Astro2020 White Paper for State of the Profession Considerations

On

The Next Great Observatories: How Can We Get There?

Corresponding Author:
Jason Tumlinson (Space Telescope Science Institute and Johns Hopkins University)
tumlinson@stsci.edu, 410-338-4553

Co-Authors:
Jonathan Arenberg (Northrop Grumman)
Matt Mountain (Association of Universities for Research in Astronomy)
Lee Feinberg (NASA/GSFC)
John Grunsfeld (NASA/GSFC Emeritus)
Ken Sembach (STScI)
Nancy Levenson (STScI)
John O'Meara (Keck Observatory)
Marc Postman (STScI)

Summary:

Flagship observatories drive astronomical discovery by pushing back the frontiers of capability. To inform Astro2020, NASA has intensively studied four flagship mission concepts spanning the electromagnetic spectrum. Working together, they can find and characterize dozens of Earth-like planets searching for life, peer into the inner reaches of galaxies back to the dawn of time, and see the hottest and coldest parts of galactic gas, stars, and black holes. Yet, recent struggles with flagships teach that the conventional path will not support the simultaneous development and operation of multiple Great Observatories. In an influential and rigorous analysis of flagship development, Bitten et al. (2019) have shown that robust investments in pre-Phase A technology development and design refinements can contain cost growth by closing the design before the final budget is set. We extend the Bitten et al. analysis to show that (1) the conventional development path will not support three or more flagship launches before 2060, if ever, (2) three flagship missions summing to $21B in total cost could launch by 2040 with effective cost-growth containment plus an increase to the total budget of NASA's Astrophysics Division of approximately $1 B per year, as was recently achieved in Planetary Science. If space astrophysics aspires to a new golden age of discovery, Astro2020 must grapple in detail with how the flagship development path should adapt to control costs and support multiple missions.

The Key Issue: Flagship-class space astronomy missions enable historic breakthroughs in our understanding of how the Universe works[1]. The first generation of "Great Observatories" spanned the gamma-ray sky to the mid-infrared and enjoyed a "discovery multiplier" effect by operating together[2]. The discoveries brought by the 15+ year overlap of *Hubble*, *Spitzer*, and *Chandra* inspire us to attempt a new generation of Great Observatories spanning the electromagnetic spectrum. As the NAS "Powering Science" report[1] stated, "It is not possible for NASA to abandon large strategic missions simply because they can be challenging and still maintain world leadership in the space sciences."

Yet, NASA, its industrial partners, and the astronomical community have struggled with cost growth and delays in the development of the most recent flagships, JWST and WFIRST. These issues arise from a combination of technological, budgetary, and human factors, some well understood and some not. Understanding how these forces act on individual missions over the long timescales of development is a research project in its own right, requiring close attention to the decisions made by all stakeholders in their proper historical context. If the astronomical community is to realize its ambitions to search for life on exoplanets, peer back to the beginning of time, and tell the whole story of galaxies and their black holes, we must maintain the ability to advance the cutting edge with flagship-scale capabilities, and doing it across the EM spectrum remains essential. To avoid future mistakes and get flagships back on the right path, it is important for the community to draw the correct lessons about problems of past and current missions.

A recent paper by Bitten, Shinn, & Emmons (2019)[3] has analyzed past and present missions such as Hubble, JWST, and the Mars Science Lab (Curiosity) and identified some of the major factors that drive cost increases and schedule delays. Regarding costs, they conclude:

> *"For the case studies identified in this paper* [e.g. Hubble, Webb, and Mars Science Lab]*, estimating the cost and schedule for these missions would have been extremely challenging. Due to the unprecedented nature of Flagship missions, there is no comparable cost to use as an analogous estimate. The development schedule for technology development is also uncertain, leading to a large uncertainty in the final launch schedule. In addition, early in their lifetime, the design of Flagships is not fully known so it is difficult to assess the cost of a "moving target" as the design evolves. The design trades and options are numerous and indeterminate through early concept development and preliminary designs as technologies mature. The final cost of Flagship missions really cannot be fully baselined until after the technology development is complete and the design has fully matured which is typically after CDR."*

The key lesson here is that costs estimated in the conceptual phases of missions — at the point when they are typically evaluated by the decadal survey — are not reliable unless they are based on essentially a complete and constrained design supported by mid-level technology design, and manufacturing maturity. On the usual path for developing flagships, these demanding conditions are not met until Critical Design Review. This dynamic creates a timing problem in which missions cannot be priced accurately until they are selected and advanced to development, while their relative immaturity at the time of decadal prioritization means that good price estimates cannot be derived, much less weighed into the evaluations. Bitten et al. show that this mismatch lies at the root of many of the issues with flagship development. If the astronomical community, acting through its decadal survey, aspires to

make discoveries with a new generation of Great Observatories, the problems of flagship development must be solved. To address these problems, Bitten et al. propose a new approach (quoted here verbatim):

1. *Conduct a science assessment and concept feasibility study to determine the value of the science and define technology challenges.*
2. *Fund technologies to TRL 6 with defined pass/fail gates for each technology where the phase is open ended with a consistent level of technology funding until technologies pass the required TRL gate.*
3. *Begin an open-ended Phase B to mature the whole system concept to TRL 6 by PDR, include prototyping of manufacturing and test activities.*
4. *Agree to a not-to-exceed annual funding level that continues until a prototype is complete (Step 6).*
5. *After the technology development phase is complete, develop a prototype of the system to work out implementation issues to know the scope of work going forward.*
6. *As prototype development is nearing completion, provide a realistic estimate of the scope of work ahead using CDR as the gate for continuation.*
7. *Get Congressional approval for all remaining development funds which is similar to working capital funds for the U. S. Navy for aircraft carrier procurement.*
8. *Conduct Phase C/D as typical, holding the Systems Integration Review (SIR), Pre-Environmental Review (PER) , Pre-Ship Review (PSR), etc., with lower level peer reviews as needed.*

Bitten et al. use a proprietary Aerospace Corporation budget and schedule analysis tool to model a number of hypothetical profiles for the more "traditional" approach and their proposed alternative. They find that their alternative approach results in higher initial cost estimates but lower total costs in the end because the stiff penalties associated with over-optimistic early assumptions and slow technology development are avoided.

Step 1 is underway now in the form of four concept feasibility studies that define their science yields and technology challenges (LUVOIR[4], HabEx[5], OST[6], and Lynx[7]). Step 1 for these flagships will culminate in Astro2020's "science assessment", which is and would remain the unique mandate of the decadal survey. By Bitten's argument, cost estimates for these missions are premature, except perhaps as loosely defined "cost boxes". If NASA and the community proceed in the traditional way, one of these missions would be ranked first and would get a new start somewhere in the mid-2020s. Experience teaches that this top-ranked project's costs will be uncertain and likely underestimated, so it will struggle to maintain viability against all the technical and fiscal headwinds that have blown before. Meanwhile, the missions ranked second or below may spend the 2020s on technology advancement, but under the current approach they will not proceed even as far as Step 2, where technology development to mid-TRL is funded consistently and with defined gates. Thus the traditional path risks failing in two ways: (1) it repeats past mistakes on the top-ranked mission by proceeding on a highly uncertain technical and budgetary path, and (2) it consigns the others to wait at least a decade before repeating Step 1 without ever getting to Step 2. Given recent experience, this traditional "winner take all" approach seems unlikely to result in a suite of cooperating Great Observatories before 2060, if ever, and even then only if they are designed for multi-decade lifetimes.

What if we followed the Bitten et al. approach, but applied it to multiple flagships in a phased program that was purpose-built to enable the Next Great Observatories?

A return to a multi-flagship, multi-wavelength fleet of space observatories would carry astronomy's aspirations for discovery well into the next decades. Based on the four NASA-led studies alone, an X-ray / UVOIR / Far-IR fleet could search for life on dozens of Earth-like planets in the nearby galaxy, unravel galaxy formation in exquisite detail back to the dawn of time, and probe both the coldest and the hottest places within stars and galaxies. *The science potential of cooperating great observatories is so great that all new avenues for making them a reality must be explored.*

Such a strategic plan would start like our current decadal process for science assessment (Step 1). The decadal would rank-order missions based on their compelling science and overall level of technical feasibility. But after that, a more comprehensive flagship development program would be implemented rather differently from the traditional approach. The key difference is that *the top two or three ranked missions would all proceed to Step 2, with a "consistent level of technology funding" for a robust TRL and design advancement plan.* This plan would define all TRL gates and possible descopes, perhaps with funding weighted by decadal priority. It might even be possible to take multiple missions to the Phase B / prototype stage (Step 4) without a "winner take all" down-select, with the "not to exceed" annual levels again weighted by priority. At some point, of course, a mission must enter full development but this need not zero out the other ranked missions if budgets are arranged properly (see below for hypothetical budgets). This approach has several key advantages:

The new Great Observatories will have better cost and schedule control: This goal is built into the Bitten et al. approach by construction, since the first few steps are meant to advance technology and make the necessary major design trades before cost baselines are established and full funding is allocated.

The new Great Observatories will develop on an accelerated timescale: Properly phased development is essential to simultaneous operation. Unlike the traditional path, high-ranked missions below the top would not be paused, only to have to run the decadal gauntlet from scratch ten or more years later. If necessary they could be evaluated again with several years of serious technology and design development behind them.

The new Great Observatories offer more flexibility in implementation: The traditional approach not only functions as "winner take all", it also does not provide an effective second choice should the top-ranked flagship encounter difficulties. The pursuit of multiple ranked missions would provide a more mature second-place mission, should the top-ranked mission prove too costly, giving NASA more flexibility than it currently possesses to rebalance the astrophysics program while remaining true to the decadal recommendations.

The new Great Observatories will advance their technology faster: It is likely that large tech development investments targeted at specific strategic missions would advance TRLs faster than the ad hoc method currently implemented via SAT and APRA. And, these investments could benefit Explorer and Probe missions on even faster timescales.

How it Might Work: Assuming a community goal of observing the sky from multiple flagships at once, contemporaneous operation implies simultaneous development. Figure 1 shows a range of possible budget scenarios for a set of three flagships. In all cases we have

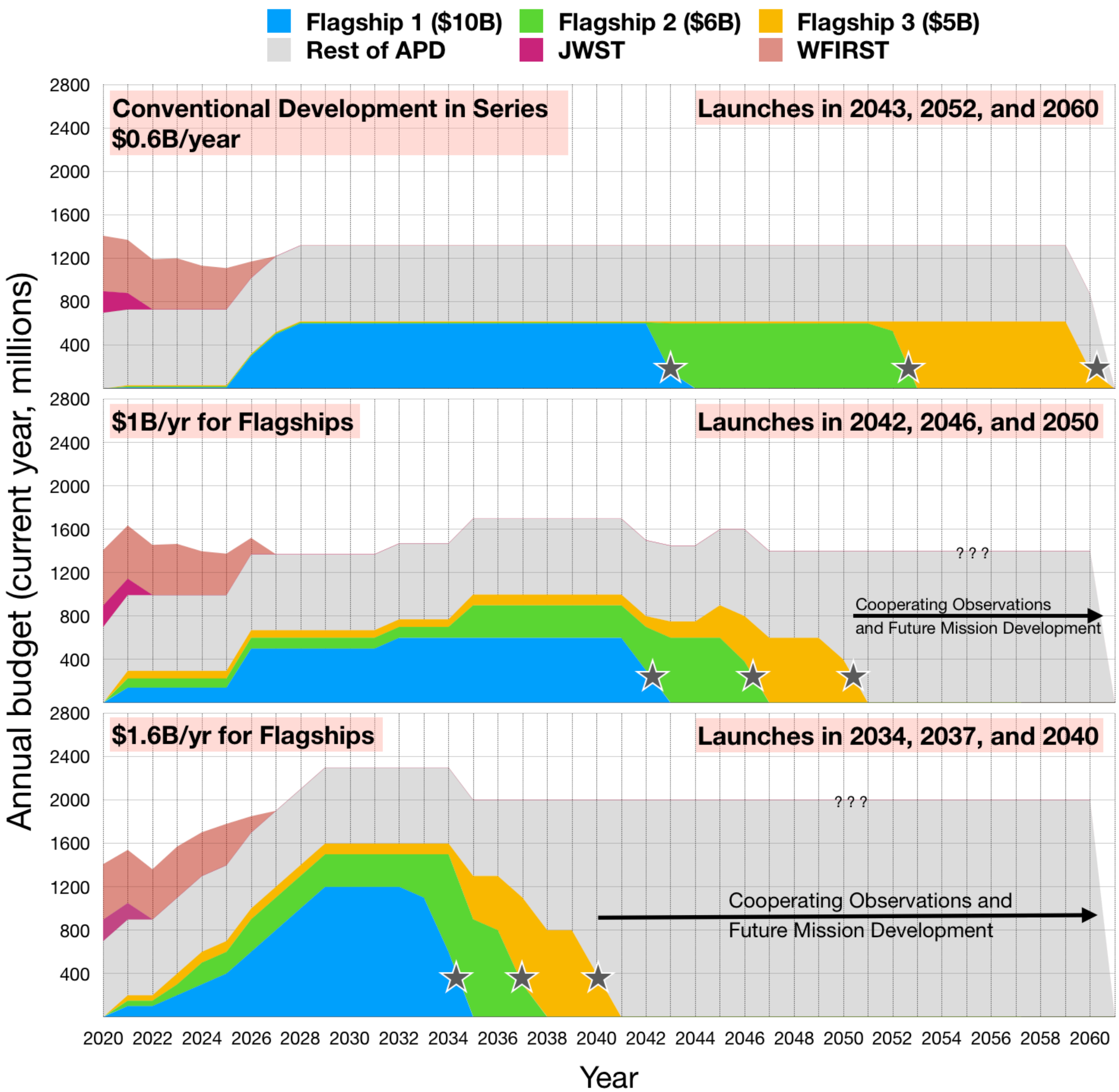

Figure 1: Hypothetical budget profiles for three different scenarios in which three flagship missions are developed and launched. The top panel shows three launches in sequence but fails to yield simultaneous operation before almost 2060, if at all. With \$0.3-1.0B growth in the Astrophysics "top line" for flagships, on the order of recent increases to NASA's planetary science division, the three flagships can launch within a decade of one another and potentially operate together for a long time.

assumed that their total realized costs (Phases A-D, operations not included) are \$10B, \$6B, and \$5B for missions 1, 2, and 3. Before discussing each scenario in detail, we emphasize that, in this exercise, NASA's non-flagship astrophysics program is assumed to continue with its current annual budget of approximately \$700M (grey wedge), leaving the "programmatic balance" of smaller missions and research funding unchanged. The scenarios are constructed as follows:

Development in Series: This scenario models the normal way of doing business by starting a mission roughly every decade, keyed to the decadal recommendations in 2020, 2030, and 2040. For this scenario we have also held the annual total budget in current year dollars fixed at $620M (approximately the current size of the JWST and WFIRST budgets combined) and not allowed for the pre-Phase A conceptual development period recommended by Bitten et al. In this case, launches occur in 2042, 2051, and 2059, spread over almost 20 years. *This is true even if we do not include cost growth and schedule delays, which are extremely likely in the absence of adequate cost control mechanisms.* Consistent with the conventional development path, this scenario does not make a "down payment" early on, so this scenario is subject to much more cost growth risk than the next two. Even though we have held mission costs fixed for this exercise, under these best of circumstances Mission 1 must operate for 20 years or more to have a decent period of overlap with Mission 3, which launches 17 years later. Clearly the conventional development path for flagships will not lead to a new generation of Great Observatories for the foreseeable future.

A Ramp-up of the "Flagship Wedge": The second scenario follows the Bitten et al. recommendation by investing a total of $300M (broken down as, e.g., $140M, $85M, and $75M) in all three flagships immediately following Astro2020, followed by a fully-funded new start for Mission 1 around the time that WFIRST should launch (2025). This scenario also includes a gradual ramp-up of the total "flagship wedge" from $0.6B to $1B for a period in the 2030s, to ensure that all three missions launch within the 2040s and could therefore enjoy a period of simultaneous operation, provided their lifetimes are 10 years or more. This "ramp-up" is assumed to come from an increase to the Astrophysics budget, not from changing the programmatic balance within the current budget. With sufficiently compelling science a $300 to 400M per year increase to enable a new generation of Great Observatories seems plausible. As recommended by the NAS "Powering Science" Strategic Science Mission report[1], "Decadal surveys should be informed by, but not narrowly restricted to, future projections of available budgets. Such flexibility may enable new and potentially revolutionary large strategic missions."

A Faster Ramp-up: Launching three flagships within 10 years would be a great achievement for NASA and would revolutionize astrophysics just as thoroughly as did the first Great Observatories. However, under the slow ramp-up scenario the first launch does not occur until more than 20 years after Astro2020. How much must be added to the Astrophysics budget to launch all three missions by 2040? We find that a total flagship budget of $1.6B, or a ~$1B increase over the current flagship budget, would place the launches clustered together in the late 2030s and would ensure operational overlap. In this scenario the total Astrophysics budget reaches $2.3 B during development of the flagships, still less than the current SMD Planetary Science budget. This would be the best outcome for a community with aspirations to rapidly push back the frontiers across the spectrum.

Between the two versions of the flagship ramp-up are a wide range of possibilities reflecting the dynamic interactions of budgets, technical and design development, and the realizable schedule of these complex missions. There are also caveats to this simple exercise. The shorter (decade-long) formulation cycles contemplated in the accelerated scenario might overtax the development of long-lead components or the workforce, such that shorter timescales are not feasible even with enhanced budgets. The dynamic nature of the Bitten et

al. approach, in which missions must reach a certain point in their design to be accurately priced, might result in a re-ordering of the three missions if they advance at different rates. The actual cost numbers could be different from the nominal values we have chosen. And finally, any multi-mission scenario gets easier to afford if one of them is accomplished through a <$5B contribution to an international partnership.

How to lead our community toward a new generation of Great Observatories is one of the key problems faced by Astro2020. Our scenarios illustrate that the typical way of developing flagships in series, within a pre-determined wedge constrained to its current annual budget, will **not** result in a new generation of Great Observatories. This finding agrees with recent technical lessons-learned studies[8,9] indicating that the problems experienced by recent flagships arise in part from the usual path of development. With increased investment in a coherent and compelling science vision, a serious "down payment" to mitigate risk and lower total costs (as shown by Bitten et al.), and with sufficient will on the part of NASA and its partners, a new generation of Great Observatories is possible.

A Note About Prototypes: Steps 3-5 of the Bitten et al. recommendations contemplate a "prototype" to close the system design from the "design trades and options [that are] numerous and indeterminate through early concept development". Indeed, component and system level demonstrations define the TRL scale at the middle levels.

For a flagship-scale mission, this need not be a working prototype of the entire system. Since the goal is to iterate through the major design and performance trades before closing on a design, prototyping should focus on those key elements that represent the most significant performance challenges or retire the greatest technical risks.

For JWST, this was the wavefront sensing and control need to achieve good image quality on a segmented mirror. For this purpose, the "Testbed Telescope" (TBT; Figure 2) was built at 1:6.5 scale (with a 1 m aperture) to demonstrate WFS&C, develop control algorithms, and validate models in all the relevant degrees of freedom[10].

The *Chandra* project employed testbed articles called the "Verification Engineering Test Articles" (VETA; Figure 3). VETA-I addressed high-risk technologies in the optics and assisted with development of the test equipment. In the words of Arenberg et al.'s lessons-learned SPIE paper[11]: "The VETA-II effort took the largest two mirrors and aligned and mounted them to the flight tolerances using the flight handling hardware and alignment metrology. Hundreds of lessons learned were documented as a result of the VETA-II effort, resulting in changes to handling procedures, support equipment, metrology hardware, and alignment and metrology processes. This enabled the installation of the flight mirrors to go very smoothly, and resulted in the telescope being completed 1 month ahead of schedule. The VETA-II is the very definition of early testing and good risk mitigation."

For the LUVOIR concept, the technological "tall pole" is the the mechanical, optical, and thermal stability of a system that must support high-contrast (10^{-10}) nulling of stellar light to detect and characterize 30th magnitude earth-like planets[4]. Here the prototyping would include both telescope and coronagraph elements in an ultra-stable vacuum chamber. This testbed would evaluate different metrology options (such as edge sensors vs. active optical monitoring). It would also explore the different coronagraphic mask technologies (e.g. VVC, PIAA, or APLC) and how they match to the telescope as an integrated system. With proper

Figure 2: The Wavefront Sensing and Control Testbed Telescope (TBT) for JWST. This testbed demonstrated TRL-6 for JWST's primary mirror Wavefront Sensing and Control[10]. It included a back end instrument simulator tested wavefront sensing over the full field of view of the telescope, validated models, and evaluated control algorithms. The TBT was built to 1:6.5 scale but demonstrated all the degrees of freedom and final performance requirements.

Figure 3: The VETA-II test article for the Chandra X-ray observatory[11]. This "Verification Engineering Test Article" enabled alignment of the main mirror elements and uncovered major design risks.

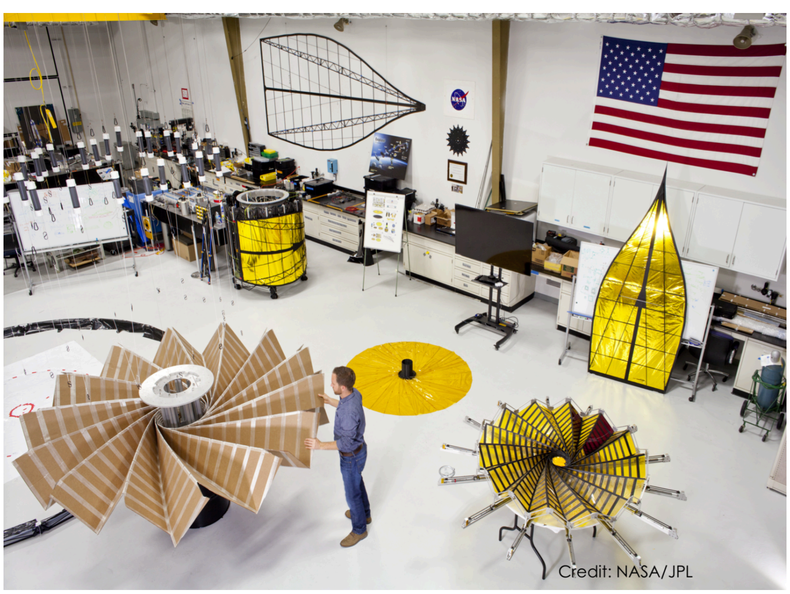

Figure 4: Starshade prototypes in the lab at JPL. Implementing a new approach to flagship development could prove this technology in the relevant space environment during the prototype phase, before a new start that commits full funding even with remaining open questions about system design and major performance tradeoffs.

investments, such a prototype system could reach TRL 6 for LUVOIR's critical technology and system design issue in a single demonstration.

If NASA's exoplanet hunting flagship needs a starshade, as in the HabEx concept[5], there must be more investment in the deployment and control of starshade technology. Considerable effort has been invested in lab-scale prototypes (Figure 4), but so far no in-flight demonstration of starshade technology has occurred. For Lynx and Origins Space Telescope this prototyping stage might encompass X-ray mirror fabrication and test and thermal controls for operating an entire telescope at 4 K. Under our proposed path for multi-flagship development, all these technology and prototype developments could occur together if recommended by the decadal and properly funded.

Possible Avenues for Further Investigation: We have performed the thought experiment of applying the Bitten et al. approach to flagships in support of a new generation of Great Observatories. Beyond the limits of this simple exercise, there are many open questions and issues that could be fruitful avenues for further study:

How could the possible budget scenarios map to the specific concepts NASA has studied? With the proper tools, it might be possible to work out a plan for these concepts through the first few steps of the Bitten et al. process within the context of Astro2020.

How could the program of Great Observatories be constructed — both scientifically and programmatically — to provide the most compelling possible case to Congress, such that the increased budgets can be achieved?

To what extent can multiple flagships save money and reduce risk using common components, such as spacecraft busses, power / control / communications hardware, and even ground based manufacturing and test equipment? To what extent can engineering and tests already done for JWST and WFIRST be reused?

How might servicing observatories (with humans, robots, or both) lengthen lifetimes, provide increases in capability, and offer risk reduction for a fleet of flagships?

How might future heavy launch capabilities reduce risk (and therefore cost) by introducing robust mass and volume budgets into the flagship equation?

Final Thoughts: We started by asking, What if we followed the Bitten et al. approach, but applied it to multiple flagships in a phased program that was purpose-built to enable the Next Great Observatories? In answer, we find that:

1) The conventional development path in which flagships are developed in sequence and subject to typical cost growth and schedule slip will not yield three flagship launches before 2060, if at all.

2) To launch three flagships within a decade, NASA must receive increases to the total Astrophysics budget totaling at least \$0.3B per year for flagship development. An increase of ~\$1B per year, commensurate with recent growth in SMD's Planetary Science Division, could get all three flagships to the launch pad by 2040 with a significant period of operational overlap.

3) *Both cost control and larger budgets are necessary to achieve a new generation of flagships.* With flat budgets, even perfect cost control will fail to pull three multi-billion dollar missions

close enough to overlap in their operations. In the absence of cost control, unplanned growth will unnecessarily waste portions of any budget increases that are achieved and thereby delay the later missions in the suite.

Our final conclusion is that adopting the Bitten, Shinn, and Emmons (2019) alternative approach for multiple flagships in parallel, with assigned decadal science priorities and serious technology developments after Astro2020, is likely to result in a generation of Next Great Observatories faster than otherwise.

References:

(1) National Academies of Sciences, Engineering, and Medicine. 2017. Powering Science: NASA's Large Strategic Science Missions. Washington, DC: The National Academies Press. https://doi.org/10.17226/24857

(2) Megeath, S. T., et al. "The Legacy of the Great Observatories: Panchromatic Coverage as a Strategic Goal for NASA Astrophysics" Astro2020 APC White Paper (see related NASA COPAG study also by Megeath et al.)

(3) Bitten, R. E., Shinn, S. A., & Emmons, D. L., 2018, IEEE Proceedings "Challenges and Potential Solutions to Develop and Fund NASA Flagship Missions" [link]

(4) The LUVOIR Mission Concept Study Interim Report, LUVOIR STDT, https://arxiv.org/abs/1809.09668

(5) The HabEx Mission Concept Study Interim Report, HabEx STDT, https://arxiv.org/abs/1809.09674

(6) The Origins Space Telescope Concept Mission Study Interim Report, OST STDT, https://arxiv.org/abs/1809.09702

(7) The Lynx Mission Concept Study Interim Report, Lynx STDT, https://arxiv.org/abs/1809.09642

(8) Arenberg, J., Atkinson, C., & Conti, A., 2016, "Ensuring the Enduring Viability of the Space Science Enterprise: New Questions, New Thinking, New Paradigms" Proc. SPIE 9904, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave, 99041F, doi: 10.1117/12.2234492

(9) Feinberg, L., Arenberg, J., Yanatsis, D., & Lightsey, P. A., 2018, "Breaking the Cost Curve: Applying Lessons Learned from the James Webb Space Telescope Development", Proc. SPIE 10698, Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave, 1069823, doi: 10.1117/12.2309661

(10) Feinberg, L. D., Keski-Kuha, R., Atkinson, C., & Texter, S.. C., 2010, SPIE Proceedings 7731 "Use of a pathfinder optical telescope element for James Webb Space Telescope risk mitigation" https://doi.org/10.1117/12.2309664

(11) Arenberg, J., et al. 2014, Proc. SPIE 9144, Space Telescopes and Instrumentation: Ultraviolet to Gamma Ray, 91440Q https://doi.org/10.1117/12.2055515