental investigations going... so we could keep one line of devices relatively static (while we worked on systems issues) while this other line is where we did all the experimentation to understand length scales and other effects and that ended up working out very well for us…* [I89]

Another Chapter in the IXO Saga

2010 marked the release of another Decadal Survey. IXO was not recommended for development as a formal program since several of the enabling technologies were still deemed too immature. It was however recommended that technology development continue, for reassessment in 2020. There is a provision in the Decadal for reevaluation should the European Cosmic Vision process rank IXO more highly. Realistically though, with continued cost increases on the JWST project (which is now estimated at ~$6B, compared to the $700M originally projected) it is unlikely that there will be any Astrophysics money left over for other projects anyway [I93]. Nonetheless, TES development continues. The team won 2010 APRA grants for both the Solar and IXO efforts. After all, once a 2 eV at 6 keV mission is flown, it will revolutionize our understanding of the universe. With technology that enables this level of breakthrough, it's just a matter of time.

Table 3-5: TES microcalorimeter Shock Sequence

Epoch	Start	Component Exploration	Architectural Exploration	Exploitation	Treading Water and Branching Out	End
Start		* Past contact (UL) communicates recent breakthrough * Con-X concept under discussion				
Component Exploration			* Technical breakthrough (Mo/Au) * Con-X ranked 2nd, but directed study to GSFC			
Architectural Exploration						
Exploit					Con-X delays	
Treading Water and Branching Out	* Solar microcalorimeters	* Technical insight about noise reducing fingers and electroplating				
End						

4 The Epoch-Shock Model

How do new capabilities traverse the innovation system as they are matured, and infused into flight projects? Chapter 1 framed a disconnect between the way innovation is currently conceptualized at NASA and the way in which new capabilities are actually conceived, matured and implemented in practice. It further argued that the extant literature cannot explain the complexity, and messiness, of the process, as illustrated by the innovation pathways recounted in Chapter 3. At first pass, the "journey" seems highly idiosyncratic; governed by chance encounters and the persistence of key individuals. However, analysis of the full set of selected cases revealed common patterns among the types of behaviors that were observed. Specifically, the system exhibits epochs of stable, identifiable behavior, punctuated by transition inducing shocks. While the sequence of epochs varies from one project to another, the dynamics they embody (and corresponding feasible management interventions) are surprisingly consistent across the cases.

Figure 4-1 and Figure 4-2 present two views of the Epoch-Shock conceptualization. A complete, successful, journey traverses from *gestation* to *flight*. An unsuccessful journey ends when the new concept "falls" to its "death" in the *technology graveyard*. The epochs are shown as boxes in the track view and stable equilibria (valleys) pocked with crevasses in the dynamic view. Overcoming the potential barrier between epochs requires a *shock*, denoted by arrows in the track view. With a large enough shock, concepts can theoretically transition from any one epoch to another. All crevasses lead to the intermediary treading water epoch; although the journey is still recoverable, unlike the other epochs, treading water is not a stable equilibrium. If no action is taken, the next step is the graveyard.

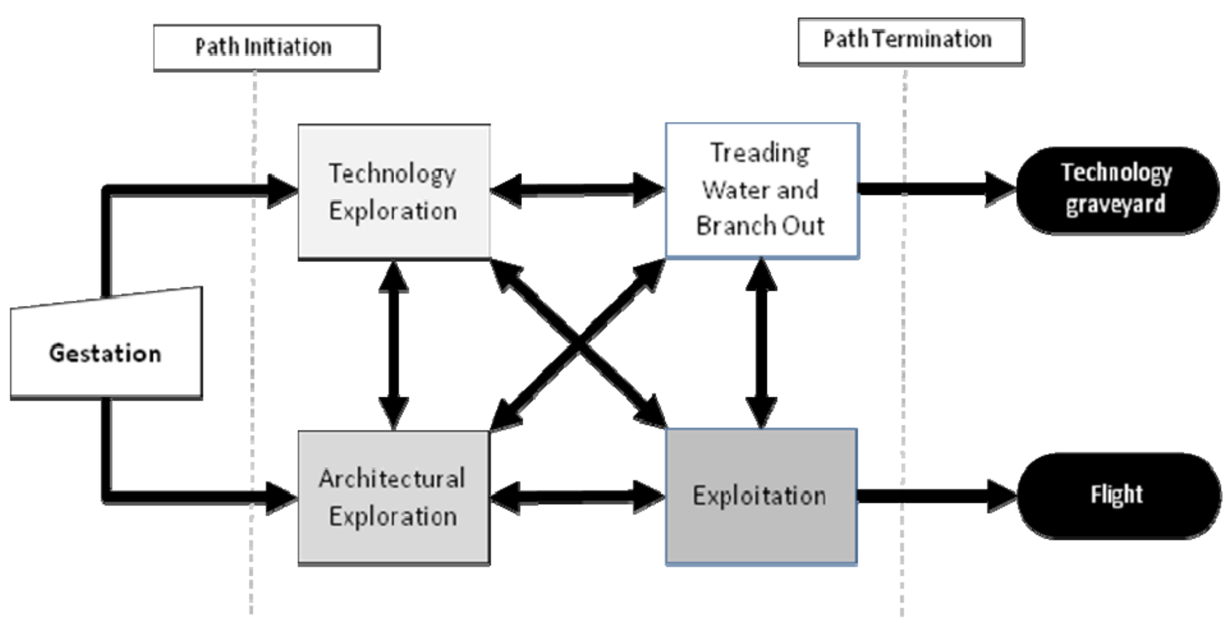

Figure 4-1: Overview of Epoch-Shock Model: Track View

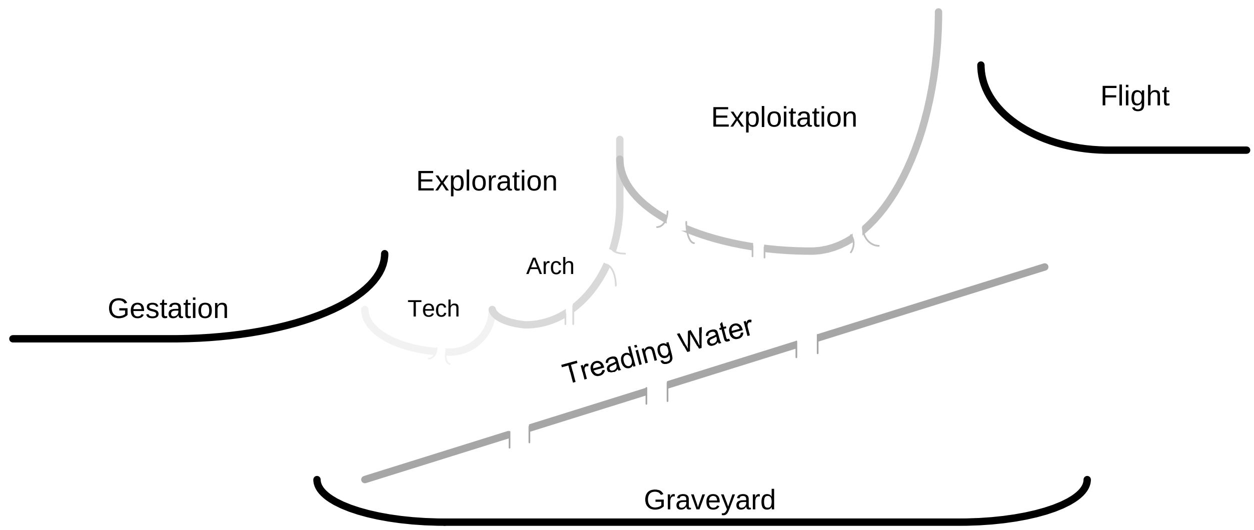

Figure 4-2: Epoch-Shock Model: Dynamic View

The sections that follow substantiate this model in detail. Sections 4.1 and 4.2 describe the characteristic epochs and transition inducing shocks respectively, in terms of (1) the Space Science funding structure, (2) personnel roles and responsibilities, and (3) the technology architecture that define and generate them. Section 4.3 concludes the chapter by integrating the previously defined model pieces.

4.1 Characteristic Epochs

Based on the longitudinal case histories described in Chapter 3 four characteristic epochs and two result states were inductively identified. The epochs are: Technology Exploration, Architectural Exploration, Treading Water and Branching Out and Exploitation. We call the pre-pathway period gestation, and identify two types of pathway terminations: the Technology Graveyard or a first Flight. These epochs are framed in the combined language of so called exploration/exploitation (c.f., (Gupta, Smith, & Shalley, 2006; March, 1991)) and component/architectural (Henderson & Clark, 1990) innovation discussed in Chapter 1. The below text describes the way these epochs can be identified and the behavior they entail. It was observed that the behaviors governing each epoch tended to persist unless a shock of some sort forced a transition. The types of transitions are introduced in context in this section and elaborated upon in the next.

4.1.1 Gestation Period

The gestation period describes the time leading up to the formal initiation of the innovation pathway. Consistent with the observations of Van de Ven et al, there was typically an extended period in which people (who would become key participants) engaged in a variety of activities that set the stage for the eventual development (Van de Ven, et al., 1999). They were typically engaged in other, often tangentially related projects at the time. Although not technically part of the pathway, it is important to consider this pre-pathway period because it sets the initial conditions which can have a strong influence on the path that follows. In this context, relevant initial conditions include the existence of relationships among future team members, experience with the relevant constituent technologies, familiarity with potential applications and access to resources. As Kingdon (2003) argues, it is often possible to trace the origins of a particular technology (or in his case policy) back in time nearly indefinitely. For consistency, we chose to fix the end of the gestation period as the first time the particular new capability was explicitly pursued by name within the innovating organization, in our case NASA. We examined the gestation period at the level of detail, and as far back as was necessary, to understand the initial conditions described above.

For example, in the microcalorimeter case (Chapter 3.4) three seemingly independent threads were brought together to initiate the pathway. First, the IR astronomer, who would eventually propose the X-ray microcalorimeter idea, joined Goddard in 1979, with a grant to build an IR spectrometer. Constrained by resources and the tools of the day, he began working with the detector development branch on bolometer fabrication. Two years into this work, the astromomer received a cold-call from a scientist at NIST – he was thinking of using thermal detectors to detect laser pulses and was wondering if bolometers could be used to detect energy pulses. The astronomer did some calculations and convinced himself that it could, and subsequently filed the calculations for future reference. Second, a colleague in the observational cosmology laboratory had become interested in the noise performance of his system for. It was through that effort that he developed what would become the widely accepted noise theory for bolometer-type devices. The main relevant contribution of this non-equilibrium thermodynamic noise model was that he presented things in such a way that one (and the IR astronomer in particular) could see the scaling laws. This finding drove the IR astronomer to begin investigating extremely low temperature devices.

This effort was already underway when, third, a group of Goddard X-ray astronomers – who happened to sit down the hall from the IR astronomer, and knew him well from frequent lunch breaks in a shared cafeteria – stopped by to get input on a new mission concept for the next grand observatory. They'd heard that he'd had some success with a smaller bandgap semi-conductor for IR spectrometers which, if it worked for x-rays, could improve the X-ray resolution as desired. After they explained the basic principles of x-ray spectrometers to him (he had no previous experience in that wavelength), he thought for a minute and concluded that InSb (Indium Antimonide – the smaller bandgap semiconductor) wasn't appropriate, but that bolometers were worth exploring. To that end, he pulled out the calculations that he'd originally performed for the NIST scientist; and combined with the noise theory, extrapolated a fundamental limit of 1eV at 6 keV if they could get the devices to 0.1 K. That was very good; much better than the x-ray astrophysicists had hoped for. So they got to work; and the microcalorimeter pathway was initiated.

4.1.2 Technology Exploration

Technology Exploration describes a pattern of behavior characterized by the simultaneous pursuit of multiple new technological approaches. In this epoch, there is typically a small core team of experts (be they scientists or technologists) internal to the organization, augmented by (potentially) multiple *ad hoc* collaborations with external counterparts. These external relationships can be quite short lived – they're often initiated through in-person encounters at technical meetings, or by introductions from colleagues and last as long as the technical interests of both parties overlap. It is not uncommon for the core team to nurture multiple of these collaborations simultaneously, limited by the resources available to them and the number of component level innovations being pursued. Resources in this epoch are obtained through some combination of R&D grants from multiple institutional levels (e.g., DDF and CETDP), supplemented by slack resources at the branch-level (e.g., if a technologist is working full-time on a flight project, any down-time and unofficial overtime can be spent on the innovation effort). Funding from different sources is applied for indiscriminately, without differentiating among target maturity levels. During this epoch the overriding goal is to fund the effort for long enough to find some strategy that works and proves the concept in a laboratory environment.

Table 4-1 summarizes the instances of technology exploration epochs observed in the innovation pathways, in terms of the funding mechanisms used, team demographic, and technology. A more detailed version of this table (as well as 4-2:4) are presented in the Appendix III. With respect to funding, multiple sources – at different institutional levels – were in play in each of the instances. Along the personnel dimension, each of the CADR, CZT and Polarimeter instances include the initiation of an external collaboration. In the case of CADR, it was through a summer research collaboration with a nearby university; for CZT it was built on a suggestion from a colleague on a past project; and for Polarimeter it was a contact from a conference. In the microcalorimeter case, although no new collaboration was initiated, a fruitful external collaboration had previously been established. Nonetheless, the team did expand in all but “Si#6”; and the expansion brought in new skill sets, hired as civil servants. Finally, in terms of technology, in each case, technology exploration epochs were characterized by “search” whether structured by multiple parallel developments, or semi-structured trial and error.

Table 4-1: Instances of Technology Exploration

	Case	Funding	Personnel	Technology
Technology Exploration	CADR#1	4xCenter	team + Inst - Tech	parallel component paths
	CZT#2	3xCenter + 3xNASA + Balloon	team +4xTech +Inst	multiple technique strategies
	Pol#3	Brainstorm + 2xCenter + 3xNASA	team + Tech	multiple readout strategies
	Si#4	NASA + Project	team + 3xInst + Tech - 3xObs	multiple materials and techniques tried
	Si#5	2xCenter + 2xNASA + Sounding Rocket + Project	team + Tech	multiple materials and techniques tried
	Si#6	2xCenter + NASA + SR +2xProject	no change	multiple readout strategies and techniques tried
	TES#7	Branch +3xCenter + 2xNASA + SR + Project	team + Tech	Exploration of new materials and techniques

4.1.3 Architectural Exploration

Architectural Exploration describes a focused form of exploration that can serve different purposes at different times in an innovation pathway, be it fleshing out a system concept or searching for a new way to reconfigure existing components to solve a new problem. It is marked by the presence of an articulated performance-oriented objective, and a focus on reconfiguration of existing system modules rather than the development of new ones per se. Scientist (users), and the corresponding project hierarchy that the act of involving users entails, play a bigger role in this epoch than other types of exploration, but the same dynamic of multiple *ad hoc* collaborations with external experts persists. Technically, the emphasis is on demonstrating feasibility of the mission concept, which may involve breadboarding major subsystems or constructing detailed simulations to explore architectural alternatives. If this exercise unearths technology-level show-stoppers a *Technology Exploration* epoch may be initiated. Activities during this epoch are funded in much the same way as during technology exploration – primarily via more substantial directorate-level R&D grants, supplemented by any other resources that can be obtained.

Table 4-2 summarizes the instances of architectural exploration observed in the innovation pathways. The two distinct modes can be observed in the data. The first instance in a particular case looks similar to a component exploration epoch, albeit at a higher level of integration. It's a ramp up of resources (drawing from as many pots as possible), as well as personnel. The identity of the additions varies across the cases, but there are more science (and UL) additions than engineers (as was the case in technology exploration). On the technology side, the activities include major architectural decisions; namely simulations to set key parameters, or preliminary material system decisions. The second mode is more structured; focused on prototyping and providing the first proof of system-level feasibility. Not surprisingly, the funding, while still distributed across multiple institutional levels, has a NASA-level component in all cases.

Table 4-2: Instances of Architectural Exploration

	Case	Funding	Personnel	Technology
Architectural Exploration	CADR#1	Branch + Center	0 + 2xTech	Simulations; begin prototyping
	CADR#2	NASA	team - Inst + Obs	Prototyping
	CZT#3	Branch + 2xNASA	0 + Admin + Obs	Top-level material alternatives
	CZT#4	Center + 4xNASA + Balloon	team + Tech	Explore different patterns and array architectures
	CZT#5	Center + 4xNASA	Team + Tech - Tech	Develop new mission concept based on updated science
	Pol#6	Center + 3xNASA	no change	Prototyping
	Pol#7	3xCenter + Project	team + Inst	Explore different architectural concepts
	Si#8	Branch + Center	team + 3xInst	Prototyping
	Si#9	Project	no change	Explore array architectural alternatives
	TES#10	2xCenter + 2xNASA + SR + 2xProject	team + Tech	Explored alternative fabrication techniques

4.1.4 Treading Water & Branching Out

The Treading Water epoch shares many surface indicators with the *Technology Exploration* epoch, but serves an entirely different purpose. As in the *Technology Exploration* epoch, funding is being applied for at multiple institutional levels, and parallel technologies are being pursued simultaneously. However the strategy motivating these pursuits is fundamentally different. This is a survival mode. The parallel technology paths leverage the same core innovation and apply them to different, but related, contexts. The goal is to increase the likelihood of finding a flight opportunity, or at least further development funding, by branching out to as many application areas as possible. To the extent that new relationships are formed, their purpose is to facilitate this branching out process (i.e., bring onboard potential customers, rather than researchers with complementary technologies). Mostly though, the core team is preoccupied with keeping key team members funded, so they won't be permanently reassigned to other unrelated projects leading to path termination in the technology graveyard.

Table 4-3 summarizes the instances of Treading Water and Branching Out observed in the innovation pathways. Observations from the cases revealed that "Treading Water" and "Branching Out" are two distinct activities that often go together. The first two instances listed below – CADR#1 and Pol#2 – were clear instances of treading water epochs. Although multiple grants were obtained, the magnitude was small and sporadic. In the CADR case, for example, the treading water epoch occurred five years into the development, with only a few key technical problems left to be resolved. Yet, given the Agency-wide R&D climate, only "scraps" of funding from ongoing programs and internal R&D could be obtained. The core team was reduced to the bare minimum and to the extent that progress was made on the core technology, it was incidental to the branching activities. Branching activities involved finding new applications for the existing technology. In the case of CADR, this was laboratory equipment; in the case of Polarimeter, this was solar polarimeters. The reclassification served to reset the maturity of the technology vis-à-vis funding mechanisms. As a result, the CADR could now apply for technology transition funding, while the Solar Polarimeter represented a new research initiative.

Table 4-3: Instances of Treading Water and Branching Out

	Case	Funding	Personnel	Technology
Treading Water and Branching Out	CADR#1	2xCenter + NASA + 2xballoon + 4xproject scraps	no change	new applications for existing technology
	Pol#2	3xCenter + 2xNASA	no change	new challenges for existing technology
	Si#3b	2xCenter + 3xNASA + SR + related Project	no change	leverage technology for sounding rocket program; explore new approaches to same mission area
	TES#4b	5xCenter + 3xNASA + SR + 3xpartial project	no change	new challenges for existing technology

The latter two instances embody branching out without treading water. The branching activities are very similar. The Si microcalorimeter leveraged the technology they had developed to start a sounding rocket program to measure the diffuse X-ray background, and began exploring TES as an alternative calorimeter approach. The TES branching investigated non-X-ray applications (including Solar) as well as a third magnetic thermometer-type. However, from a funding and personnel perspective, Si#3 and TES#4 do not look like epochs of treading water at all. The funding is not at risk, and if anything the team is expanding. As will be explained in more detail in Chapter 4, branching activities during a non-treading water epoch may be associated with technologies developed for Flagship missions.

4.1.5 Exploitation

Exploitation describes a set of structured actions to mature the components of a particular systems architecture towards flight readiness. These actions are governed by formal institutional regulations regarding e.g., the types of testing that must be performed. While problems identified during this epoch can certainly lead to novel technical solutions, the search process is much more focused than in either of the exploration epochs. Where exploration is looking for an approach that will work, exploitation is looking to ensure that the selected approach works efficiently and reliably. The cost of activities in this epoch is proportionally much higher than corresponding laboratory demonstrations; as a result it is often, but not always, conducted as part of a flight project. Further, this epoch tends to entail a major expansion in the size of the innovation team; an increase by a factor of 10 is not uncommon. Within NASA, exploitation in the project context is the domain of engineers and project managers, and most of the new additions will be of this cadre.

Table 4-4: Instances of Exploitation

	Case	Funding	Personnel	Technology
Exploitation	CADR#1	Balloon + Project	no change	Customize prototype for balloon application
	CADR#2	Center + 2xBalloon + Project	no change	Space qualify analogous design
	CZT#3	Center + 3xNASA + 2xProject Bid	no change	Mission-level prototype
	CZT#4	NASA + Project	team - Tech - Admin	Detailed design of commercially sourced array
	Pol#5	3xCenter+ 2xNASA	team + Inst + Obs	Demonstrate robustness of TFT design
	Pol#6	NASA + Project	Team - Inst	Mature TPC design towards flight
	Si#7	2xNASA + Project	team + Inst (project already around)	Improve array uniformity
	Si#8	3xCenter + NASA + SR + 2xProject	no change	Modify existing array for next flight

Table 4-4 summarizes the instances of Exploitation observed in the innovation pathways. To the extent that non-project funding is evident in the funding column, it is either a carryover from previous epochs, or the beginning of a new technology branch. For example, the NASA (APRA) listed under Pol#6 is for the Solar polarimeter pathway initiated during the proceeding treading water epoch. In the personnel column two types of exploitation epochs can be identified: non-project and project. In the case of non-project exploitation (i.e., CADR#1, CZT#3, Pol#5, and TES#9) the team size stayed fairly stable, adding technical members, as in Pol#5, if there was a particular capability gap in the team. For project-based exploitation, the team expands significantly. In either case, the technology activities are similar as defined by NASA systems engineering guidelines (NASA, 2007).

In defining these epochs, care was taken to minimize the association of particular dynamics to project phases. This is because all the above described behaviors can exist, both inside and outside approved projects; they are fundamental to the innovation pathway, not the imposed institutional structure.

4.1.6 Path Termination: Technology Graveyard or Flight

As was alluded to above, an innovation pathway can terminate in one of two ways; the new capability can be implemented on a flight project, or find its final resting place in the technology graveyard. The first termination point is self-explanatory. The second termination point runs contrary to the conventional wisdom that technological concepts never die, or at least have very long shelf lives. As will be explained further in Section 5.3, the data suggests that capabilities can in fact die (for all intents and purposes), if the team and the associated tacit knowledge disband.

4.2 Transition Mechanisms and Shocks

The behavior, characteristic of the epochs described above, persists unless a transition is induced (or shocked) from one epoch to another. This section describes the set of transition mechanisms observed in the cases in general terms and then illustrates how they can combine to produce shocks, using examples from the pathways.

4.2.1 Breakthrough Concepts and Capabilities

A change in what is technically possible clearly has an impact on the state of an innovation pathway. Several different types of technical breakthroughs can occur both inside and outside of the core development team and at the component and architectural level. The two main types of breakthroughs that occur are new concepts "*Insight*" and practical demonstrations "*Demo.*"

A. Insight refers to fundamentally new ideas about the architecture of a system. For example, in the CADR pathway, the technologist realized (had the insight) that cooling power could drastically be increased, and hold time could be effectively eliminated, by employing several staged ADRs instead of the traditional single ADR, one-shot system. Depending on the timing, an insight can initiate a new innovation pathway (typically at the architectural level) or cause a major shift in development direction. Whether this type of breakthrough occurs internal (i) or external (e) to the team influences the level of the initial internal exploration. This is because if the breakthrough is external, key component-level roadblocks may have already been identified.

B. Demos are marked by the first time that the space mission utility of the new capability is demonstrated. Where proofs of theoretical feasibility often exhibit poorer performance than incumbent technologies, practical demonstrations overcome important hurdles (be it demonstrating that the device can be sufficiently broadband, as in the polarimeter case, or attaching microstructures to a new semiconductor compound, as in the CZT case). From a utility perspective Demos are critical, but their role as shocks is weaker than Insights.

This is because they serve to legitimize the current path rather than initiate a new one. Demos can occur at either the component (c) or architectural (a) level, and need to occur inside the team to be relevant.

Table 4-5 summarizes the instances of breakthrough concepts and capabilities that were observed in the cases. It notes the coded type of breakthrough and provides a brief description of the event.

Table 4-5: Instances of Breakthrough Concepts and Capabilities

Case	Type	Description
CADR#1b	Insight(i)	CSE#2 has architectural insight (multiple ADRs in cascade = CADR)
CADR#3	Demo (c)	New gas-gap heat switch invented
CADR#4	Demo (a)	Working 4-stage prototype demonstrated in lab
CZT#1a	Insight (e)	CSA#11 recognizes value of external breakthrough in radiation detector manufacturing
CZT#3	Demo (c)	Patterning and attachment techniques invented
CZT#5a	Insight (e)	GRB Afterglow discovered by Bepo sax
Pol#1c	Insight (e)	Implications of first polarized X-ray spectrum recognized by CS#1
Pol#2b	Demo (a)	First practical (broadband) polarimeter demonstrated
Pol#3a	Insight (i)	TPC idea increases potential sensitivity by orders of magnitude
Si#1c	Insight (i)	CSA#2 has architectural insight (bolometer as calorimeter)
Si#3	Demo (c)	Good enough results (with HgTe absorber) shift focus to architecture
TES#1b	Insight (e)	Former advisor communicates promise of TES breakthrough
TES#2a	Demo (c)	Feasibility of new material system (Mo/Au) demonstrated
TES#4	Insight (i)	Performance improvements associated with noise reducing "fingers" and electroplating demonstrated

4.2.2 Use Opportunities & Needs

Development opportunities are idiosyncratic to the particular organizational context under study. Nonetheless describing the patterns of those present at NASA will shed light on the types of transition mechanisms that are possible. Development opportunities can take the form of identified technical needs or funding opportunities. Ideally, funding opportunities correspond directly with identified needs, but this is not always the case. The key conceptual difference between the two types of opportunities is that, where technical needs are typically identified while trying to implement a predefined concept, a funding opportunity, provides the means through which to define said concept. Both the identification of a technical need and the prospect of a funding opportunity serve to focus the efforts of technologists and scientists; however they create different types of catalysts.

C. Flagship: The science objectives for *Flagship* class missions are developed by a consensus process within the science community, coordinated by the NRC Decadal Surveys and implemented by NASA. One implication of this extended process is that the corresponding instrument community is given approximately two years advanced warning about what will be requested next. Given the revolutionary aspirations of Flagship class missions, this type of opportunity typically prompts a search for a radically new approach in particular technology areas, funded by several additional years of dedicated technology development prior to a formal instrument selection. Thus, the prospect of a Flagship class call can “pull” significant advancements in fairly specific areas. The announcement of the next Flagship mission happens fairly predictably in the year following a Decadal Survey.

D. Explorer: Explorer class mission announcements, on the other hand, are explicitly non-specific (in terms of target science), come with little advanced warning and expect a high level of technology maturity in the proposed baseline. Where Flagship missions incorporate multiple advanced instruments, explorers are typically organized around one or few targeted measurements, as proposed by a PI (Principal Investigator). As a result, they provide great opportunities for first flights of new technology; however, that technology must have been previously developed outside of the mission context for it to be selected. Thus the timing with respect to other types of shocks (particularly breakthroughs) modifies the impact of the Explorer mechanism significantly. The spacing of Explorer-calls is somewhat unpredictable, and depends on budgetary constraints from other program elements (e.g., if a Flagship is delayed or overrun, the next explorer call will be delayed too).

E. Gap: Mission or subsystem concepts sometimes evolve outside the context of any particular Flagship or Explorer opportunity. As the details are worked out, technical roadblocks are sometimes identified. Similar to the Insight transition described above, a Gap can initiate a new innovation pathway or significantly change an existing trajectory. The key conceptual difference here is the direction of the impetus, which often relates to the background of the initiator (e.g., a scientist is more likely to start from a system concept, identifying required technology improvements as the concept is fleshed out, whereas a technologist may identify opportunities for improvement in the context of the technology s/he is working on).

Table 4-6 summarizes the instances of use opportunities and needs that were observed in the cases.

Table 4-6: Instances of Use Opportunities and Needs

Case	Type	Description
CADR#1a	Flagship	Con-X baseline specification comes out in 1998 (requires much more cooling/mass than pervious missions => need for better cooling technology)
CADR#2	Gap (c)	Initial prototype reveals heat switch roadblock
CZT#2	Gap (c)	Selected detector architecture requires technical breakthrough in terms of packaging and yield
CZT#4	Explorer	1995 and 1996 Explorer AO: BASIS not selected… twice, but proposal writing exercise demonstrates new technical challenges
CZT#5b	Gap (a)	New science need motivates new technical requirements
CZT#6	Explorer	1999 Explorer AO (SWIFT selected) => flight development
Pol#1a	Gap (c)	Costa demonstration reveals need to make broadband
Pol#2c	Explorer	2003 Explorer AO (AXP not selected), but brings group together and expands team to include project scientist (management role)
Pol#4	Demo (a)	Practical demonstration of TPC
Pol#6	Explorer	2008 Explorer AO (GEMS selected) => flight development
Si#1a	Flagship	AXAF ranked 1st in (80s) Decadal Survey
Si#2	Gap (c)	Concept demo reveals technical challenges (absorber material etc…)
Si#4	Flagship	Mission timeline forces emphasis on array optimization
Si#8	Explorer	SRX selected as MoO for Astro E2
Si#11	Gap (c)	NeXT spec drives new component requirements
Si#12c	Explorer	SRS selected as MoO for Astro H
TES#1a	Flagship	1996-98 pre-discussions formulating Con-X specification
TES#2b	Flag study	Con-X ranked 2nd (in decadal) but study contract directed to GSFC

4.2.3 Changes in Context

Innovation pathways don't proceed in a political vacuum. The prioritization of particular science questions (e.g., when Flagship class missions are ranked) drastically improves the chances of success for some developing technologies and limits the prospect for others. Similarly international collaborations force compromises with respect to who develops what – while agency competencies are certainly taken into account, sometimes the desire for friendly relationships can shelve a decade's worth of development. This category of shock is idiosyncratic and highly unpredictable (from the perspective of the innovation pathway) but can have a substantial impact on it. For the most part, changes in context correspond directly to one of the previously described shocks (e.g., Flagship); however three types of context changes merits individual consideration.

F. Drought: Non-project technology development funding opportunities arise on a yearly basis at both the center- and directorate-level, targeted at new concepts of various levels of maturity. Technologists rely heavily on these grants to sustain their research; however as shocks, the availability of such grants only act negatively on the technology trajectory. Namely, receiving funding has a neutral affect on the shape of the trajectory, but not securing funding leads to an epoch of treading water (whatever the current epoch). It's worth clarifying that drought in this context refers to an unexpected loss of funding. For example, in the CADR case, the technology was progressing well and would likely have received continued funding except that a top-level policy change eliminated all potential funding sources simultaneously.

G. Priority is in a sense synonymous with mission context, in that the fate of a technology that has been earmarked for a particular Flagship mission concept will live and die with that mission. For example, the TES microcalorimeter pathway was initiated for the Con-X flagship mission and consequently has been subjected to the fluctuations in priority associated with it. This has manifested both in terms of variable ease of access to R&D funding and the ability to recruit and retain key personnel. Priority can be positive, as for example, in the CZT case, when the then NASA administrator publicly noted that a GRB finding mission would be an excellent fit for the next Explorer call, maturing the CZT GRB-finding concept became a Goddard priority. Priority can also be negative, as for example, in the TES case, when Con-X was delayed indefinitely, mentioning Con-X as a potential near term application stopped being a near guarantee of funding. Clearly, the same Priority change can have different impacts on different technologies. What was a negative impact on the TES case worked in the favor of the Si microcalorimeter; NeXT became a viable near term opportunity to leverage past work on the incumbent (Si) technology.

H. Failure: To this point, the discussion has implicitly assumed that once a mission has been approved (i.e., has entered Phase B) the pathway has terminated with success. While the technology has achieved the standard of "implemented," until the instrument returns data, NASA doesn't considered it to have been demonstrated. For example, in the eyes of many, microcalorimeters have yet to be proven because, although they have been launched twice, they have never been operated on-orbit (due to unrelated mission failures). Each of these failures resulted in a next opportunity for exploration (in preparation for the redo at the mission level). Instances of failure in the relevant mission area will be tagged as Failure. Failure can be positive or negative, depending on one's perspective. For example, in the CADR case, the sequence of Astro mission failures gave priority to redundancy, consequently prioritizing one of the CADRs main advantages over the incumbent.

Table 4-7 summarizes the instances of changes in context that were observed in the cases.

Table 4-7: Instance of Changes in Context

Case	Type	Description
CADR#5a	Drought	CETDP program eliminated as part as post-VSE restructuring
CADR#5b	Priority(-)	Con-X's schedule (the target flagship) was pushed back nearly a decade
CADR#6a	Failure	The Astro E2 cooling system failed on orbit (negating the potential for SRX data return)
CADR#6b	Priority(+)	Astro H recast as MoO to replace Astro E2, with much more priority place on redundancy in (failed) cooling system => Multi stage ADR proposed as solution. Also, ASP mission (which requires continuity explicitly) enters study phase
CZT#6a	Priority(+)	NASA Administrator states that next Explorer should be a GRB mission during public address
Pol#5	Drought	Limited R&D funding available due to post-VSE restructuring; ran out of resources waiting for next Explorer AO
Si#5	Priority (-)	AXAF-S (target mission) split, then demanifestation creates uncertainty on "S" portion, revived as international collaboration (Astro E)
Si#7a	Failure	Astro E fails on orbit
Si#7b	Priority (+)	Replacing Astro E with Astro E2 becomes priority
Si#10b	Priority (-)	Con-X's schedule (the target flagship for TES technology) was pushed back nearly a decade, creates opportunity to propose Si for NeXT
Si#12a	Failure	The Astro E2 cooling system failed on orbit (negating the potential for SRX data return)
Si#12b	Priority (-)	Instead of working towards NeXT (an ambitious X-ray mission) emphasis is now on replacing Astro E2 with Astro H (A version of NeXT that used the engineering model detector plane)
TES#3	Priority (-)	Con-X's schedule (the target flagship) was pushed back nearly a decade

4.2.4 Actions of Individuals/Teams

As seen in the QWIPs story, the actions of individuals can have a huge impact on the shape, and persistence, of an innovation pathway. For example, when QWIPs was rejected for a phase II SBIR grant, the pathway may have terminated there (despite major technical progress), had the lead technologist not called his contacts and advocated for the technology. The presence of "Agency" in our cases seems to be most observable during times of strife. Specifically individuals clearly matter during treading water epochs, but their impact is less clear when progress is being made and funding is readily available. Nonetheless, technical breakthroughs were often closely preceded by new collaborations, and the fresh ideas and equipment that the collaborators brought to the team. Thus, while individual actions don't typically shock a transition on their own, they contribute importantly to other transition mechanisms, often putting them over the top.

I. Advocacy: The dictionary defines advocacy as the active support of an idea or cause. Within the space science innovation system "active support" can take several forms. A technologist or scientist may fight a funding decision, or use backchannels to influence it (e.g., the QWIPs instance described above). A manager or director, can use his or her positional power to influence which emerging concepts are supported by the "A team" (e.g., in the CZT case, when the Space Science Director called an engineering branch head to make his personal priorities clear). Or, outside of NASA, the science community can lobby in order to clearly communicate their priorities to decision makers (e.g., in the QWIPs case, congressional lobbying lead to a TIRS line-item in the appropriations bill).

J. Join: In the context of the relatively small core teams, each member (particularly when outside researchers are recruited) tends to bring a particular capability; be it expertise in a component of the innovation, or access to a critical facility. In the pathways studied herein, additions had much more impact than departures (as most departures not caused by funding droughts coincided with a loss of interest by one or both parties). New collaborations were most commonly initiated following the annual or bi-annual technical conference (of the relevant domain), but some meetings happened off cycle as well. Additions are coded as join (+) and departures are coded as join (-).

Table 4-8 summarizes the instances of changes in context that were observed in the cases.

Table 4-8: Instances of Actions of Individuals/Teams

Case	Type	Description
CADR#6c	Advocacy	CSE#2 presents multi-stage ADR as only solution to redundancy requirement
CZT#1b	Join(+)	Science director tells BH#1 and CSA#11 to work together; initiates collaboration between science and engineering
Pol#2a	Join(+)	CS#1 meets IR#1 at conference, make arrangement to use IR#1's equipment for experiments
Pol#3a	Join(+)	CS#1 learns of UL#4's research; gets invited to lab for demo; prompts new approach, and collaboration
Si#1b	Join(+)	X-ray Astrophysics group walk down the hall to get IR Astronomers perspective on design problem
TES#1c	Join(+)	Former adviser suggests new approach to CSA#9

4.2.5 *Observed Transitions*

Some of the above described transition mechanisms can induce certain transitions from one epoch to another on their own. Other transitions require the confluence of more than one mechanism. Table 4-9 summarizes the transitions that were observed in the innovation pathways. The contents of the cells describe the shocks or set of shocks that induced a transition from the row epoch to the column epoch (i.e., "1. Insight(e) + Gap(c)" initiated a pathway, transitioning from gestation to component exploration). The number designates the sequence of that particular transition in the innovation pathway. The colors correspond to different cases. Blue is CADR; Red is CZT; Green is Polarimeter; Purple is Silicon Thermometers; and Black is TES. The listed mechanisms correspond to the nomenclature described above. In theory, a shock could induce a transition from any current epoch to any next epoch. However, as is evident in the NxN matrix captured in Table 4-9, some transitions are more likely than others. Chapter 5 explores the patterns of transitions in more detail.

Table 4-9: Summary of Observed Transition Inducing Shocks

Epoch	Start	Component Exploration	Architectural Exploration	Exploitation	Treading Water	End
Start		1. Insight (e) + Gap (c) 1. Flagship + Insight (e) + join	1. Flagship + Insight (i) 1. Insight (e) + join 1. Flagship + join + insight (i)			
Component Exploration			3. Demo (c) 3. Demo (c) 3. Demo (c) 2. Demo (c) + Flag study	8. Explorer (MoO) 12. Failure + Priority + Explorer (MoO)		
Architectural Exploration		2. Gap (c) 2. Gap (c) 2. Gap (c) 11. Gap (c)		4. Demo (a) 4. Explorer 6. Priority + Explorer 2. Join + Demo (a) + Explorer 4. Demo (a) 4. Flagship	3. Priority	
Exploitation			5. Insight (e) + Gap (a) 3. Join + Insight (i)		5. Drought + Priority 5. Drought 5. Priority	7. Approved 7. Approved 9. Approved 13. Approved
Treading Water	6a. Branch 6a. Branch 5a. Branch 3a. Branch	4. Insight (i)		6. Failure + Priority + Advocacy 6. Explorer		6. Approved
End		7. Failure + Priority	10. Priority			

4.3 Putting the pieces together: Innovation as an Expedition

The preceding sections have articulated the pieces of the process model. Chapter 2 described the underlying systems architecture, in terms of funding structure, personnel roles and responsibilities, and technical architecture. Section 4.1 defined a set of characteristic epochs in terms of the dynamics in those underlying tracks. The epochs embody stable behaviors which persist unless a transition is induced. The set of transition inducing shocks were enumerated in Section 4.2, which categorized mechanisms as technical breakthroughs, use opportunities and needs, changes in context and agency. It assessed the extent to which individual mechanisms can induce transitions independently versus requiring the confluence of multiple mechanisms. This Section puts the pieces together.

The act of categorizing epochs and shocks tends to overemphasize structural characteristics of the process and shocks that act on the technology. However, as illustrated by the innovation pathways documented in this thesis, the journey is very much a human undertaking, albeit subject to serious environmental constraints. In order to clarify this aspect of the model, we present a synthesized version of the CADR innovation pathway (details in Section 3.1), explaining it in the language of the model, with the aid of a mountaineering expedition metaphor.

Gestation Period

People don't just climb mountains like Everest on a whim. It takes months of training that typically includes summiting other technically difficult mountains. They put together a team of fellow climbers with complementary skills (e.g., a team doctor is a useful asset) that had been previously developed over many years. Sherpas, are also a must, to help carry extra supplies and equipment that will be needed over the course of the expedition. Finally, while many recreational hiker/climbers dream of someday attempting Everest, most never try. With the exception of the few extreme climbers who are driven to climb it "because it's there," for most teams it takes a specific catalyst (e.g., Erik Weihenmayer was motivated to prove that blindness didn't limit his potential).

Initiating an innovation pathway requires analogous elements. In the case of the continuous adiabatic demagnetization refrigerator (CADR), the 1998 inception was preceded by a two decade gestation period. In 1980, the Goddard cryogenics branch was formed in order to develop a core sub-Kelvin cooling competency. Something that the new branch head believed would be important in the future. They chose the ADR as the key technology, developing the first one for AXAF (one of NASA's grand observatories, later renamed Chandra). Over the next two decades, the group developed several increasingly functional ADRs. When CSE#2 joined the Goddard cryogenics group in the mid 1990s, he was trained on the then standard (albeit, highly customized), ADR flight developments. He observed some clear limitations in the way things were done, that would become particularly important as cooling requirements increased with future large arrays. And in 1998 he was struck with an architectural insight that would overcome these challenges. He shared his idea with his office mate, also in the cryogenics branch, but with a complementary perspective. The colleague agreed that this was something worth pursuing, and together they sound funding to begin the journey.

Shock: The transition from gestation to architectural exploration was induced by an insight, bolstered by an identified technical showstopper for a future Flagship mission.

Exploration (in general)

When climbing Everest, the first major milestone is arriving at Base Camp. There is very little assumed risk at this point. While many expeditions make the trek by foot, it is accessible by truck, so if any equipment is forgotten, resupply is fairly straightforward. The first leg is thus a good opportunity to test out the capabilities of the team, discover if new additions are required etc… as the team acclimatizes to the altitude. Most teams acclimatize in stages, starting with the Sherpa capital Namche Bazaar, stopping for a day, followed by Dingboche for another day, and finally Everest Base Camp. Many more people climb to Everest Base Camp, than actually attempt the full journey.

Architectural Exploration

We find it useful to differentiate between the (technical) level of the exploration (i.e., architectural and component). This could be considered analogous to working in equipment (architectural), versus base conditioning (component). If the conditioning isn't there, great equipment won't help that much, but sometimes you need to try to work in the equipment to realize that there's a more fundamental challenge.

In the case of CADR, since the architectural insight was, in theory, a straightforward reconfiguration of existing technology, they jumped straight into an epoch of architectural exploration. Although the first tranche of resources was much less than would eventually be necessary, they were confident that more could be obtained along the way. The initial activities involved developing computer simulations to define a useful staging scheme, and building a first prototype. In building the prototype, they realized that they would have to overcome a component-level show stopper after all. In the incumbent ADRs, heat switches were used to maintain a stable cold state, and opened during recycling. In the CADR design concept, they would have to operate at multiple (staged) temperatures, with continuous, controlled partial cycling. Not unlike a fit climber, without previous altitude experience, learning that a new kind of conditioning would be required, the team dropped their level of exploration to focus on this key enabling component.

Shock: The transition from architectural exploration to component exploration was induced by the identification of a gap – the need to invent a two-way heat switch.

Component Exploration

The team was successful at obtaining more resources. This enabled them to pursue three parallel approaches to inventing a suitable heat switch. They initiated a collaboration with an external university partnership for this purpose; assigning a different graduate student to each approach. One of the approaches proved successful. While focusing at the component-level, they also recognized that a supporting piece of equipment would be extremely valuable moving forward. They purchased an Electric Discharge Machine (EDM) and recruited an expert operator (technician).

Shock: The transition back to architectural exploration was induced by the demonstration of the new heat switch technology.

Architectural Exploration

Component hurdle eliminated, they returned to the prototyping activities from before. Now they were able to demonstrate a continuous ADR in a laboratory environment. Armed with a four stage prototype, they seemed poised to receive the additional funding that would be required to make the next push on the path to flight.

Shock: The transition from architectural exploration to exploitation was induced by the demonstration of a four-stage CADR prototype and the prospect of future funding.

Exploitation

After Everest Base Camp, the next major milestone is called Advanced Base Camp (ABC). It's at a substantially higher altitude which has several implications for its stability as a base. Resupply is much harder and the higher altitude begins to affect you. There is an inherent tradeoff between acclimatizing and running out of resources. Going forward these challenges only becomes more intense. The climbing becomes significantly more technical. Falling, during a fixed rope climb, or down a crevasse become a real concern. Further, the atmosphere thins as climbers approach the “death zone.” The “death zone” is so named because climbers can typically only endure a maximum of two or three days at this altitude for making summit bids. It’s only 1000 meters up from Camp IV to the summit, but even with the most favorable weather the last leg is harrowing. If a good weather window doesn’t open within those two days, climbers are forced to descend, often all the way to Base Camp.

In the case of CADR, the momentum from the Demo(a) was enough to get them into an exploitation epoch, but their position wasn’t stable or sustainable. The R&D funding that they expected to sustain them as they prepared for a summit push never came. Neither did a relevant mission opportunity which could have provided them with the extra momentum. The expected R&D funding never came because of an agency wide policy decision to focus on President Bush’s Vision for Exploration (and translated into cutting the agencies R&D budget by nearly three quarters). At the same time, the Flagship application for the new technology was running into problems of its own. Ranked second in the decadal survey, and funded only as a directed mission study, it was further scaled back in the middle of the 2000s.

Shock: The transition from exploitation to treading water was induced by the combination of a funding drought and a shift in future mission priorities.

Treading Water

In the climbing scenario, injuries and falls are an ever present risk. The landscape gets more treacherous the higher you go. While the goal is obviously to stay out of emergency situations, if one does occur, more experienced teams, with access to better resources are more likely to survive.

In the CADR case, the sudden lack of funding shocked the pathway from a state of impressive progress, to survival mode. While the cryogenics branch head had previously been very supportive of CSE#2s initiatives in the past (e.g., encouraging him, during the early stages, to spend down time working on his new idea), as resources tightened, she suggested the effort be mothballed until the funding climate improved. CSE#2 resisted because he was concerned that a temporary mothball would quickly become permanent. In particular, he was concerned about losing one key team member – the technician with expertise in EDM use, who was near retirement, and would likely choose that over being transferred to a new project for the interim. Training a replacement, and bring back other team members once funding was restored would cripple progress.

So CSE#2 scrounged for resources, and tangentially related project work. He secured just enough funding to keep the core team alive, while peripheral members found other expeditions to join. While the forward progress seemed minimal, several of the sustaining activities proved relevant later. During this period, they developed a CADR for a balloon flight, which mitigated certain operational concerns. They also developed a laboratory CADR, which involved mating the system to a mechanical refrigerator, rather than the standard helium bath. It also involved experimenting with different temperature stability control schemes. These side projects would become critical precursors, when rescue did eventually come.

Shock: The rescue – if one can call it that – came in the form of a series of related mission failures. The Japanese Astro program used standard ADRs to cool its relatively modest detector arrays (recall that larger arrays require greater cooling power). The first X-ray mission, Astro E failed on launch. Five years later, the second mission (Astro E2) failed on orbit – due to a failure in the cooling system. Japan and NASA decided to collaborate on a third attempt (Astro H), but his time failure could not be tolerated. This put a huge premium on redundancy in the cooling system, which translated into a desire for an ADR system that could mate to both an He bath and a mechanical refrigerator (just what had been demonstrated during the branching out). Around the same time, the scientist for whom the team had developed a ballon-based CADR, began formulating a scanning mission that would explicitly require the continuous capability. This further solidified the sortie from the crevasse.

Exploitation

Although the team was back on the trail again, the team members that had left during the treading water epoch could not return. As CSE#2 had feared, their interim projects had become permanent responsibilities, which could not be shirked. The system that was baselined for Astro H wasn't a CADR, but it was a multi-stage ADR with a complex control scheme. These requirements allowed the team to mature and qualify a nearly equivalent system. The main remaining aspect was to prove-out the thermal stability during continuous operations. Now that infusion to a flight project appeared imminent, the team felt justified in requesting a strategic investment from Goddard. Using IRAD funds, they were able to mature the remaining thermal stability capability.

As they approach the "death zone" once again, success is out of the hands of the team. They are as ready as they can be, but if the weather doesn't clear up (i.e., neither IXO – the Flagship – nor ASP – the scanning mission – get approved) they may be forced to return to Base Camp, potentially for the foreseeable future.

This model can similarly be used to explain the pathways taken by each of the innovations described in Chapter 3. Figure 4-3 represents this dynamic view of the innovation pathways. In the figure, the colors differentiate among the pathways as indicated in the legend. The sizes of the bubbles are scaled to the time spent in each epoch. The longest is 5 years and the shortest is one year.

While the particulars are unique to each case, there is notable structure across the pathways. Once initiated the team moves from gestation to either component or architectural exploration – depending on how well defined the concept is and the level of the insight. If the first push is to architectural exploration, a drop to component is often required before taking on the higher stakes of exploitation. If component exploration is first, architectural exploration is a natural next step, but may not even be necessary, if architectural questions have already been worked out. The transition from exploration to exploitation requires a bigger commitment. While, a system prototype may seem like a sufficient shock to make a summit push, it's a big risk, if a mission opportunity doesn't come along in time, much of the forward progress may be lost. Similarly, changes in mission priority during this critical phase can easily lead to an epoch of treading water. Treading water is highly survivable, if it's temporary and advocacy is strong, but if it persists, the team can quickly fall apart. The uncertainty associated with this epoch often spawns new, related, pathways. Assuming the team stays together, the summit can be restored at multiple points. We observed second starts in all of the component exploration, exploitation and flight epochs. For them there, the path is similar to any journey at that stage. These observations will be further elaborated upon in the sections that follow.

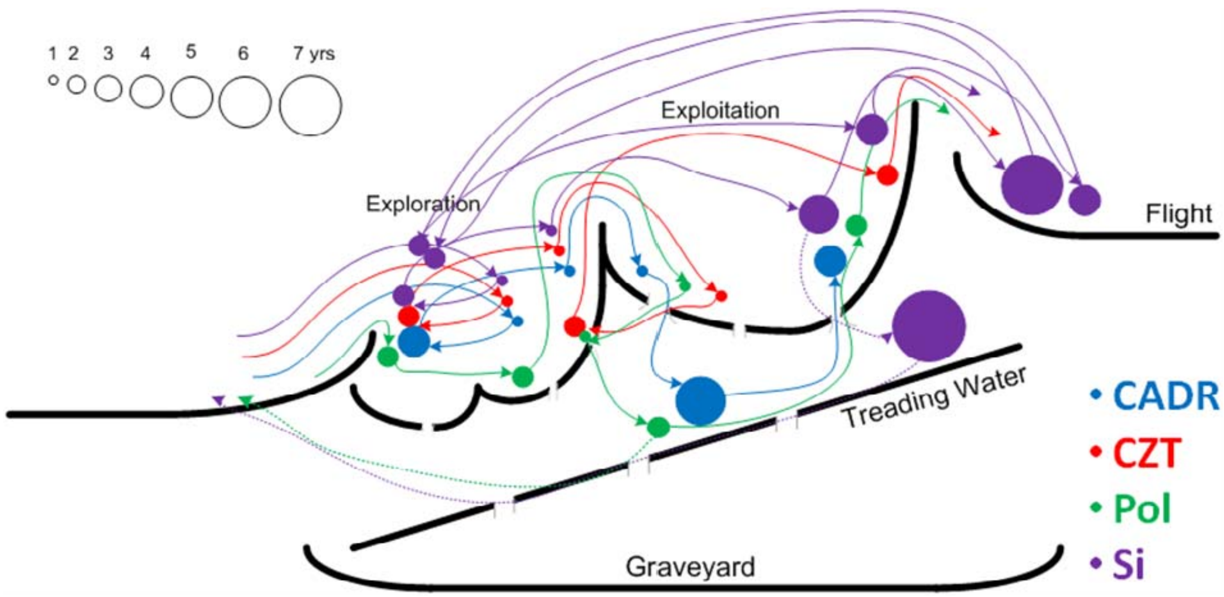

Figure 4-3: Overlay of Innovation "Expeditions"

5 Patterns and Implications

To this point, the emphasis of the discussion has been descriptive; identifying characteristic epochs and transition inducing shocks that govern the dynamics of innovation in NASA's SMD, and potentially in other technology intensive bureaucratic organizations. In describing the model, we argued that this epoch-shock formulation is not just stage-gates with new names. The two conceptualizations are meaningfully different, both in terms of the underlying model, and the policy interventions they lead to. Specifically, where improving the yield of a stage-gate innovation system is a matter of solving a coupled resource allocation problem and setting appropriate gate criteria; the epoch-shock view forces managers and innovators to recognize the implication of being in the different epochs at different times, and harnessing the momentum provided by the variably predictable directional shocks.

To that end, this chapter begins by illustrating how the Epoch-Shock model can be used to understand the dynamics that underlie innovation in this context. It then uses this understanding to compare the expected outcome of policy interventions currently being considered, under the Stage-Gate and Epoch-Shock formulations. Finally, it explores how the Epoch-Shock model provides new insight into the policy problem; and to theory and practice more generally.

5.1 Common Pathways: The Roads Most Travelled

The Epoch-Shock formulation provides a structured basis for comparing and understanding the pathways studied herein. Figure 5-1 captures the observed sequences; the bracketed numbers show the number of instances of the particular epoch and the thickness of the arrows are scaled by the number of times that the particular transition was taken. This representation gives a top level sense of the relative importance of particular pathways. Comparing Figure 5-1 to Figure 4-1, it is apparent that while the majority of the potential pathways are used at least once, not all bidirectional flows were observed, and some unexpected transitions occurred.

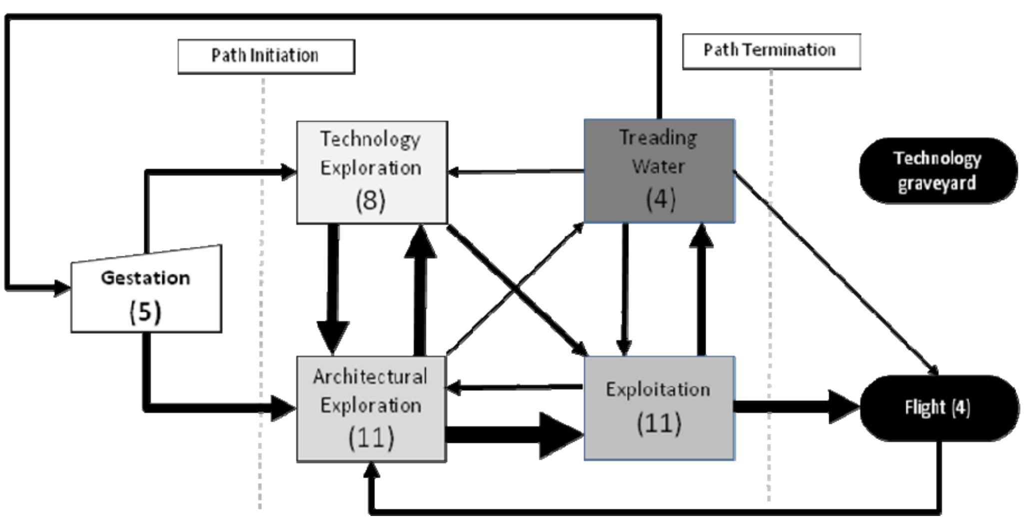

Figure 5-1: Overlay of Paths Travelled

Since not all shocks act independently, nor do they have the same impact on the process, it is also useful to unpack Figure 5-1 one level further, in terms of the four classes of shocks. This view is shown in Figure 5-2. As indicated in the legend, the colors correspond to the type of shock (e.g., Red represents a technology shock). Primary colors were selected so that combined shocks could be illustrated as mixtures (e.g., Red + Blue = Purple).

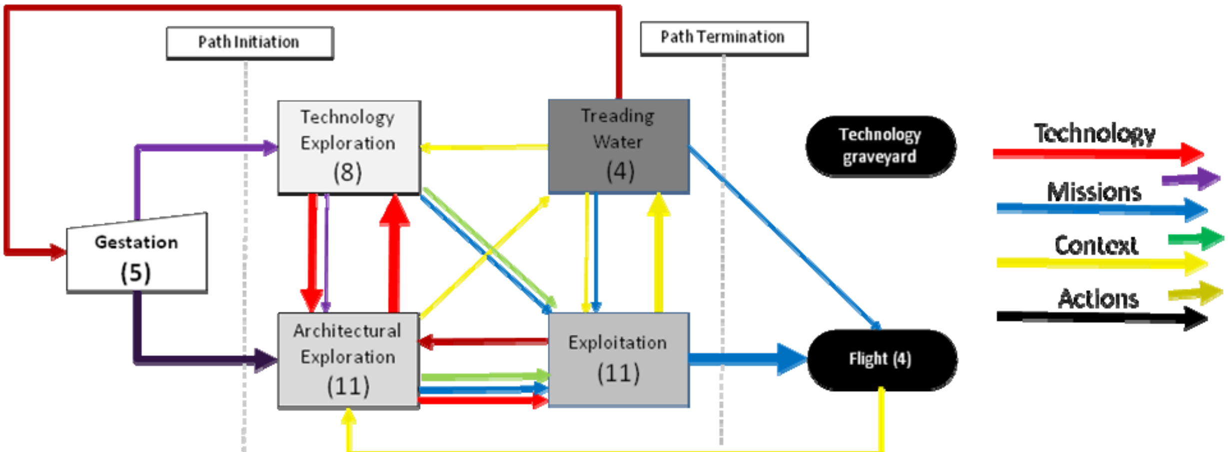

Figure 5-2: Differentiated Track Overlay View

Recognizing that the five cases, used to generate this figure, do not in any way capture all of the possible transitions, the map is used as a guide to identify potentially interesting behaviors worth unpacking. The sections that follow focus on the dynamics that underlie two "most travelled paths."

5.1.1 The Breakthrough Window Lag

The thickest arrow in Figure 5-1 maps the flow from architectural exploration to exploitation. While passing through exploitation, much of the flow is diverted to treading water rather than ending up in flight. This observation is of potential concern because by the time an innovation has reach exploitation, feasibility has been demonstrated and a significant investment has already been made. Since it is desirable for useful capabilities to be used, it is important to understand what is causing the forking. Upon resolving the different shock catalysts (e.g., Figure 5-3) it is apparent that this fork is related to the type of shock that caused the initial transition. Specifically, if a technology shock (red) initiates the transition to exploitation, and there is delay before the next mission opportunity, or there is any negative context shock, then treading water occurs. If, on the other hand, a mission shock initiates the initial transition to exploitation (blue or green), successful termination in flight is likely, barring a major negative context change.

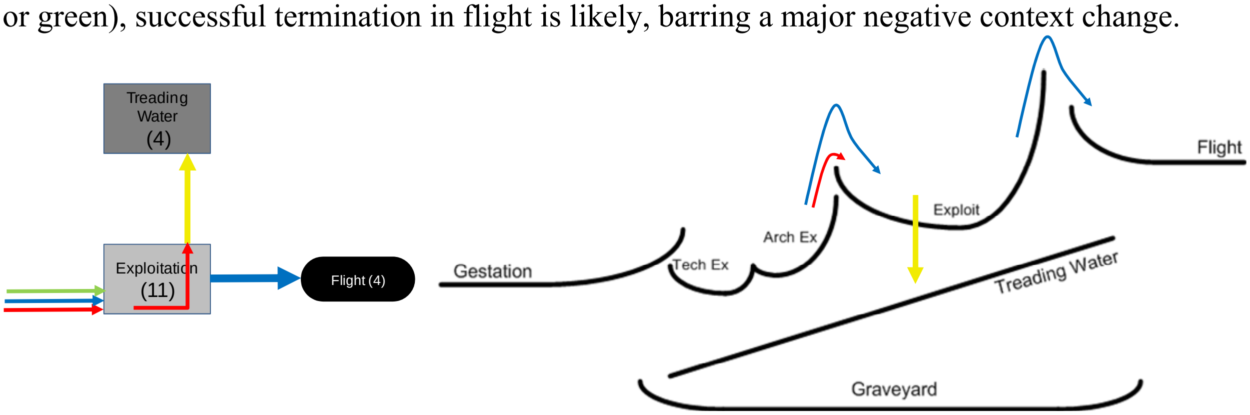

Figure 5-3: Breakthrough Window Lag

The diversion to treading water occurs because exploitation is a time-limited state. This is because the activities that define the exploitation epoch are inherently expensive (e.g. vibration testing a prototype is much more expensive than a bench experiment.) Also, more mature capabilities qualify for far fewer funding mechanisms. As a result, if the team is rejected from one source, it is harder to recover. Combined, this means that innovation teams can't loiter in exploitation. Instead, innovation teams must survive delays by treading water and branching out. This explanation raises an important tradeoff: the treading water epoch builds slack into the system so that innovations can survive until the next flight window opens. Some slack is necessary, but too much slack wastes resources by enabling prolonged treading water.

Recall that the treading water epoch is characterized by diversification both in terms of the technical level of the activities and the scope of their application. In terms of the technical activities, technologists exploit the hierarchy and complexity of the system to strategically re-scope their efforts. Specifically, since components, subsystems and systems mature at different rates, and the R&D activities, corresponding to different hierarchical levels, require drastically different levels of funding, a common treading water strategy (i.e. when there are no resources to proceed at the system level) involves pursuing the less expensive R&D activities at the component level. This allows the team to write proposals for "early stage" resources, even though they are more than 10 years into the project. For example, when the CADR team struggled to get funding for the system level vibration testing they desired, they focused their efforts on the (much less expensive) thermal regulation control schemes.

Another behavior in this epoch involves "Branching Out." Rather than dropping the level of development to component R&D (as described above), the technologists recast their capability, as new in the context of a new application. For example, when the X-ray polarimeter pathway was struggling to get resources to sustain their mature capability, they wrote a proposal to develop a solar polarimeter. While there were legitimately some new technical hurdles associated with the new application, it also leveraged much of the technology they developed for the X-ray application and funded the extensions they would have wanted to develop anyway. In this way, the recasting was strategic rather than real. Nonetheless, a new solar polarimeter was initiated in this way. This is an example of how, both treading water and branching out occasionally lead to a discovery that initiates a new pathway (i.e., the arrow from treading water to gestation in Figure 5-1.)

Some of the forking also manifests as a reverse flow from exploitation to architectural exploration. If the innovation team has access to sufficient resources, it can loiter in exploration instead of treading water. As with the behaviors associated with treading water, this intentional drop in apparent maturity is at least partially a strategic move. We observed this reversal twice. Although it happened for very different reasons, both had similar implications. In the polarimeter case, following the AXP mission rejection, the concept was given a category 3 rating (i.e. money explicitly targeted to mature the capability so that it would be ready for the next flight opportunity). While this was explicit exploitation money, the lead scientist chose to use it to pursue a new technical path. Specifically, he returned to exploring an idea that would eventually make the capability far more functional. While the performance was improved, the resources were squandered leaving the capability no more mature than the earlier iteration had been when the next mission opportunity came along. This is an example where the team could have made forward progress but chose not to.

In the CZT case, the rejection of the BASIS proposal did not faze the team. Since the Gamma-ray group had access to significant slack resources, rather than treading water, they returned to the drawing board and continued exploring new detector technologies without missing a beat. This illustrates how the existence of a strong group can dampen the impact of contextual factors. This notion was also discussed in detail in the microcalorimeter case, in terms of the team's ability to emphasize and de-emphasize different aspects of their technical portfolio depending on the political climate.

5.1.2 Hierarchical Switchbacks

The next thickest arrows in Figure 5-1 embody the bi-directional flow between technology and architectural exploration. To understand the significance of this pattern, we must also consider the arrows initiating from gestation. Traditional views of the development process, formalized in measures like TRL, assume that new component ideas are developed, then integrated into systems architectures and used. However, in the systems we studied, more than half were initiated as architectural ideas. Specifically, the most common sequence was for a technologist, with an architectural idea, to initiate an architectural exploration. Progress would subsequently be stymied by a component level roadblock. That roadblock would be removed through technology exploration (i.e., solved with a new component) allowing architectural exploration to resume. For example, in the CADR case, the initiating technologist, while working on a flight ADR, realized that future missions would be limited by scaling laws in the current paradigm. His insight was to sequence ADRs in a temperature cascade to create continuous operation. Although the change was nominally a modular reconfiguration of existing technology, prototyping revealed that the new operational requirement required the invention of a new type of heat switch that could be operated over a wider range of temperatures. It was only after the new heat switches were created that the realm of the architectural possible could be fully understood.

The symmetry of this "push and pull" flow has implications both for the way maturity is conceptualized and how new ideas are integrated. In terms of maturity, the single TRL metric conflates progress at both the component and architectural levels, assuming that components are invented first and then integrated into increasingly complex subsystems and systems. In reality, innovation pathways bounce up and down the architectural levels over the course of their evolution. This leads to the switchbacks in perceived maturity evidenced in the funding plots aggregated in Figure 5-5. Rather than the monotonically increasing function that one would expect (since the funding levels nominally correspond to increasing TRL), in all cases, multiple levels of funding are uses simultaneously, and trajectories loop back to lower level funding sources later in the process. This is partially attributable to treading water and branching out behavior discussed above, but also explained by the exploratory fluidity discussed here.

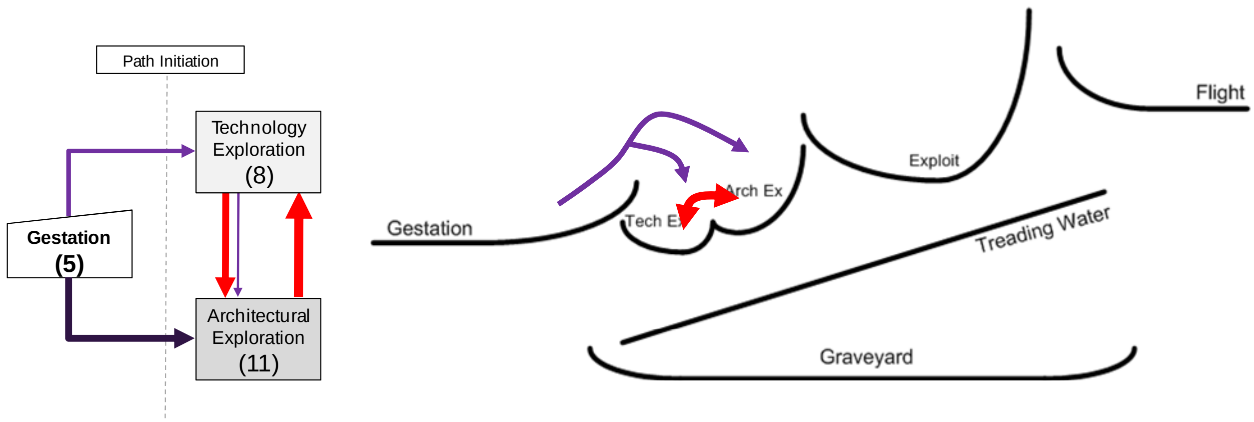

Figure 5-4: Hierarchical Switchbacks

In terms of new ideas, whether the path initiates in technology or architectural explorations seems to be related to the origin of the idea. If the insight is external in origin, the NASA team tends to dive directly into solving the revealed component challenge. If, on the other hand, the insight is internal, the architectural concept is explored first. This is because many of the science innovations are need-driven. NASA scientists and technologists work on both projects and R&D's and many of the R&D endeavors are inspired by lessons learned on the projects. Recall that the CADR technologist found his inspiration on the previous flight project he was working on. Since the innovations are need driven, the component developments are explored on an as required basis. This leads to a dynamic of technological satisficing. Specifically, the searches for component solutions seek to find adequate (rather than optimal) solutions and are guided by the materials and equipment that are readily available. For example, the semi-conductor microcalorimeter absorbers are made from HgTe because "*the alternatives we turned to were the materials that we had lying around an infrared lab.*" "*We don't have any illusions that what is being used now is in any sense optimal. It's something that works.*" [I88]

This dynamic of technological satisficing is perpetuated by pressures to be ready for the next mission opportunity, whenever it comes. Specifically, the trajectory that gets selected is the one that is most ready when a mission need arises (recall how QWIPs suddenly became the most feasible when production time rose in priority). This "moving on" with "good enough" solution accounts for both the ease of bouncing between technological and architectural explorations and also enables the re-scoping treading water behavior discussed above (i.e. there is always a known set of outstanding component problems to be addressed next time when progress at the system level lags). An extreme example of the former point is illustrated by the arrow from flight to architectural exploration in Figure 5-1. It draws the path of two failed Japanese X-ray missions which created opportunities to revisit the microcalorimeter pathway. In both cases, the team revisited known component level problems that had previously been ignored due to project pressures.

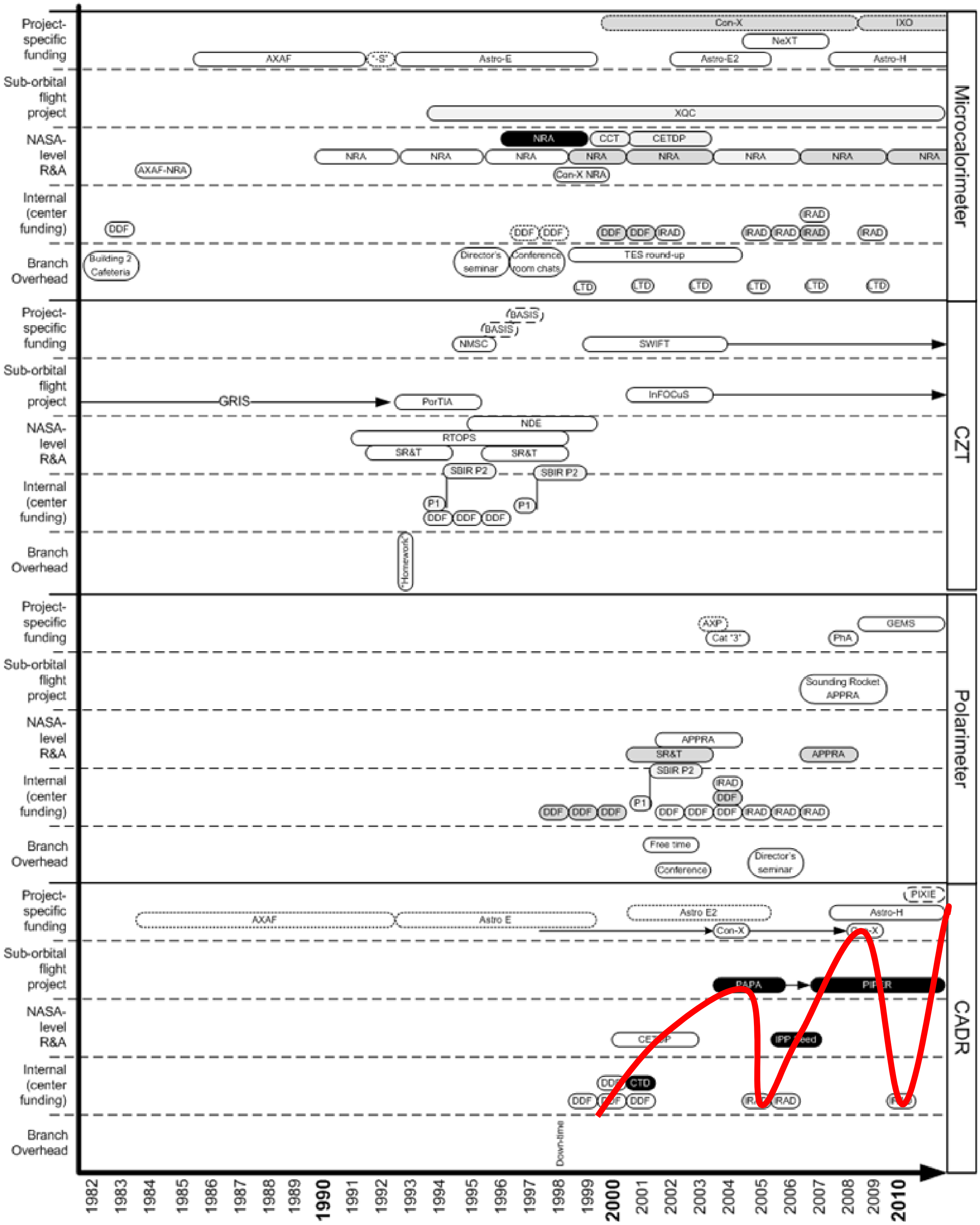

Figure 5-5: Switchbacks in Perceived Maturity

5.1.3 The road not travelled

To this point, the discussion has focused on high traffic pathways. We turn now to a conspicuously absent pathway, the road from treading water to the technology graveyard. It is absent, in part, do the selection criteria used; pathways were identified that had at least progressed to Milestone B (baselined for flight). Thus, it should not be a surprise that none of the five pathways, selected based on that criterion, terminated in the graveyard. That being said, there is a common perception that ideas never die, even when they should. And, many of the pathways came very close to terminating, but didn't. Insights can thus be gained from examining the teams' survival concerns while treading water. As discussed above the Polarimeter and CADR pathways faced the gravest risk of "dying," and in both cases, the concern was for loosing key team members. Excerpts from the cases, illustrate theses ideas richly:

> <u>*CADR (5 years from 2004-2009):*</u> Neither the branch head nor the champion CSE#1 – were ever concerned that the technical capability would become obsolete. [I23, 24] …her suggestion that the project be temporarily mothballed was for purely financial reasons; her job was to keep her staff funded, and money was extremely tight. [I23] …he was worried about losing one key technician – the expert in electric discharge machining – who was "***the kind of guy who would rather retire and work on his motorcycle***" [I24] than transition to another project while waiting for CADR funding to be restored. And rebuilding that kind of expertise would have taken a very long time. So, the CSE#1 found just enough funding to <u>keep the project alive</u>.
>
> <u>*Polarimeter (1.5 years from 2004-2006)*</u>: Once the AXP development funds ran out, the contractors wrote APRA (HQ level funding) proposals to sustain themselves and found "***day jobs***" to help free up resources to keep the project alive. Yet, by late 2005 the contractors were questioning the sanity of sticking with a program that couldn't pay them: "***frankly, we thought we were done. I was actively looking for a job, [the other contractor] had some job offers, but we decided to write two last proposals for the TPC stuff.***" [I40, I2, I55] Both APRAs were awarded in 2006, <u>keeping the team alive.</u>

These accounts highlight the reality that maintaining certain skills is critical, and if those skills are lost it will effectively terminate the pathway. What's interesting is that contractors and technicians – who have less flexibility with respect to charging hours to non-paying projects – are often the ones who fill these roles. They are the most likely to be reassigned to a different project when budgets get tight, and have the least task discretion which could give them the freedom to return once/if the climate changed. This idea of the one-way path away from an R&D effort is a critical one, because it explains why keeping the team together (treading water) is so important. Two other examples from the cases provide further support for the prevalence of these concerns.

> The QWIPs pathway – our pilot study, traced prior to the formal research design decisions and as a result not subject to the selection criteria – did sojourn in the graveyard, following the breakup of the original team. The pathway was restored when a policy change brought the old team back together.

IXO Mirror Assembly: In describing the impact of funding uncertainty on the microcalorimeter pathway, the IXO project manager recounted the following regarding another R&D intensive assembly: "*We've had fits and starts a) because of the money being erratic (for several years we've had midyear budget cuts that are really drastic… and then b) you have to* ***reduce your staff… and it's hard to build it back up again.*** *An example is that we had this really good mechanical team working on the mirror* [mounting… He] ***got called off because we're low priority****…* […, but] *technology requires the right kind of person, and it just takes time to get that* [new] *person up to speed… [I57]*

These examples illustrate the team – and tacit knowledge – aspect of path termination, and show how strong water treaders can survive nearly indefinitely. Conversely, if teams do not have the capacity, or do not recognize the importance of treading water, paths can easily be prematurely terminated. This calls into question the normal conception of shelf-life as a function of technical obsolescence. Clearly there is a people part that needs to be considered too.

5.2 Stage Gates vs. Epoch Shocks

An agency's R&D strategy is formulated based on underlying assumptions about how the system works. NASA's current management strategy assumes a stage gate-like system. However, our observations suggest that this representation is worse than just coarse; it's wrong. This section illustrates how strategies that presume a stage gates system may not lead to the intended result.

Recall that a stage gate system has three main conceptual elements: "stages" during which technologies are matured; "gates" where decision-makers choose which technologies to promote to the next stage; and "shelves" where partially matured technologies are stored for future reintegration into the process. Thus, the main policy levers include: (1) changing the relative resources allocated to the different stages; (2) creating more distinct stages; and (3) promotion decisions (i.e. how many and which technologies move on to the next stage and which get shelved).

While the elements of the epoch shock model are roughly analogous to those of the stage gate conceptualization – as shown in Figure 5-6, technology exploration roughly maps to basic R&D, architectural exploration, to applied R&D, exploitation to project specific development and treading water to the shelf – the dynamics they embody are fundamentally different. In order to illustrate the implications of these differences this section discusses the policies, as conceived in the stage gate view, and assesses their efficacy based on the epoch shock model.

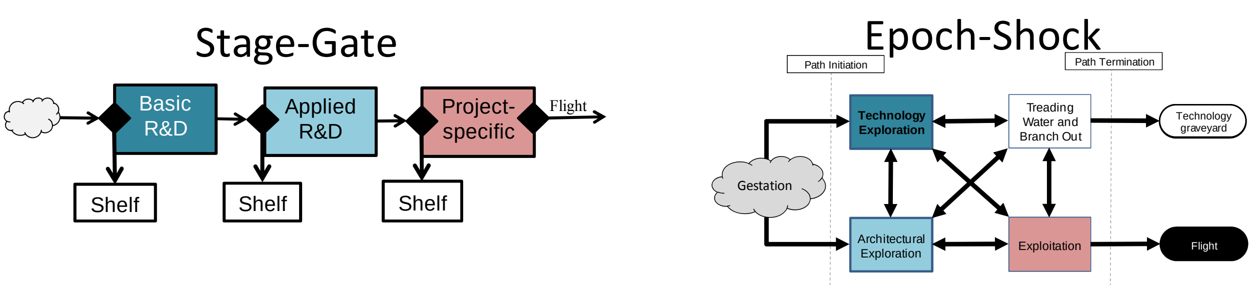

Figure 5-6: Comparison of Stage-Gates and Epoch-Shocks

5.2.1 *Relative Resource Allocation*

The first standard policy change under consideration is to add proportionally more funding for basic R&D. The idea is to increase the pool of new concepts that will be available for future capitalization on a mission. As reasoned in the stage gate view, the logic is clear: increasing funding in the first stage of a linear sequence will allow more new concepts to be explored and, as a result, more new good ideas to be identified. In the past, however, when similar policies have been implemented, the intended benefit has not been observed conclusively (see for example the policy swings illustrated in Figure 1-1). The epoch shock model allows us to understand why.

Our observations suggest that resources cannot be earmarked for early stage/basic R&D so simply. As discussed above, that basic funding stream will be split between a) truly novel concepts and b) other more mature concepts that are either i) treading water and branching out (i.e., they re-scope their work to focus on less expensive component level challenges or recast the technology as "new" in the context of a new application area) or ii) bouncing between component and architectural levels of exploration. While these diverted resources are not going to waste, it is important to realize that they are also not reaching the intended target of the policy. Thus, if the goal is really to encourage more early stage ideas to be pursued, a more targeted approach is required.

5.2.2 *Adding More Stages*

Another policy currently under discussion involves adding more stages to the process. The idea is that the challenges associated with the stage transitions will be mitigated if neighboring stages are more similar. For example, the transition from applied R&D to project specific development has been colorfully labeled the "valley of death" (Wessner, 2005). The idea is that there is both a wealth of resources available to early stage concepts as well as project funding for mature capabilities but there is a dearth of resources available to maturities in between. The extra stages would thus form a bridge across the valley of death, as illustrated in the below cartoon.

However, given that none of the pathways studied herein respect the nominal linear progression (see Figure 5-5) bridging transitions, as framed in a monetary way, may not solve the underlying problem. Firstly, there is a previously unrecognized, people side of the valley of death. Specifically, having the resources to make the transition does not guarantee that it will happen. Recall that in the Polarimeter case, the lead technologist opted to use his category 3 funding to pursue an exciting, albeit tangential, new path (i.e. exploring) rather than mature his existing capability. This speaks to a broader distinction between individual temperaments. Some scientists and technologists are natural explorers while others are natural exploiters. Explorers value new ideas and the potential for improvement, while exploiters value implemented results that materially improve the mission. Both orientations are critical to sustained innovation but it is important to recognize that the same incentives will work differently on each.

Second, the expectation of a linear progression is simply not respected in practice. In the epoch shock conceptualization, there is no need to build in additional epochs because the conceptualization does not impose linearity on the progression. We observed multiple instances of bouncing between technology and architectural exploration, reverse flow from exploitation to architectural exploration and even from treading water and flight. This new conceptualization allows us the flexibility to capture the chaos endemic to the system.

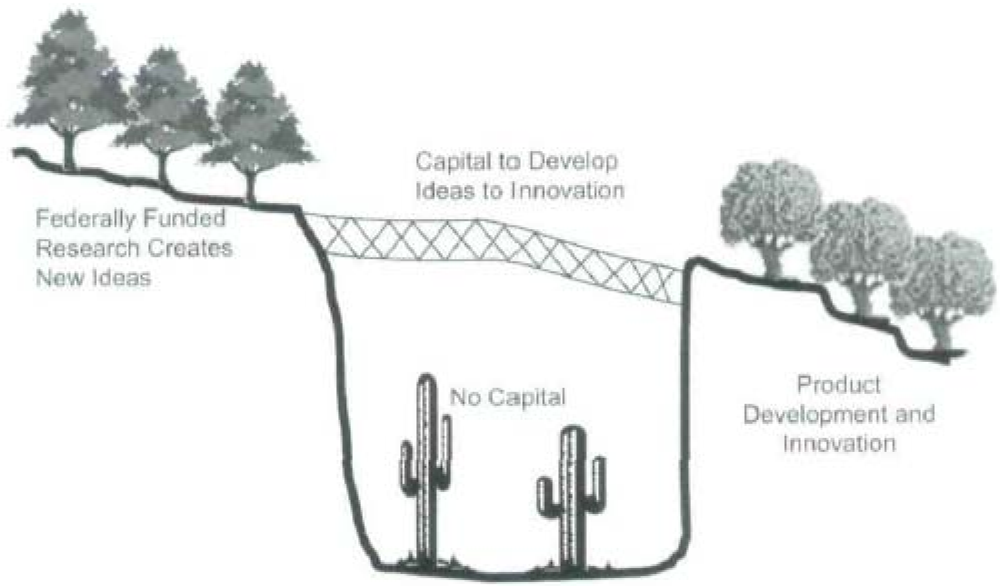

Figure 1.—Innovation award programs like the Advanced Technology Program (ATP) and the Small Business Innovation Research Program (SBIR) can help companies cross The Valley of Death (1).

Figure 5-7: Bridges across the Valley of Death

5.2.3 Promotion Decisions

Another commonly discussed intervention involves controlling the percent progression from one stage to the next. Based on the stage gate conceptualization, the idea is to control the flow of maturing concepts to achieve the desired output of new capabilities. This assumes that gates embody active decisions by a centralized decision maker. However, this is not the case in practice. Actively controllable gates just don’t exist.

A core difference of the epoch shock conceptualization is the probabilistic nature of the transitions. Where stages end when the technical maturity of the system reaches a sufficiently high level, transitions between epochs need to be induced by shocks. The shocks may embody technological breakthroughs but they can also relate to contextual changes and mission opportunities. For example, to the extent that there was a gate controlling the entrance to the treading water epoch (in the breakthrough window lag dynamic described above) it was based on the co-timing of a technical breakthrough (which is unpredictable) and the announcement of the next relevant mission call (which is semi-cyclical). Neither of these activities is controllable by any single decision-maker. To the extent that individual decision-makers can serve a winnowing function, it is by rejecting unworthy proposals. However, as we have discussed, technology teams tend to apply simultaneously to funding at multiple institutional levels, limiting the impact of any single rejection.

The concept of being shelved because of a rejection also manifests differently than expected. The baseline assumption that technologies can be restored from the shelf unless they have become obsolete (i.e. a better technology has superseded them) oversimplifies the situation. The team component of obsolescence must be considered as well. Specifically, an innovation pathway terminates in the technology graveyard when the team, and their associated tacit knowledge, breaks up. For example, in the CADR case, the lead technologist talked about his fear that one key machinist would retire to work on his motorcycle rather than join another project until funding for the CADR was restored. The implication of this observation is that even if decision-makers can shelve particular technologies, some level of maintenance is required to keep the team intact if they hope to restore the project at a later date.

5.3 Rethinking the Policy Problem

The previous section demonstrated how the Epoch Shock model provides new insights into why policies developed under Stage Gate assumptions may not produce the intended outcome. However, the prevailing explanation was an unsatisfactory: reality is more complicated. In this section, we show how this new lens allows us to rethink the policy problem in a prescriptively useful way. Following the traditions of (Cohen, March, & Olsen, 1972; Kingdon, 1984; Van de Ven, et al., 1999) we accept that there is significant uncontrollable stochasticity in the real system and focus on patterns of behavior that can be harnessed if understood. In other words, the epoch shock model seeks to clarify which processes are worth trying to influence and which must be recognized and accepted as unalterable.

The expedition analogy elaborated upon in section 4.2.5 provides a useful framework for thinking through potential interventions. Potential levers will be discussed at three levels: (1) Identifying and influencing key decisions made by individual actors in the system; (2) recognizing the institutional landscape that exists and what about it can actually be changed; and (3) playing with the forecastability of context and mission shocks, which aren't all as exogenous as they seem. To clarify the realm of potential actions and decisions, this section will focus on the trade-offs associated with interventions that could remedy the breakthrough window lag and the architectural switchbacks. In describing these interventions, we will draw heavily on the underlying systems architecture described in Chapter 2.

5.3.1 Smoothing the Breakthrough Window Lag

Recall the observation that when the transition from architectural exploration to exploitation is precipitated by a technical breakthrough, the pathway tends to stall in an epoch of treading water, unless a relevant flight opportunity crops up almost immediately. By the time a pathway reaches the exploitation epoch, significant resources will have been invested and general feasibility demonstrated. Thus, one wouldn't want pathways to terminate just because of a lag between two types of unpredictable shocks. The idea of slack has been studied in the context of decision processes under ambiguity (March & Olsen, 1976; Mintzberg, Raisinghani, & Theoret, 1976). One key insight from this literature is that the presence of slack changes the dynamics of trial-and-error-learning in that it enables technologists to continue pursuing an idea despite negative outcomes (and consequently overcome institutional inertia)(Garud & Van de Ven, 1992). In our context, some amount of slack is required (and desirable) serving to permit these technologies to survive the breakthrough window lag, but it's also important to recognize that too much slack is inefficient; one wouldn't want teams to tread water indefinitely.

To resolve this trade-off, three related interventions can be considered. The first question to consider is: Where should this slack be built in? Would it be better for teams to choose to delay a transition from architectural exploration to exploitation, to have access to dedicated support while treading water or to have an option to drop to architectural exploration if a window isn't imminent? The answer determines where you want the slack and how it can be created. For example, staying in the more stable architectural exploration epoch may be a matter of re-education. If technologists recognize the time constrained nature of entering the exploitation epoch (i.e. the death zone) they may be more reticent to enter without a clear summit strategy. The flip side is that technologists may never choose to adopt the risk (and loiter in exploration indefinitely). This is a problem because it would be nearly impossible for managers to identify these exploratory loiterers. Increasing support for treading water could have similar negative implications because of the resource drain associated with the branching out behaviors discussed above.

Actually targeting resources to create organizational slack is a more complicated issue. Unlike the stages of the stage-gate model, which are defined in terms of technology maturity, the definition of epochs incorporates technical activities, manpower movements and resource patterns. While epochs and shocks better captures reality, they provide less clarity in terms of policy insights. As a first step towards the solution, we need to recognize at least two dimensions of resource targeting. There is maturity of the technology as well as the level of the technical hierarchy. This is important because the magnitude of resources required correlates both with increasing maturity and increasing system complexity. As illustrated in Figure 5-8, moderate resources are useful for both immature systems and mature components. Thus, if the goal is to create slack without poaching funding targeted at early stage concepts, the technology readiness ladder needs to incorporate the second dimension as well.

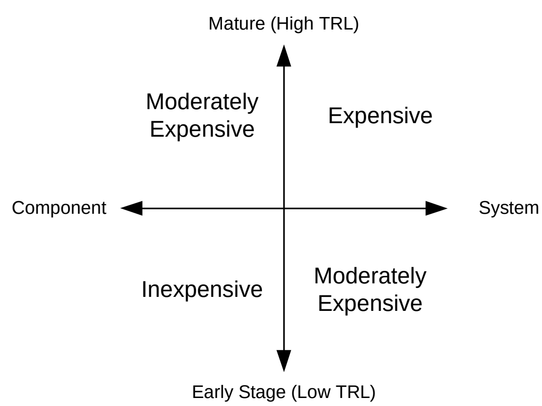

Figure 5-8: Two dimensions of R&D costs

The second question is: How predictable should mission windows be? Part of the challenge of allocating slack is the uncertainty associated with the timing of the next window. Within NASA there are nominally regular intervals between missions. Flagships are on a decadal cycle, each focusing on a different mission area, and Explorers, which are supposed to be announced every few years and are mission area agnostic. In reality, the spacing of the Explorer calls is highly dependent on the budgetary pressures imposed by flagship overruns. As a result, historically, the calls have been unpredictable. Given the importance of the timing of the windows, and the monopsony position of NASA in the space science market, there is no reason why the windows shouldn't be more predictable. In terms of the expedition analogy, when climbers enter the "death zone," no matter how prepared they are, their success hinges on a window of favorable weather in time for a summit attempt; unlike on the mountain, NASA administrators have the potential to control the weather. This is not to say that NASA can control the length of the lag (the technical breakthrough is still unpredictable) but it can lessen the uncertainty associated with the wait.

The third question is: Should there be a mechanism for path termination? At minimum, we need to recognize that there is currently no mechanism for management to actively shelve loitering technologies. While some slack is important, we observed instances of treading water that lasted five or more years. Given that survival requires the team to stay together, this is a significant resource drain. Implementing a centralized terminating mechanism, though, would require major restructuring. It would likely undermine the decentralized market flavor of NASA's R&D system and restrain the shadow organization of stable research groups that are one of NASA's hallmarks. To understand the issue we need to consider how the system currently works and what changing it might mean.

The notion of group formation played importantly in both the CZT and microcalorimeter pathways. In the former, the technology development was protected from two explorer proposal rejections because they had their own access to slack resources. In the latter, as the team solidified and gained favorable status, they were able to expand their technology portfolio, which protected against major context changes. For example, when Con-X was touted as the next grand X-ray observatory, they downplayed their semi-conductor microcalorimeters and emphasized the next generation TES. Then, when the tides changed and Con-X was delayed and Astro E2 failed, they emphasized how their semi-conductor technology was perfect for NeXT. This is in stark contrast to the experiences of the CADR and polarimeter teams, who, as new technology areas, were put into jeopardy by what were arguably comparatively minor context changes. The idea of groups as stability draws important parallels to Leonard-Barton's concept of core competence and core rigidities (Leonard-Barton, 1992). The trade-off is that the stability advantages associated with strong groups are also a major drawback in that these groups are both more likely to continue along similar technology trajectories and their technologies become increasingly difficult to kill.

NASA has intentionally cultivated a relatively free market for R&D funding. Each proposal is nominally evaluated on its own merits each time it is submitted. Resource distributors explicitly try to focus on the proposals at hand without consideration of the broader system and future opportunities. As explained by one resource manager, it's not his job to advocate for promising technologies; if the mission guys ask him for input, he'll give an assessment, but he feels that it is important that he remains separate and impartial. Of course, some amount of coordination does occur. A clear example of this was described in the TES pathway. When Con-X's R&D funds were cut, much of the development funds were provided through the NRA program; in this context NRA program managers communicated with Con-X leadership to understand which technologies were their frontrunners and how much total resourcing they were getting. This current environment enables individual innovation teams to hedge their bets by applying for resources at multiple institutional levels simultaneously; but it comes at the cost of potentially significant inefficiencies. In this current structure, meaningful portfolio planning is nearly impossible and no single element manager has the power to shelve or promote any particular technology. This is why adding a winnowing mechanism would require a significant centralization of decision authority. It is impossible to assess what kind of adverse impact that would have on the current system.

5.3.2 Encouraging Path Initiation

Many of the current interventions that are being considered seek to increase the number of new ideas that are explored. However, in section 5.2.1 we argued that these interventions wouldn't be effective because they target the wrong problem. Specifically, we observed that in the current system the availability of small amounts of funding suitable for exploring these early stage ideas is already quite high. In other words, technologists who are committed to pursuing an idea are not really constrained in the early stages. Neither is the problem one of a lack of ideas. Every technologist interviewed listed numerous brainstorming ideas past, present and future. The key issue that needs to be understood relates to the conditions that cause a technologist to push a particular idea forward and the availability of pre-early stage slack needed to decide whether an idea is worth pursuing. This section also addresses the related question of whether there is a role for management to play in encouraging this initial leap.

First, recall that we observed a relatively even split between paths initiated with technology vs. architectural exploration and that the flow between the two exploration epochs was fluid. In all cases, the path initiation shock was a composite technical insight and mission opportunity. Superficially one could argue that this is a canonical window merging a problem and solution stream. However, the more important insight is at the level of how the streams combine and what this means about organizational structure. Unlike either the DoD and the European Space Agency (ESA), which explicitly separate their R&D organization from their project processes, NASA's differentiation happens at the individual level. Specifically, a scientist or technologist will spend parts of his or her time on R&D and the rest on flight projects. The relative mix of the individual's time is determined organically by their relative preferences and propensity for grant writing. The result is a normal career progression that involves more project work early on, increasingly R&D focused in the middle and a transition to more management or project work depending on the success of the R&D stage. Stable R&D groups change this dynamic by hiring some individuals into a baseline of high proportion R&D work.

In theory, having individuals bridge R&D (exploration) and project work (exploitation) creates a strong feedback mechanism for lessons learned and mission needs, and makes the joining of the streams more robust in terms of timing mismatches between breakthroughs and windows. In fact, many of the mission breakthrough couplings were achieved at the level of individuals, rather than a question of good timing. For example, the CADR technologist had been developing flight ADR's when he noticed the limitation of the current approach with respect to future requirements. This allowed him to appreciate the value of his technical insight and be prepared when the imminent change in needs occurred. As a result, it wasn't necessary for his insight to coincide exactly with a mission need. Similarly, the microcalorimeter team was involved with the requirements definition discussions for Con-X even before the initial decadal call. These examples illustrate the value of forcing individuals to span the R&D and flight divide and as a result anticipate and smooth out timing mismatches between breakthroughs and mission opportunities.

It is important to recognize, though, that this spanning is not enforced in practice. Because these individuals have substantial discretion in how they spend their time, some of the best innovators manage to avoid the spanning altogether. When this happens, the NASA system 'as implemented' is more similar to the ESA/DoD system 'as conceived' (where R&D is separate from projects). This can result in the worst of both worlds: NASA doesn't get the advantages of having specialized, and separate, R&D and Project organizational subunits, but neither does it have constant feedback path through individual spanners. An important strategic question is thus, should individuals be forced to span (i.e. should there be a requirement for individuals to spend some of their time on projects.) Or should the free market for time allocation be allowed to control the division. To really address this key tradeoff will require a more focused study and likely some modeling work.

Second, there is an important difference between early stage resources and the freedom to brainstorm (which doesn't necessarily require monetary support). While there are ample instances of the former, the latter are becoming increasingly scarce and their importance cannot be overlooked. Given the personal investment associated with undertaking an innovation expedition, the technologist needs to convince him or herself of the viability of the concept before formally initiating the journey. Prior to NASA's adoption of full cost accounting, individuals did not have to account for every minute of the day and were encouraged to (or at least not discouraged from) spending a small amount of time exploring novel ideas. For example, one scientist, when recalling the good old days, told how he used to walk down the hall to the machine shop and get any technician who was free to mock up the idea he was currently working on. Within a very short time he would have a good sense whether or not the idea was worth pursuing. Nowadays, even if that machinist is sitting idle, he would not do the work without an explicit charge code. It is hard to value the extent to which these tiny amounts of resources enabled innovation, but it is noteworthy that nearly every scientist and technologist we interviewed commented on how stifling they find the current requirements to write proposals to account for their hours during this important brainstorming phase.

5.4 Broader Relevance of Findings

In the preceding discussion, five potential feasible interventions were identified: (1) ways to build in slack; (2) more predictable timing of mission opportunities; (3) mechanisms to terminate pathways; (4) enforcing time-splitting between R&D and project activities; and (5) allowing free time for brainstorming. The potential implications of these changes were discussed in the context of NASA's Science Directorate. More important, though, than the NASA-specific interventions, is the basis that the Epoch Shock model provides for exploring future strategic trade-offs. This section explores the extent to which we expect these findings to be relevant more broadly and why.

Previous work by the author (Szajnfarber, et al., in press) identified the set of unique characteristics that define the space context. These included the complexity of the project being developed, the monopsony structure of the market, and the limited access to, and harshness of, the normal operating environment. The paper theorized ways in which these characteristics would affect, the fundamental dynamics of innovation, in this context. Now, having conducted an empirical study, we use these ideas as a basis for generalizing the results.

First, in terms of the nature of the product, the innovations studied herein were largely key component changes that could be integrated into a relatively stable system. In other words, a revolutionary change at the level of a sensor could have a major impact on the satellite functionality but if it did not work it could be changed out for the incumbent without a huge amount of rework. Also, compared to the project costs (on the order of $100's of millions to billion) the pre-infusion R&D is relatively inexpensive (totaling at most in the $10's of millions). The implication of these characteristics is that the R&D can reasonably proceed relatively independently as an unmanaged free market.

Second, the space market is characterized by a monopsony-oligopoly (i.e. one buyer and few sellers). This creates a discrete and specific market which only exists when the buyer wants to buy. As a result, user needs must be specified explicitly and transactions only occur occasionally. The buyer (i.e. NASA) must bear the R&D burden and innovators must take advantage of the mission opportunities as they arise. This monopsony dynamic influences the model we induced in terms of the shock structure and the lack of technological obsolescence. Unlike in a competitive market where it is reasonable to assume that useful innovations will be used as they come available, space-specific innovations will only ever be used on a particular mission. With respect to obsolescence, while NASA is certainly not the only innovator in this field, the competition is so strongly tied to the interests of the government, when there is a shortage of resources, it affects everyone.

The last characteristic, the harsh and unserviceable environment, has less of a direct impact on the epoch shock model because we did not consider the operational phase explicitly. Nonetheless, as a governing reality of this context, it influences many of the expensive aspects of the exploitation phase and has a strong second order influence on the discrete opportunities.

While it may seem on first read that these criteria impose stringent restrictions on the generality of the theory developed herein, these types of systems represent an important class. In addition to most space developments (which account for an annual budget of more than $50 billion in the U.S.), it is also likely applicable to other defense acquisitions, infrastructure investments and large energy technologies, like nuclear power technologies, wind turbines etc. Testing this generality is an important area of future work.

It should also be noted that while outside of the direct characteristics of this market, a comparison to other documented innovation pathways, reveals similar underlying patterns. Specifically, the Minnesota Innovation Research Project (Van de Ven, et al., 1999), identified cycles of convergence and divergence in the development of specialized medical devices. We noted similar patterns at NASA even though the organizational structures were quite different. Further comparisons like this may provide a more nuanced understanding of uniqueness of the space sector.

6 Wrapping Things Up

Despite a rich legacy of impressive technological accomplishments (e.g., project Apollo, the Hubble Space Telescope) in recent years, the ability of government space agencies to deliver on their promises has increasingly been called into question. Although multiple acquisition systems and organizational structures have been tried, there remains a fundamental lack of understanding of how new technology development can, and should, be encouraged in this unique market structure and product context. This thesis set out to address that gap, by developing a more nuanced, and empirically grounded, explanation.

The empirical basis for this work stems from a retrospective multi-case study of six NASA technology *innovation pathways*, each spanning decades of pre-project development. The pathways were constructed from a combination of more than 100 hours of expert interviews, cross-referenced with 150 archival documents, and the equivalent of two months of informal observations. Each case was validated by multiple respondents. A visual mapping framework was developed which allowed the process data to be plotted, and enabled a structured cross-comparison of the cases.

This strategy resulted in two methodological contributions. Firstly, it demonstrated a quasi-real-time process tracing technique for studying governmental organizations. Within the process tracing paradigm, there is an ongoing debate regarding the relative merits of real time vs. retrospective approaches. The key advantage of real time, or forward looking approaches, is that they are not biased by the clarity imposed by history. However, the study must last over the time period that it takes the process to unfold. This is clearly infeasible with the innovation pathways studied herein, which unfolded over two or three decades. Nevertheless, we found that a nearly real-time lens could be achieved in the context of governmental engineering systems by tracing contractual documents produced on a yearly basis.

Second, it showed that the innovation is a viable unit of analysis. The process tracing method has previously been used by management scholars and political scientists. In their applications the levels of potential unit of analyses typically include the individual, the team, the business unit, the organization, the sector or the nation. We chose, instead, to use the innovation itself as the unit of analysis. This choice allowed us to consider together events at the political, agency and technology level, simultaneously, which is necessary in our context. This approach and unit of analysis is expected to be useful for studying engineering systems more broadly.

Individually, each innovation pathway represents an important empirical contribution since the pre-infusion technology development process has not previously been described. These pathways provide an important insight into the link between macro-level policies and micro-level behaviors. In particular this work uncovered the following behaviors:

- **The impact of technologist strategies for keeping the team together.** Where, from a management perspective, obsolescence has been previously thought of as a characteristic of the technology, this research illustrates the human element of shelf-life. Technologies can only survive on the shelf as long as the team tacit knowledge remains intact. Recognizing this, innovation champions engage in a number of different strategies for keeping the core team together during periods of scarce resources. These include re-scoping to work on less expensive component technologies, or branching out to related applications and recasting the current work as new. These strategies can maintain team competence for years, and are relatively immune to policy changes designed to re-allocate resources to different stages of development. This explains a significant portion of the maturity regressions observed at the system level.

- **The impact of allowing the market for grant funding to distribute human resources across R&D and project work.** Within NASA, PhD scientists are afforded significant autonomy. They are expected to spend some proportion of their time on R&D and the rest on flight projects. In practice, the proportion is largely determined by an individual's success in writing grants to support their R&D time. The existence of time-splitting is important because it serves as a main feedback mechanisms between the two organizational modes. However, the cyclicality of R&D funding and the high incentives to win freedom through grants, means that time-splitting doesn't always happen at any one time. For the most part, NASA scientists are mission focused, and as a result are motivated to implement the technologies they develop, but some are more concerned with the next exciting technology. The latter group may never transition a useful technology without external direction, and there is currently no mechanism to enforce the transition.

This empirical evidence was used to assess the validity of current conceptualizations of the innovation process. R&D management practices typically conceptualize complex product innovation as a Stage-Gate process whereby novel concepts are matured through a succession of development stages and progressively winnowed down at each sequential gate. This view implicitly assumes that maturity is a monotonically increasing function of the technology, and that partially matured technologies can be restored from the "shelf" for future maturation, baring obsolescence. However, in practice, the pathways taken by new capabilities do not respect these assumptions, with important implications. Specifically:

- Rather than being a monotonically increasing function in time, particular innovations draw simultaneously from funding mechanisms targeted at different TRL ranges, and loop back to win "early stage" grants decades into their development when system level progress is stymied. As a result, increases to early stage R&D funding may not reach the targeted concepts.

- Rather than being a purposeful management decision, getting shelved is something that happens to innovation teams due to the lack of co-timing of technical breakthroughs and mission opportunities. As a result, maintaining a shelved capability is as much a matter of keeping the team together as it is a question of technical obsolescence.

Instead, it was argued that the observed dynamics can be better explained by four epochs of persistent, stable, and identifiable behavior, punctuated by transition inducing shocks. Acting individually, or in combination, these shocks can induce transitions from any one epoch to another. This new Epoch-Shock formulation enables a re-thinking of the policy problem. Where the Stage-Gate model leads to an emphasis on centralized flow control, the Epoch-Shock model acknowledges the decentralized, probabilistic nature of key interactions and highlights which aspects may be influenced. In particular, five potential feasible interventions, and their associated tradeoffs, were identified:

(1) Building flexibility into the system so that teams can survive breakthrough-window lags;
(2) Improving predictability of mission opportunities, in order to mitigate lag uncertainty;
(3) Mechanisms to terminate pathways or encourage branching out if treading water persists too long;
(4) Enforcing some level of time-splitting between R&D and project activities or developing an alternative feedback mechanism; and
(5) Facilitating free time for brainstorming, not just early stage funding.

The potential implications of these changes were discussed in the context of NASA's Science Directorate, and more generally, in terms of their implications for complex product innovation in a monopsony market.

These conclusions take the form of articulations of key tradeoffs rather than clear implementable prescriptions. While this may seem, at first pass, to be an unsatisfying result, its contribution as an important first step should not be underestimated. This research has shown why current simplified models do not capture the fundamental dynamics of the system, and consequently may lead to inappropriate interventions. The new model clarifies the tradeoffs inherent in harnessing an extremely complex and messy system. As an exploratory study, this research lays the foundation for more focused future investigations, which can begin to develop the more easily implementable strategies which are the eventual goal.

This future work will require multiple methods and a broader set of cases. Three natural extensions are outlined here. First, the methodological contribution of adapting the process tracing method for application to complex engineering systems can be leveraged to study technology innovation in other settings. For example, ESA's development of the X-ray Astrophysics Observatory Xeus represents a perfect natural experiment. Many of the enabling technologies studied in this thesis were developed for infusion onto NASA's Con-X mission, which have sister technologies under development, over the last two decades, on the other side of the Atlantic and in Japan for Xeus. In 2008, Con-X and Xeus were merged under the guise of the International X-ray Observatory (IXO). If IXO moves forward, an expert selection of which technology to use will be made. What makes this quasi-experiment particularly exciting is that differing contextual factors drove teams in the different countries to pursue different technological paths, relatively independently. For example, where the US chose to pursue CdZnTe in the early 1990's, Japan and Europe have focused their development efforts on CdTe. By comparing the pathways taken by these equivalent functional technologies, one could both develop important insights about the ESA/Jaxa innovation ecosystems while also leveraging the cross-comparison to identify implications of the different institutional approaches.

Second, one of the key motivations for pursuing an in-depth qualitative study of innovation at NASA was the exploratory nature of the work. There was simply insufficient prior knowledge of the phenomenon to develop measureable metrics for a large N quantitative study. Now that this research has developed propositions about the dynamics that underlie innovation at NASA it is feasible to follow up with a theory-testing survey study. For example, it would be interesting to collect data on the career progressions of the different actor/archetypes (e.g. technologists, observers, instrumentation specialists etc.) to see if the proposed patterns hold more broadly.

Third, empirical research, in the quasi-experimental paradigm, is fundamentally limited by the extent of historical variability. In other words, we will never be able to discover alternatives that have not yet been tried. A natural response to these limitations is to use computer experiments to test theoretical alternatives. However, in the domain of innovation the success of the simulation work has been limited; there are just too many contributing factors that may be important. As a result, modeling work has largely focused on diffusion patterns which take place after the innovation has been developed, assume random interactions or focus on equilibrium solutions. However, particularly in the space sector, getting to the first implementation in a highly non-equilibrium market is the interesting part. What this research contributes are theoretical propositions about types of actors and the interaction routines that govern their behaviors. Coupled with recent advances in modeling frameworks, like Agent-Based modeling, there is an interesting opportunity to make contributions both to the field of computational social science and to extend this nascent theory of innovation in space agencies developed herein, making it more useful for guiding institutional policy.

References

Adams, W., & Adams, W. J. (1972). The Military-Industrial Complex: A Market Structure Analysis. *The American Economic Review, 62*(1/2), 279-287.

Anderson, P., & Tushman, M. L. (1990). Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change. *Administrative Science Quarterly, 35*(4), 604-633.

Augustine, N. R., Austin, W. M., Chyba, C., Kennel, C. F., Bejmuk, B. I., Crawley, E. F., et al. (2009). *Seeking a Human Spaceflight Program Worthy of a Great Nation*. Washington D.C.

Beard, E. (1976). *Developing the ICBM: A Study in Bureaucratic Policitics*. New York: Columbia University Press.

Braun, R. D. (2010). *Investments in the Future: NASA's Technology Programs*. Greenbelt, MD: NASA.

Brown, S. L., & Eisenhardt, K. M. (1998). *Competing on the Edge: Strategy as Structured Chaos*. Cambridge MA: Harvard Business School Press.

Burgelman, R. A. (2002). Strategy as vector and the inertia of coevolutionary lock-in. *Administrative Science Quarterly, 47*, 325-357.

Buthe, T. (2002). Taking Temporality Seriously: Modeling History and the Use of Narratives as Evidence. *American Political Science Review, 96*(3), 481-493.

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A Garbage Can Model of Organizational Choice. *Administrative Science Quarterly, 17*(1), 1-25.

Collier, R. B., & Collier, D. (1991). *Shaping the Political Arena. Critical Junctures, the Labor Movement, and Regime Dynamics in Latin America*. Princeton: Princeton University Press.

Cooper, R. (1990). *Winning at new products*.

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *The Academy of Management Review, 14*(4), 532-550.

Falletti, T. (2006). Theory-Guided Process-Tracing in Comparative Politics: Something Old, Something New. *APSA-CP, Newsletter of the Organized Section in Comparative Politics of the American Political Science Association, 17*(1), 9-14.

Freeman, C. (1987). *Technology and Economic Performance: Lessons from Japan*. London: Printer.

Garud, R., & Van de Ven, A. H. (1992). An Empirical Evaluation of the Internal Corporate Venturing Process. *Strategic Management Journal, 13*, 93-109.

Gehman, H. W. J., Barry, J. L., Deal, D. W., Hallock, J. N., Hess, K. W., Hubbard, G. S., et al. (2003). *Columbia Accident Investigation Board Report*. Washington DC: NASA.

Gersick, C. J. G. (1991). Revolutionary Change Theories: A Multilevel Exploration of the Punctuated Equilibrium Paradigm. *Academy of Management Review, 16*(1), 10-36.

Glaser, B., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies of qualitative research*. London: Wiedenfeld and Nicholson.

Government, U. S. (2005). *Federal Acquisition Regulation (FAR)*.

Greve, H. R. (2007). Exploration and exploitation in product innovation. *Industrial and Corporate Change*, 1-31.

Grigsby (2009). *Landsat Data Continuity Mission HQ Actions/Issues*: NASA.

Grissom, A. (2006). The Future of Military Innovation Studies. *The Journal of Strategic Studies, 29*(5), 905-934.

Guglielmi, M., Williams, E., Groepper, P., & Lascar, S. (2008). *The Technology Management Process at the European Space Agency*. Paper presented at the International Astronautical Congress.

Gupta, A. K., Smith, K. G., & Shalley, C. E. (2006). The Interplay Between Exploration and Exploitation. *Academy of Management Journal, 49*(4), 693-706.

Hall, P. A. (2003). *Aligning Ontology and Methodology in Comparative Politics*. New York: Cambridge UP.

Henderson, R., & Clark, K. (1990). Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms. *Administrative Science Quarterly, 35*, 9-30.

Jhabvala, M., Reuter, D., Choi, K., Jhabvala, C., & Sundaram, M. (2009). QWIP-based thermal infrared sensor for the Landsat Data Continuity Mission. *Infrared Physics & Technology, 52*, 424-429.

Jhabvala, M., Reuter, D., Choi, K., Jhabvala, C., & Sundaram, M. (2009). *QWIPs-based Thermal Infrared Sensor for the Landsat Data Continuity Mission*. Greenbelt, MD: NASA.

Kingdon, J. W. (1984). *Agendas, Alternatives and Public Policies*. Washington DC.: Addison-Wesley Educational Publishing Inc.

Langley, A. (1999). Strategies for theorizing from process data. *The Academy of Management Review, 24*(4), 691-710.

Langley, A., & Truax, J. (1994). A process study of new technology adoption in smaller manufacturing firms. *Journal of Management Studies, 31*, 619-652.

Launius, R. D., & McCurdy, H. E. (1997). *Spaceflight and the Myth of Presidential Leadership*: University of Illinois Press.

Lawler, A. (2009). Trouble on the Final Frontier: NASA's scientific missions have enjoyed spectacular success. But significant cost overruns and lanuch delays jeopardize future missions. *Science, 324*.

Leonard-Barton, D. (1992). Core Capabilities and Core Rigidities: A Paradox in Managing New Product Development. *Strategic Management Journal, 13*, 111-125.

Lindsay, J. (2006). War upon the Map: The Politics of Military User Innovation. MIT.

Logsdon, J. M. (1970). *The Decision to go to the Moon: Project Apollo and the National Interest*. Cambridge: MIT Press.

Lundvall, B. A. (1992). *National Innovation Systems: Towards a Theory of Innovation and Interactive Learning*. London: Printer.

Mankins, J. (1995). *Technology Readiness Levels*: Advanced Concepts Office, NASA.

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. [Special Issue: Organizational Learning: Papers in Honor of (and by) James G. March]. *Organizational Science, 2*(1), 71-87.

March, J. G., & Olsen, J. P. (1976). Organizational learning and the ambiguity of the past. In J. G. March & J. P. Olsen (Eds.), *Ambiguity and Choice in Organizations* (pp. 54-69). Bergen: Universitetsforlaget.

McDougall, W. A. (1985). *... the Heavens and the Earth: A Political History of the Space Age*. USA: Basic Books Inc.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage.

Mintzberg, H. (1979). An emering strategy of "direct" research. *Administrative Science Quarterly, 24*, 580-589.

Mintzberg, H., Raisinghani, O., & Theoret, A. (1976). The structure of unstructured decision processes. *Administrative Science Quarterly, 21*, 246-275.

Mohr, L. B. (1982). *Explaining Organizational Behavior*. San Francisco: Jossy-Bass.

Morison, E. E. (1966). Gunfire at Sea: A Case Study of Innovation *Men, Machines, and Modern Times* (Vol. Chapter 2). Cambridge: MIT Press.

NASA (2004). *The VIsion for Space Exploration*.

NASA (2007). *NASA Systems Engineering Handbook*.

NRC (2001). *Astronomy and Astrophysics in the New Millennium*.

NRC (2007). *SBIR and the Phase III Challenge of Commercialization: Report of Symposium*.

Nutt, P. (1984). Types of organizational decision processes. *Administrative Science Quarterly, 29*, 414-450.

O'Reilly, C., & Tushman, M. L. (2007). Ambidexterity as a Dynamic Capability: Resolving the Innovator's Dilemma.

Peck, M. J., & Scherer, F. M. (1962). *The Weapons Acquisition Process: An Economic Analysis*. Cambridge: Harvard Business School School Press.

Pentland, B. T. (1999). Building process theory with narrative: From description to explanation. *Academy of Management Journal, 24*, 71-724.

Pettigrew, A. M. (1990). Longitudinal field research on change: Theory and practice. *Organization Science, 1*, 267-292.

Posen, B. (1984). *The Sources of Military Doctrine: France, Britain, and Germany between the world wars*. Ithaca: Cornell University Press.

Romaneli, E., & Tushman, M. L. (1994). Organizational Transformation as Punctuated Equilibrium: An Empirical Test. *Academy of Management Journal, 37*(5), 1141-1166.

Rosen, S. P. (1994). *Winning the Next War*. Ithaca: Cornell University Press.

Rothwell, R., & Zegveld, W. (1994). Reindustrialization and Technology. In R. Rothwell (Ed.), *Towards the Fifth-generation Innovation Process* (Vol. 11, pp. 7-31): International Marketing Review.

Rubin, H. J., & Rubin, I. S. (2005). *Qualitative Interviewing: The Art of Hearing Data*. Thousand Oaks: Sage.

Sapolsky, H. M. (1972). *The Polaris System Development: Bureaucratic and Programmatic Success in Government*. Cambridge: Harvard University Press.

Sauser, B., Ramirez-Marquez, J., Magnaye, R., & Tan, W. (2008). A Systems Approach to Expanding the Technology Readiness Level within Defense Acquisition. *International Journal of Defense Acquisition Management, 1*(3), 39-58.

Sherwin, C. W., & Isenson, R. S. (1967). Project Hindsight. *Science, 156*(3782), 1571-1577.

Smith, K., & Tushman, M. L. (2005). Managing Strategic Contractions: A Top Management Model for Managing Innovation Streams. *Organization Science, 16*(5), 522-536.

Sonenshein, S. (2010). We're Changing - Or are We? Untangling the Role of Progressive, Regressive, and Stability Narratives During Strategic Change Implementation. *Academy of Management Journal, 53*(3), 477-512.

Stone, D. (2002). *Policy Paradox*. New York: Norton & Co.

Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: grounded theory procedures and techniques*. Newbury Park, CA: Sage.

Szajnfarber, Z., Richards, M. G., & Weigel, A. L. (in press). Challenges to Innovation in Government Space Agencies. *Defense Acquisition Review Journal*.

Tushman, M. L., & Romaneli, E. (1985). Organizational evolution: A metamorphosis model of convergence and reorientation. In L. Cummings & B. Staw (Eds.), *Research in Organizational Behavior* (Vol. 3, pp. 171-222). Greenwich, Conn: JAI Press.

Tushman, M. L., & Smith, K. W. (2002). Organizational Technology. In J. Baum (Ed.), *Companion to Organizations* (pp. 386-414). Malden MA: Blackwell.

Ulrich, K. T., & Eppinger, S. D. (2008). *Product Design and Development*: Andy Winston.

Van de Ven, A. H., Angle, H. L., & Poole, M. S. (2000). *Research on the management of innovation: The Minnesota studies*: Oxford University Press US.

Van de Ven, A. H., Polley, D. E., Garud, R., & Venkataraman, S. (1999). *The innovation journey*. New York: Oxford University Press.

Wertz, J. R., & Larson, W. J. (2007). *Space Mission Analysis and Design*. USA: Microcosm Inc.

Wessner, C. W. (2005). Driving Innovations Across the Valley of Death. *Rsearch Technology Management, 48*(1), 9-12.

Yin, R. (2009). *Case Study Research: Design and Methods* (Vol. 5): Sage.

Appendix I: Innovation Pathway Shortlist

Innovation	Selected	Rationale/Comments
Quantum Well Infrared Photodetectors	Y	Pilot study
Pop-up Bolometers	N	Studied tangentially to microcalorimeters but not a focus. Microcalorimeters have better pairing with other X-ray technologies
Spacewire	N	Chose to focus on hardware system
Microcalorimeters	Y	Split into two cases: semiconducting and superconducting microcalorimeters = time-spaced pair
Cadmium Zinc Teluride detectors	Y	Closest contemporary to semiconducting microcalorimeter
X-ray polarimeter	Y	Contemporary with TES microcalorimeter and CADR, parallel target mission to CZT
Continuous Adiabatic Demagnetization Refrigerator	Y	Shares mission context with semiconducting microcalorimeters, and level of integration with X-ray polarimeter
Microshutters	N	JWST project pressures made study impossible
Segmented Glass Optics	N	Backup case, not used in the end
Wave front sensing	N	Chose to focus on hardware system
Corelated Noise Diode	N	Chose to focus on Astrophysics theme for comparability reasons
Radiofrequency Interference Mitigation	N	Chose to focus on Astrophysics theme for comparability reasons

Appendix II: Summary of Interview and Document References

Interview References

Interview#	Code	(Relevant) Functional Title	Case
I1	CSA#2	Staff Scientist, COTR	Microcalorimeters, General
I2	CSE#1	Detector Technologist	QWIPs
I3	FM#9	Center-level Funding Manager	Multiple
I4	CSS#2	Scientist on unused case, COTR	General
I5	CSS#3	Scientist on unused case, COTR	General
I6	SB#1	SBIR Company	QWIPs
I7	PM#7	(Past) Director of Code R	General
I8	SB#2	SBIR Company	General
I9	SL#1	NASA Senior Leadership	General
I10	SL#2	NASA Senior Leadership	General
I11	SL#3	NASA Senior Leadership	General
I12	SL#4	NASA Senior Leadership	General
I13	FM#10	SBIR Funding Manager	General
I14	SL#5	NASA Senior Leadership	General
I15	CSS#4	Scientist on unused case	General
I16	FM#11	SBIR Infusion Manager	General, QWIPs
I17	CSE#8	Staff Engineer	General
I18	CSS#5	Research Scientist	QWIPs
I19	CSS#5	Same as I18	QWIPs
I20	FM#12	SBIR Funding Manager	General
I24	CS#1	Scientist contractor	Polarimeter
I25	CSE#1	Same as I2	QWIPS
I26	N/A	Scientist on unused case	General
I27	N/A	Scientist on unused case	General
I28	N/A	Technologist on unused case	General
I29	N/A	Scientist on unused case	General
I30	CSE#2	Low Temperature Physicist	CADR
I31	CSE#3	Detector Technologist	CZT
I32	N/A	Scientist on unused case	General
I33	N/A	Scientist on unused case	General
I34	N/A	Scientist on unused case	General
I35	M#1	Director of Engineering	Multiple
I36	FM#1	ROSES/SBIR element manager	QWIPs, Multiple
I37	FM#1, 2	ROSES element manager	QWIPs, Multiple
I38	M#2-10	Tech Federation	General
I39	FM#3	ROSES proposal support	Multiple

Interview#	Code	(Relevant) Functional Title	Case
I40	N/A	Scientist on unused case	General
I41	CSS#1	Project Scientist	LDCM(QWIPs)
I42	CSA#2	Same as I1	Microcalorimeters, General
I43	PM#1	Project Manager	Multiple
I44	FM#4	ROSES proposal support	Multiple
I45	CSE#2	Same as I7	CADR
I46	BH#1	Engineering Branch Head	CADR
I47	N/A	Engineer on unused case	General, microcalorimeters
I48	CSE#3	Same as I31	CZT
I49	N/A	Project Scientist	General
I50	N/A	Financial Officer	General
I51	CSA#3	Project Scientist	General, microcalorimeters
I52	CSA#4	LHEA Scientist	CZT
I53	CSA#5	LHEA Scientist	CZT
I54	N/A	Scientist on unused case	General
I55	CSE#4	Detector Technologist	CZT/QWIPs
I56	CSE#4	Same as I55	CZT/QWIPs
I57	PM#2	Project Manager	IXO (CADR/microcalorimeter)
I58	CSE#5	Quality Engineer	CZT
I59	BH#2	Engineering Branch Head	CZT/QWIPs
I60	CSA#6	LHEA Scientist	Polarimeter
I61	CSA#7	Project Scientist	GEMS(Polarimeter)
I62	CS#2	Scientist Contractor	Polarimeter
I63	PM#3	Project Management team	GEMS(Polarimeter)
I64	FM#5	ROSES Proposal Support	Multiple
I65	FM#6	Funding Manager (R&A)	Multiple
I66	FM#7	ROSES topic manager	Multiple
I67	FM#8	ROSES topic manager	Multiple
I68	CSA#8	Project Scientist	Microcalorimeter
I69	CSA#9	Research Scientist	Microcalorimeter
I70	CSE#5	Low Temperature Physicist	CADR
I71	PM#4	Instrument Manager (unused)	General
I72	CSA#10	Research Scientist	Microcalorimeter
I73	CSA#11	Project Scientist	SWIFT(CZT)
I74	CSA#12	Past Director of Space Science, Project Scientist	CZT/Microcalorimeter
I75	PM#2	Same as I57	IXO (CADR/microcalorimeter)
I76	CSE#2	Same as I7	CADR
I77	CS#2	Same as I62	Polarimeter
I78	PM#5	Systems Engineer	Astro(Microcalorimeter/CADR)

Interview#	Code	(Relevant) Functional Title	Case
I79	CSA#13	Project Scientist	Future SMEX(CADR)
I80	CSE#3	Same as I31	CZT
I81	PM#6	Project Manager	SWIFT(CZT)
I82	CSE#4	Same as I55	CZT/QWIPs
I83	CSE#6	Detector Technologist	Microcalorimeter
I84	CSA#8	Same as I68	Microcalorimeter
I85	CSA#11	Same as 150	SWIFT(CZT)
I86	CSE#3	Same as I31	CZT
I87	CSE#7	Detector Technologist	Microcalorimeter
I88	CSA#2	Same as I1	Microcalorimeters, General
I89	CSA#10	Same as I72	Microcalorimeter
I90	CSE#8	Detector Technologist	Microcalorimeter
I91	CSE#7	Same as I87	Microcalorimeter
I92	CSE#1	Same as I2	QWIPs
I93	UL#3	University Professor - Astrophysicist	Microcalorimeter/CADR
I94	CS#3	Scientist contractor	Mircocalorimeter

CADR Document References

Document#	Type	Description
D1	Grant proposal	DDF FY1999
D2	Grant proposal	DDF FY2000 – 1
D3	Grant proposal	DDF FY2000 – 2
D4	Grant proposal	DDF FY2001
D5	Grant proposal	ROSES FY1999
D6	Grant proposal	CETDP FY2000
D7	Funding records	IRAD FY2005
D8	Funding records	IRAD FY2006
D9	Funding records	IRAD FY2010
D10	Grant proposal	IPP 2006
D11	Journal Paper	Cryogenics (2001)
D12	Journal Paper	Cryogenics (2004)
D13	Journal Paper	NIM-A (2006)
D14	Journal Paper	Cryogenics (2010)
D15	Journal Paper	J. of Low Temp. Physics (2007)
D16	Patent	2005: Passive gas-gap heat switch
D17	Journal Paper	Cryocoolers 12 (2002)
D18	SPIE Proceedings	multiple papers in SPIE
D19	Report	ADR "X-Doc" (1980)
D20	Press coverage	Repository of material on Astro missions
D21	Press coverage	IPP newsletter

Document#	Type	Description
D22	Grant solicitation	Con-X NRA 1998
D23	Presentations	Internal presentation (2008, 2009)
D24	Grant proposal	ROSES FY2003

CZT Document References

Document#	Type	Description
D1	Grant Proposal	SR&T 1992
D2	Grant Proposal	SR&T 1996
D3	Grant Proposal	DDF 1994
D4	Grant Proposal	DDF 1995
D5	Grant Proposal	DDF 1996
D6	Report	DDF 1994 year end report
D7	Report	SR&T progress report (1995)
D8	Journal Paper	Solar Physics 118 (1988)
D9	Journal Paper	IEEE Trans. Nucl. Sci. 35 (1988)
D10	Journal Paper	IEEE Trans. Nucl. Sci. 39 (1992)
D11	Journal Paper	Mat. Sci. and Eng. B16 (1993)
D12	Journal Paper	SPIE proceedings (1994)
D13-15	Journal Paper	SPIE proceedings (1995) x 3
D16-21	Journal Paper	SPIE proceedings (1996) x 6
D22	Journal Paper	Nucl. Inst. Meth. Phy. Rev. A (1996)
D23-27	Journal Paper	Materials Research Society, Smiconductors for Room-Temperature Radiation Detector Applications (1997) x 3
D28	Journal Paper	AIP proceedings 1998
D29	Project Proposal	BASIS MidEX 1996
D30	Project Proposal	BASIS SMEX 1997
D31	Project Proposal	SWIFT MidEX 1999
D32	Archive	SBIR archive 84-99
D33	Press	"Great Moments in GRIStory" (2009)
D34	Press	"New Technology With Many Potential Applications Incorporated in NASA's SWIFT Satellite" (2001)
D35	Press	"NASA Selects Mission to Search for Planetary Systems and Observe Cosmic Explosions" (1999)
D36	Brochures	Company tech specs (1995)
D37	Company Histories	Early history of radiation detectors
D38	Personal communications, NDAs etc…	Among the technologists within and outside of NASA, not quoted directly, to protect anonymity

Document#	Type	Description
D39	Journal Paper	Astrophysics and Space Science (1995) (GRBs, not CZT)
D40	Text book	Sold state detector design notes
D41	Archive	History of GRIS: http://lheawww.gsfc.nasa.gov/docs/balloon/New_GRIS_homepage/grisover.html
D42	Archive	Hisotry of InFOCus: http://infocus.gsfc.nasa.gov/
D43	Mission home	SWIFT: http://heasarc.nasa.gov/docs/swift/swiftsc.html
D44	Mission home	NuSTAR: http://www.nustar.caltech.edu/

Microcalorimeter Document References

Document#	Type	Description
D1	Proposal	1983 AXAF NRA (OSSA-3-83)
D2	Proposal	NRA-93-OSS-03
D3	Proposal	New Millenium (1995-1997)
D4	Proposal	MTI proposal 1996
D5	Proposal	NRA-95-17-SZ-040 + evaluation
D6	Proposal	1998 Con-X NRA (2eV Final)
D7	Proposal	CETDP (not funded) 1998-2002
D8	Proposal	CCT Proposal (1999)
D9	Proposal	NRA-99-01-HEA-032
D10	Proposal	ROSES2000
D11	Proposal	CETDP 2000-2001
D12	Proposal	ROSES2001
D13	Proposal	CETDP 2002
D14	Proposal	NRA 2004
D15	Proposal	NRA 2007
D16	Proposal	NRA 2010
D17	Proposal	DDF 2000 "vias"
D18	Proposal	DDF 2001 "trench"
D19	Proposal	IRAD 2002 "interconnects"
D20	Proposal	FY05 IRAD "jfets"
D21	Proposal	FY06 IRAD
D22	Proposal	FY07 IRAD
D23	Proposal	FY07 IRAD
D24	Proposal	FY09 IRAD
D25	History	"Quantum Calorimetry History" - Physics Today (1999)
D26	Press	"US to Participate in Astro-E2"
D27	Press	"IRAD innovators of the year 2008"
D28	Publication	"Bolometer noise: nonequilibrium theory" Applied Optics (1982)

Document#	Type	Description
D29	Publication	"Thermal detectors as x-ray spectrometers", Moseley, S. H., Mather, J. C., McCammon, D., J. App. Phys., 56 (1984) 1257-1262.
D30	Publication	preprint link "Experimental tests of a single-photon calorimeter for x-ray spectroscopy", McCammon, D., Moseley, S. H., Mather, J. C., Mushotzky, R. F., J. App. Phys., 56 (1984) 1263-1266.
D31	Publication	"Thermal Detectors for High Resolution Spectroscopy" IEEE Transactions on Nuclear Science (1986)
D32	Publication	"Advances Toward High Sepctral Resolution Quantum X-ray Calorimetery" IEEE Transactions on Nuclear Science (1988)
D33	Publication	"Development of Microcalorimeters for High Resolution X-ray Spectroscopy" Journal of Low Temperature Physics (1993)
D34	Publication	Calorimeters for very high resolution x-ray spectroscopy" AIP Conf. Proc. 1997
D35	Publication	Toward a 2-eV Mircocalorimeter X-ray spectrometer for Constellation-X (SPIE proceedings 1999)
D36	Publication	"First results from Mo/Au transition-edge sensor X-ray calorimeters" Nuclear Instruments and Methods in Physics Research (2000)
D37	Publication	US Patent # 5,880,468: Superconducting Transition Edge Sensor (1999)
D38	Publication	"A hot-electron microcalorimeter for X-ray detection using a superconducting transition edge sensor with electrothermal feedback" Nuclear Instruments and Methods; Phys. Res.. A (1996)
D39	Publication	"The Science and Technology of Microcalorimeter Arrays" NIM A (2004)
D40	Publication	Other relevant papers listed: http://phonon.gsfc.nasa.gov/Publications/Index.html
D41	Publication	2000s Decadal Survey: "Astronomy and Astrophysics in the New Millenium" (2001)
D42	Publication	2010s Decadal Survey: "New Worlds, New Horizons, In Astronomy and Astrophysics" (2010)
D43	Websites	Chandra (formerly AXAF) http://chandra.harvard.edu/
D44	Websites	Archived details about Astro-E/E2: http://heasarc.gsfc.nasa.gov/docs/suzaku/aehp_about.html or http://www.astronomy.csdb.cn/heasarc/docs/astroe_lc/about_ae2/about_ae2.html
D45	Websites	About Con-X/Xeus/IXO: http://ixo.gsfc.nasa.gov/
D46	Proposal	NRA-96-OSS-07 (TES bolometers)

Polarimeter Document References

Document#	Type	Description
D1	Grant Proposal	DDF 1998
D2	Grant Proposal	DDF 1999
D3	Grant Proposal	DDF 2000
D4	Grant Proposal	DDF 2002
D5	Grant Proposal	DDF 2003
D6	Grant Proposal	DDF 2004
D7	Grant Proposal	DDF 2004-2
D8	Grant Report	Annual report 1999
D9	Grant Report	Annual report 2000
D10	Grant Report	Annual report 2001
D11	Grant Proposal	SR&T1999
D12	Grant Proposal	SR&T2000
D13	Grant Proposal	SR&T 2000-1
D14	Grant Proposal	APRA 2001
D15	Grant Proposal	APRA 2004
D16	Grant Proposal	APRA 2005
D17	Grant Proposal	APRA 2006
D18	Grant Proposal	APRA 2006-1
D19	Grant Proposal	APRA 2006-2
D20	Grant Proposal	APRA 2010
D21	Funding records	FY04 IRAD
D22	Funding records	FY05 IRAD
D23	Funding records	FY06 IRAD
D24	Funding records	FY07 IRAD
D25	Funding records	FY09 IRAD
D26	Funding records	SBIR P1, P2 1999
D27	Presentation	Director's seminar 2005
D28	Presentation	IEEE Talk 2006
D29	Presentation	AXP concept
D30	Journal Paper	Nature 2001
D31	Journal Paper	NIM-A 2002a and b
D32	Journal Paper	IEEE Trans. on Nucl. Sci. 2002
D33	Journal Paper	Proc. SPIE 2000 – 2004
D34	Journal Paper	NIM-A 2003
D35	Journal Paper	J. of Physics Conf Series 2007
D36	Press release	APRA Selections 2006
D37	Press release	SMEX AO 2003
D38	Press release	SMEX AO 2007
D39	Press release	Down select note

Document#	Type	Description
D40	Press release	SMEX awards press release
D41	Press release	AXP selection as category 3

QWIPs Document References

Document#	Type	Description
D1-5	Contract awards	2x EST: ATIP-99-0100; ACT-02-0005; 1x SBIR: SBIR-06-1-S4.02-8429; Lab book notes for DDR and SDI; financial data provided
D6-10	Contract materials	"Quad charts" final presentation for each contract, accessible at http://esto.gsfc.nasa.gov/ and http://sbir.nasa.gov, more details obtained from the personal records of the technologists
D11	Scientific publications	Multiple publications consulted, main source (M Jhabvala, D Reuter, K Choi, C Jhabvala, & M Sundaram, 2009)
D12	Internal presentations	More than 10 internal review meetings re: include TIRs or not (e.g., "Landsat Data Continuity Mission HQ Actions/Issues"(Grigsby, 2009) and "QWIPs-based Thermal Infrared Sensor for the Landsat Data Continuity Mission"(M Jhabvala, D Reuter, K Choi, C Jhabvala, & M Sundaram, 2009))
D13	Websites	LDCM home page: ldcm.nasa.gov
D14-15	Legislative discussion	Debate archived at: http://www.idwr.idaho.gov/GeographicInfo/Landsat/landsat-thermal-band.htm; Landsat remote sensing act of 1992
D16	Press material	Press releases when QWIPs was inducted into the Space Technology Hall of Fame, Successful contract results etc.

Appendix III: Detailed Support for Epoch Identification

This appendix provides a more detailed version of each of tables 4-1:4. The utility of the model specified in thesis is contingent, in part, on a manager's ability to identify epochs in real-time (i.e., as they unfold). In describing the epochs, propositions about were implicitly made on the characteristic behavior of each of the three underlying structural tracks: Funding, Personnel and Technology. The below tables summarize the activities along each track, by epoch and instance, as a basis for corroborating these propositions. Since activities (e.g., funding) do not turn on and off at precise epoch transitions, new activities are highlighted in bold. Carryovers (i.e., activities continuing from the previous epoch) are listed with normal text.

	Case	Funding	Personnel	Technology
Component Exploration	CADR#1	3xDDF +CTD +CETDP	2 +UL -CSE	3 comp heat switch paths
	CZT#2	2XDDF + SBIR P1&2 + RTOPS + SR&T + PoRTIA	2 +SBx2 +CSA + CSEx2	multiple surface deposition techniques
	Pol#3	"free time" DDF + SBIR P1&P2 + SR&T + APRA	3 + IR	trial and error on TFT readout
	Si#4	AXAF-NRA + AXAF	4 +CSAx2 +CSE +St -CSAx3	multiple absorber materials, implantation and thermal isolation tried
	Si#5	DDF + IRAD + CETDP + NRA + XQC + Con-X	4 + CSE	2nd round of experimentation with thermal isolation and a new base material
	Si#6	2xIRAD + NRA + XQC + NeXT +Con-X	6	JFET and automatic attachment R&D
	TES#7	"conference room chats" +2xDDF + Con-X_NRA + NRA + CCT + XQC + Con-X	5+ CSE	Exploration of new materials and electroplating fab techniques

	Case	Funding	Personnel	Technology
Architectural Exploration	CADR#1	**"free time" +DDF**	0 + CSEx2	run simulations of CADR operations; begin building two stage prototype
	CADR#2	CETDP	2 -UL +CSA	Develop first 4-stage CADR prototype
	CZT#3	**"homework"** + RTOPs +SR&T	0 +BH + CSA	Investigate alternative base materials (HgI2, CdTe)
	CZT#4	**DDF** +SBIR P2 + RTOPS + **NDE** + PoRTIA + **NMSC**	8 + CSE	Explore different patterns and array architectures
	CZT#5	**SBIR P1&2** + RTOPS + SR&T + NDE	9 + SB - SB	Revisit different patterns and packaging; also new coarse spatial resolution concept
	Pol#6	**DDF** + SBIR P2 + SR&T + APRA	4	Develop first practical polarimeter
	Pol#7	2xDDF + IRAD + "category 3"	5 + UL	Exploring strategies for suppressing diffusion (e.g., TPC and neg ions)
	Si#8	"hallway discussions" + **DDF**	1 + CSAx2 + UL	Develop first X-ray microcalorimeter in lab
	Si#9	AXAF	4	Explore array architecture alternatives (1D, bilinear)
	TES#10	DDF +**IRAD** + NRA +**CETDP** +XQC +Astro E2 +Con-X	6 +CSE	Explored alternative noise suppression techniques (fingers) and fabrication uniformity

	Case	Funding	Personnel	Technology
Exploitation	CADR#1	**PAPA** + **Con-X**	3	Optimizing design for PAPA application + solving thermal stability questions
	CADR#2	**IRAD** + PIPER + Astro H + **PIXIE**	4	Vibration testing + thermal stability demonstration + relevant applications to missions
	CZT#3	**DDF** + RTOPS + **SR&T** + NDE + **BASIS proposal x 2**	6	Proof of concept for specific application to BASIS design
	CZT#4	**SWIFT** + NDE	Project team - CSE - BH	Procuring and verifying commercial detectors; assembly and test
	Pol#5	**2xDDF + IRAD** +APRA + **AXP("cat 3")**	4 + CSA + CS	Demonstrate robustness of TFT design
	Pol#6	APRA +**GEMS**	Project team – UL	Mature TPC design towards flight
	Si#7	AXAF => **Astro E** + **NRAx2**	4 + CSA (project already around)	Focus on array uniformity for AXAF Spec
	Si#8	**IRADx3** +**NRA** +XQC +NeXT + **Astro-H**	no change	Reuse Astro E2 array for Astro H

	Case	Funding	Personnel	Technology
Treading Water and Branching Out	CADR#1	**IRADx2** + PAPA + **PIPER** +**Con-X peanuts x4** + **IPP seed**	2	core: limited progress branch: developed laboratory CADR (thermal stability and different mating temperatures)
	Pol#2	**IRADx3 +** APRAx2	5	core: continued experimenting with TPC branch: started investigating solar applications and wide area
	Si#3b	**DDFx2** +**NRAx3** +XQC +Astro E	No change	core: focused on developing XRS for AXAF-S or future incarnation branch: started XQC sounding rocket, investigating TES
	TES#4b	**IRADx5 +NRAx3** +XCQ +**Astro E2 and H +NeXT** +Con-X and **IXO**	no change	core: continued improving existing technology branch: started investigating new applications (e.g., solar)