IDA Ideas

A podcast by the Institute for Defense Analyses

Episode #15

Addressing Challenges in Weapons Sustainment

Guests: V. Bram Lillard and Benjamin A. Ashwell

Host: Rhett A. Moeller

February 2024

Product 3001363

Institute for Defense Analyses
730 East Glebe Road
Alexandria, VA 22305

About This Publication

The views, opinions, and findings should not be construed as representing the official positions of the Department of Defense or the U.S. Government.

For More Information

V. Bram Lillard, Director, Operational Evaluation Group
vlillard@ida.org, (703) 845-2230

Copyright Notice

Addressing Challenges in Weapons Sustainment

IDA Ideas Host Rhett Moeller spoke to V. Bram Lillard and Benjamin Ashwell. Bram is the Director of the Operational Evaluation Division (OED) in the Systems and Analyses Center, an IDA-operated federally funded research and development center. Benjamin is a researcher in OED.

[Begin transcript]

Rhett Moeller: Hello listeners. I'm Rhett Moeller and I'm the host of IDA Ideas, a podcast hosted by the Institute for Defense Analyses. You can find out more about us at www.ida.org. Welcome to another episode of IDA Ideas.

Today, we're going to talk about weapons systems readiness and sustainment. This, of course, is a topic of great importance to the defense community, and one that IDA is particularly well poised to address through its vast history of work on the topic. For today's discussion, I'm thrilled to welcome two of our experts into the studio. Bram Lillard and Benjamin Ashwell. Both work in our Operational Evaluation Division or OED. Benjamin is one of our knowledgeable researchers focused in this area, and Bram is the director of the division and also quite steeped in system sustainment. Can you take a moment to introduce yourselves?

Bram Lillard: Sure, thanks Rhett. Again, I'm Bram Lillard. I'm the Director of the Operational Evaluation Division here at IDA and I've been at IDA actually, almost 20 years now. It has been a fantastic place to work and to pursue topics helping our men and women in uniform. One of those many topics that the Operational Evaluation Division is passionate about is today's topic, weapon system sustainment and readiness. And I'm excited to talk about how we are helping the Department of Defense [DOD].

Rhett: Great. Welcome to the show. Benjamin?

Benjamin Ashwell: Hi, I'm Benjamin Ashwell. I'm the lead for the sustainment analysis group in OED. I've been at IDA for eight years, half of that on naval surface warfare analysis and the last four years on weapon system sustainment analysis.

Rhett: Great. Also, welcome. We have a lot of experience here and I am excited to get going. So, as we considered questions to bring into this conversation, one of the first that came to mind is let's start by defining what a weapon system is.

Bram: Sure. So, when we refer to weapon systems in the context of the Department of Defense, we're referring to the equipment that men and women in uniform rely on should we ever have to be engaged in a conflict or war. This equipment can be everything from the hand-held radios that soldiers use all the way up to fighter jets, carriers, destroyers, ships, tanks, trucks, as well as computer systems.

Rhett: So, with that, what is readiness when it comes to weapon systems and what is sustainment and why are they so important?

Bram: So, readiness is a term that is very well known in the Department of Defense and is one of the highest priorities of the department's leadership. It refers to the ability of a unit, which can be a platoon of soldiers, a ship, a squadron of aircraft, to be able to have everything they need, be that equipment, manning, training, so that they can do the mission should they be called upon to do it. And there's a lot that goes into readiness. We are today going to be talking primarily about the equipment aspect of it. Sustainment is typically a term that is very similarly related where over time these weapon systems need to continue operating and they need to be available when they're needed. And what it takes to do that is an art and a science. And there's a tremendous amount of data that is necessary to sustain these weapon systems so that they are available to the units that need them both in peacetime and in wartime.

Benjamin: Sustainment is the way you achieve and maintain readiness, and readiness is the goal. A lot of attention and focus in the DOD and in the wider defense media is focused on big budget new projects, the newest exciting things, and that's very important. New technology development is essential to maintaining our nation's defense. But a tremendous amount of the budget goes into making sure the stuff we already have — tanks, trucks, networks, IT [information technology] infrastructure — making sure those things continue to work year after year, decade after decade. More money is spent on sustainment every year than is spent on either R&D [research and development] or new acquisition. So, while it maybe doesn't get the attention it might deserve in the press, it certainly does get the lion's share of the budget. And therefore, it's very, very important that we understand how to spend this money efficiently to make sure we are getting the most bang for our buck.

Rhett: Given that, what kinds of analysis does the Department of Defense want to be able to do?

Bram: So, one of the biggest challenges that the Department of Defense faces is the ability to tie cause and effect together. If I spend $10 on these spares or this maintenance or helping improve the training quality or the logistics, then how much does that actually give me in terms of improving readiness? There's a variety of metrics that the department keeps track of and tries to improve through investments and resource decisions, but tying those outcomes to what goes into it and the budgeting process is not obvious. So, there is a deep need in the Department of Defense for analysis that can help tie those two things together.

Benjamin: I often liken the metrics the DOD uses to evaluate how well it's supporting readiness as blood pressure, having a bad blood pressure is bad. If these metrics are low, that's bad. But having a good blood pressure does not mean you're healthy. These intermediate metrics are useful, but they don't directly relate to the, no kidding, take flight-line readiness, the number of Super Hornets that can take off and do their mission today. And that link between resource decisions, money, people, time and flight-line readiness, pier-side readiness, the ability of the systems to do their missions is often lacking because people focus on the intermediate metrics.

Bram: The motto that we often use in the sustainment analysis group here at IDA is that we are trying to tie resource decisions to readiness outcomes, resources to readiness. This is not easy to do. There is a tremendous amount of data, which is dirty data or very difficult to interpret data, but we have found a tremendous amount of recent success in building an end-to-end picture. So, for example, let's take a fleet of aircraft. You can imagine there's hundreds of aircraft that are positioned all over the world at various U.S. bases and beyond to ensure that those aircraft can fly when we need them, takes everything from the maintainers, the spare parts, the logistics train to get the material to where it needs to be, as well as the supply chains to ensure that that equipment can be available when it is needed. There is a tremendous amount of financial analysis that needs to be done to optimize these decisions. And we think we are being successful by building models and simulations that connect all these dots together.

Benjamin: We use a variety of tools to try to understand the sustainment system and its interrelationships. One of the most powerful is a discrete event simulation, which is kind of fancy language for a computerized board game. Imagine a really big board game, you have 600 Super Hornet fighter jets all over the world as little game pieces. And each game piece if you zoomed in has 1,000 different subcomponents, generators, radars, landing gear — we have each of those and each of those has its own rules in some massive rulebook. This part fails this often, this part fails this often. If this fails, you can't do any air-to-ground missions. If this one fails, you can't do air-to-air refueling.

Bram: Exactly.

Benjamin: And then we put little Monopoly pieces all over the board for the spare parts and things break and bad parts get pulled off the aircraft, new parts get put on. And there's even more layers because we're trying to model the way the sustainment community is managing the procurement and repair of spares by putting all of this into a big board game, pressing go, and having it run and roll dice. In a stochastic random way, we can see how all these pieces interact because we're kind of modeling the physics if you like of the sustainment system. So, we can see how, if you perturb one spot, you increase the time it takes to move parts from the West Coast to Japan, you change the failure rate of a component. You can tie those upstream effects to the number of aircraft on [an] aircraft carrier that are capable of performing their mission. And it's only through that physics of the sustainment system simulation that you can see cause and effect across the entire sustainment system.

Bram: I'd like to add one more thing, the leaders across our military services, the folks that have to make these resource decisions. There is no lack of a good idea on what could be done. So, the output metric here we're talking about, that Benjamin has mentioned here, is that we're trying to increase the number of available airplanes to fly, the number of available ships, the time that the radar is functioning, the time that the network is able to perform its mission. There is no lack of ideas on how we might improve those measures. We could buy more spare parts, we could change the maintenance structure or the number of personnel or their specialties. We could stand up a new depot repair facility. We could improve the reliability of certain critical components on the aircraft or the ship. There are numerous ways to invest in improvements. But what is often missing in the decision-making process in the budgeting process across all of the military services is that if I put, again, $10 towards this, $1,000 towards that, what am I [going to] get from it? And is there a return on investment for these many different investment options and can I rack and stack them in a way to make data-informed decisions and make the best decision for the investments that the department is making?

Benjamin: One reason this is difficult is because there's rarely a single culprit behind poor readiness.

Rhett: Sure, sure.

Benjamin: It'd be nice if it's just, if we fix the generator, the aircraft's [going to] be great. Depending on the weapon system, there may be a couple of what we call top degraders, things causing real headaches. But aside from those, it's death by 1,000 cuts. Typically, many DOD organizations like to tackle the top 10 or top 20. And that's fine, but that rarely gets you the most efficient outcome because you need to be tackling the top 100, top 200 and spreading your resources in a way that achieves an optimal outcome, because again, there are many, many problems that are interconnected. So, you have to have a systems view of how changes to what you're doing influence the readiness outcomes.

Bram: So ultimately, we need these capabilities, these analytical approaches, like the models we just talked about, to be able to answer a wide variety of different types of questions, everything from the tactical to the strategic. So, a tactical question might be, as Benjamin has previously mentioned, how many spare parts do I need to buy? How many additional generators do I need to procure in order to ensure this squadron of aircraft are more available and ready to fly when they are needed? Alternatively, we might need to ask and answer very strategic questions, like should I build an entire new maintenance facility at a particular base or change the entire maintenance construct for the fleet of ships? That breadth of different complexity of questions are the challenges that we are taking on with these approaches, and building an analytical construct and a modeling baseline that enable us to go from the tactical to strategic is something that we have become very passionate about and we are contributing to.

Benjamin: Yeah, this tactical to strategic range, as Bram said, on the one end, you're thinking about individual purchasing decisions, this component versus that component right now. The same models that are helping you figure that out can answer really big picture, strategic questions like what would happen to the supply system to our stockpile of spares if we ended up in a conflict where we suddenly started operating 50% more and we started changing the kind of missions we flew and the resupply times were longer because of enemy interdiction. You can start to see how those big picture questions are supported by the same model, the same physics of the supply system work across the board. And so, it's a powerful flexible tool.

Bram: It helps with answering those "what if" questions. What if this scenario happens? What if I run out of that part? What if I need to preposition equipment? How do I handle that? And what are the consequences if I don't do anything?

Rhett: That's obviously a complex problem with a need for a very comprehensive solution. And certainly, the things that you have talked about already show a great start toward that. So, what challenges does the Department of Defense face?

Bram: So, there are many challenges. I'll say a few and then Benjamin, I think you have a couple of ideas as well. But the first one that comes to mind is that sustainment is hard, number one, and it's the timelines [that] can be extremely long. The problem with these timelines is that you essentially need to know now what you need three years from now, and we need to be making decisions today which will affect the readiness of our systems in the future. Unfortunately, or maybe understandably, the DOD is typically focused on what can be done immediately. And many of the processes that we're talking about that we're modeling are on much longer timelines that require analysis today for future outcomes.

Benjamin: It makes the cause-and-effect piece of this difficult because the flash-to-bang, you know, the lightning-to-thunder delay between when you make the decision and when you see the outcome is long, it might be longer than an officer spends in a certain position at a certain command. The person who made the decision is gone by the time its effects are seen. Additionally, beyond just things taking a frustratingly long time, the scale of the problem swamps people. There are hundreds of thousands, maybe millions, of unique spare parts that different agencies are managing: radar circuit cards, landing gear, showerheads, ice makers — a staggering range of different things. There is no way any one or small group of humans can get their minds around the entire trade space — I will if I buy five fewer showerheads and one more bomb rack — what does that do? The scale of it defeats people. People end up focusing on solving local problems. I'm making a small trade-off between landing gear and tail hooks. But because the problem is so big, people aren't asking that question — should I be funding submarines instead? — in the first place. The trace base is so vast. So, we've got hundreds of thousands of components, varying prices, lead times, a qualitative difference in the importance of a component. How much is a landing gear worth versus a bomb rack? Which one's more important is a value judgment.

Bram: So, it's just all of this together. So, one example of this is that we recently completed an analysis of the V-22 Osprey aircraft. There are actually multiple fleets of V-22s. Many are operated by the Marines, but then also many are operated by the Air Force for different missions. And then most recently, the Navy has begun adopting a different variant of the same aircraft. Those three fleets all share a common supply system. And so, making decisions about one, a set of components for a particular squadron of aircraft, could have reverberative effects across multiple other fleets as well as military services. If you do not have a single, common, end-to-end picture of how that all fits together, then you could be making a decision that has an outsized effect on other parts of the military without even knowing it.

Benjamin: For instance, the timelines of these processes are so long that the people who made the decision to buy something, repair something, begin a process improvement or reengineering, those people might not be in the same position when the process is complete. So, you're an officer, you got assigned to say NAVAIR [Naval Air Systems Command] and you oversaw a plan to improve depot repair times. Two years later, you were promoted and moved to new posting. By the time the depot improvement process is complete, it might be your replacement or your replacement's replacement who is there trying to defend decisions made years ago.

Bram: So, in addition to the two problems, we've already talked about, timelines and the scale of the problem, the other major challenge is data. And in order to do this analysis and make very good decisions, we have to have good data and interpretable data and reproducible analysis that enables us to make better decisions and do the modeling. Right now, it is not always clean and easy. Let me give a couple of examples. Number one, the databases that the DOD maintains are often disconnected from one another. Secondly, they are often interpreted differently by different communities, so much so that they are incompatible with one another. So, bringing sort of order to chaos is one of the things that we're doing here at IDA.

Furthermore, just because they exist in databases, these data, does not mean that the data themselves actually reflect reality. Let me give you an example. We have many of our service members who spend their entire careers performing maintenance on these weapon systems and they are extraordinarily hard-working and knowledgeable people, but they are under a lot of stresses as well. And so even tasks that can be either simple or complicated, in order to gain the data that we need for some of these analyses, the data entry must be high quality and sometimes the systems that these maintainers work with are not always easy to use or complicated to fill in all the pieces of information that we might desire for some of these analyses.

Benjamin: So, as Bram mentioned, some of these problems are driven by problems with the data systems or the data collection process or the fidelity of the data. But there are also some fundamental problems dealing with small sample size. If you think about, say, my favorite example, the Navy Super Hornet, there are 600 of those, approximately, and by DOD standards, that's a very large fleet of aircraft, great sample size, but compared to civilian equipment, cars, iPhones, TVs, it's miniscule. So, when you ask a question that you think there should be an authoritative answer to, what is the failure rate of the generator? That's a complicated question. First of all, your sample size isn't that big. Second of all, what do you mean by failure? How about replacement because it was old? Is that a failure? Does that count? Also, these failure rates change over time: as aircraft age, failure rates change; as missions change, failure rates change. So certain kind of missions may stress the system and make it break, and others might not. And then you go to cases that are just fundamentally ambiguous. There are three Seawolf-class submarines in service, and there will only ever be three, n equals three. How do you get the statistical failure rate of a component from there? Oh, this strut, it's never failed so far. Does that mean it'll never fail again? Just the small sample size and these questions about what is failure about operational tempo assumptions, these are challenges you face when trying to come up with key metrics like failure rate. That sounds simple, but in fact, [it is] complex and it is critical that the assumptions you're making are clear, they're documented, and that they are consistent across different analyses so that we have apples-to-apples comparison across the fleet.

One of the things that's key for us at IDA to be able to do this is the long-term relationship we've built with our sponsors in the Navy and the Air Force, other services, the people in and out of uniform, who know what they're talking about, the people who fill out these forms, the people who actually go out there and fix aircraft. I don't know what I don't know about aircraft. I'm pretty decent at looking at data and working with data, but to translate the data into what actually matters, what actually counts, that relationship with the subject matter experts is key. Without our collaborators in the Navy [and] the Air Force, other services, we'd be lost. It's that marriage of expertise with data and modeling and abstraction and the experts who are the engineers, the maintainers who understand, again, what actually matters in real life. That's kind of the secret sauce that makes this possible, that makes this applicable, that makes it relevant.

Rhett: So, you've talked about what Department of Defense wants to do. You've talked about the challenges involved. How is IDA helping?

Bram: That's a great question. I view our role is to bring some order to an extremely complex situation and really helping the Department of Defense understand how best to make optimal decisions. We've talked about data, we've talked about the challenges, one of the ways that we think we are aiding our partners in the Department of Defense is to create data pipelines. In order to do that, we have to take all of these different disparate data sources, as well as the folks who work with them, and bring them together into a process to understand what each of the individual data pieces mean, how do we appropriately interpret them, and then how do we create a repeatable process to create all these inputs for these end-to-end sustainment models?

Benjamin: So, Bram mentioned a pipeline, and what that means in practice is a process where we start with raw data. And by raw data, I mean we get all the maintenance records from naval aviation for the last five years. We get all the requisitions to support those systems. We get stock levels, contracting actions, flight hour records, we get all of the ground truth for what happened and is happening right now, we put it in one place, we curate it, and then we commune with it. We figure out what does this column mean? When this column is an R and that column is an 801, does that mean something important to us? Does that change our interpretation of what happened? Because we need to be generating a ground truth for modeling, an internally consistent truth that is relevant for the kinds of decisions we're helping the DOD make. So, we get the raw records and then we run them through code.

We've written a software package that takes all the records, applies filters — oh, these records don't count because of this column; oh, this sounds like a critical failure and this one isn't — and condense it until we have something much more manageable for millions of records. We now have just maybe thousands of rows, one row per item per aircraft type. This is how often the generator fails on an MH-60 Romeo helicopter. And this is how often it fails on an MH-60 Sierra helicopter and it's different. This is how bad it is if the item fails in flight, this is how much it costs. This is how long it takes to get a new one.

Now, we've got the metrics, the parameters that define the essential qualities of each item. But we did it in a repeatable way, where, for example, almost all naval aviation platforms share the same databases. So, we can apply the same logic and filters for almost every platform in naval aviation. It's consistent, the models make the same assumptions. And when we get new data, you know, we get the November data, we can throw it into the hopper and turn the crank. And within an hour have updated metrics for every type of aircraft remodeling.

Rhett: Bram?

Bram: So, one unique aspect to how we're contributing is that we are able to do all of these analyses and share them with our government sponsors and ensure that the analysis methods, the calculation algorithms, and everything that Benjamin just described is transparent and can be used by others. And if we see improved readiness from the F-35 or the F-18 or an Aegis destroyer, then we have succeeded.

Benjamin: Another way we're improving this process is making sure the existing metrics and inputs get better and better. That the inputs, say, failure rate or the lead time to manufacture something, that the method we use to calculate that takes advantage of the latest and greatest in terms of statistical analysis, machine learning, AI, to try to make sure that we always have the most predictive, most accurate numbers flowing into the decision making. Bram mentioned that we're sharing our tools and methods with the services. Enduring improvements require the services to see these tools as their own. As long as we're on the outside, turning the crank, it can be helpful, it can be useful, but real change will happen as the services internalize these tools, adopt them and begin running them on their own as it becomes part of "the process," the official way they think about sustainment becomes written into the procedures. That's how we scale up the research of a dozen or two dozen analysts into a DOD-wide change.

Rhett: Yeah, that makes sense. So, with all of this said, what do you see as being the next thing? What's the next objective?

Benjamin: On a technical level we are always trying to improve the fidelity of the models, the accuracy, trying to capture more nuance. One area we're focusing on right now is trying to understand how to match up the maintenance data with the manpower data to understand the bottlenecks, because you might have all the spares you need on the aircraft carrier and you might have a lot of maintainers on the aircraft carrier, but we might still see maintenance backlogs. Why? Well, different maintainers have different specializations and capabilities and different priorities. And we don't necessarily know what they're doing throughout the day. How they're being pulled in a bunch of different directions at once. Trying to capture that linkage will allow us to understand the bottlenecks and again, tie cause and effect. What would happen if you added more ordinance experts to the maintenance team? Would that streamline a process or is that not the bottleneck to begin with? Additionally, there are special classes of components managed by different organizations. Every month, we're getting smarter on all the exceptions and special cases. And we're trying to incorporate this into the model to make sure that the answers have the highest fidelity we can achieve.

Bram: So maybe just to take it up a level to the 50,000-foot view, one of the most important next steps that needs to happen is that the leadership across the military services, Office of the Secretary of Defense, and so on really need to continue to emphasize and require these kinds of data-driven analytical approaches. It is only through these rigorous approaches, these modeling efforts that we've mentioned today that readiness can actually improve across the department. Benjamin mentioned earlier that we view the secret sauce to all of this is pairing the experts who wear the uniform, who do this on a day-to-day basis, with the data analysts here, and that has to be encouraged and grown. We also need to have long-term support for these efforts in the Department of Defense. It takes years to build up the competence, the knowledge base, the data pipeline, continuing to improve the quality of the data, as well as the models, and those commitments can ensure that readiness overall is going to improve, and ultimately our national security will be enhanced.

Rhett: Well, gentlemen, thank you very much for taking the time to discuss this timely topic with us and for providing us with a framework to understand both the importance and the challenges involved. As somebody who has been on the receiving end of military sustainment and concerned about readiness, this is a topic of interest to me, too. Thank you for your work and helping to support our service members.

Bram: Thank you. We really enjoyed it.

Benjamin: Yes, thanks.

Rhett: This has been most illuminating for me, and I hope it has for you too, listeners. As always, if you want more information on IDA and its ongoing work, please do check us out at ida.org. We also have a presence on [X, formerly] Twitter, at IDA_org, and we have a channel on YouTube. The show is hosted by the Institute for Defense Analyses, a nonprofit organization based in the Washington, D.C., area. Once more, you can find out more about us and the work we do at ida.org. Thanks for tuning in, and we hope you'll join us again next time as we discuss another big idea here at IDA Ideas.

Show Notes

Learn more about the topics discussed in this episode via the links below.

Lillard, V. Bram, and Han G. Yi. "Building End-to-End Sustainment Models for Weapon Systems." IDA Document NS D-22672. June 2021. https://idalink.org/d-22672.

Lillard, V. Bram, Benjamin A. Ashwell, and Edward J. Beall. "Tools for Building End-to-End Readiness Models with OPUS/SIMLOX." IDA Document NS D-13183. April 2020. https://idalink.org/d-13183.

Ashwell, Benjamin A., V. Bram Lillard, and Megan L. Gelsinger. "DATAWorks 2022: An Introduction to Sustainment: The Importance and Challenges of Analyzing System Readiness." IDA Document NS D-32961. March 2022. https://testscience.org/wp-content/uploads/formidable/13/D-32961-Lillard-et-al.pdf.

Ashwell, Benjamin A., Andrew C. Flack, and Han G. Yi. "M&S Approach for Quantifying Readiness Impact of Sustainment Investment Scenarios." IDA Document NS D-32983. April 2022. https://testscience.org/wp-content/uploads/formidable/13/D-32983-Flack-et-al.pdf.

Ashwell, Benjamin A. "DATAWorks 2022: From Gripe to Flight: Building an End-to-End Picture of DOD Sustainment." IDA Document NS D-32988. April 2022. https://testscience.org/wp-content/uploads/formidable/13/D-32988-Ashwell.pdf.

Lillard, V. Bram, Kyle E. Remley, and Joseph M. Fabritius. "DATAWorks 2022: Taming the Beast: Making Questions About the Supply System Tractable by Quantifying Risk." IDA Document D-32982. April 2022. https://testscience.org/wp-content/uploads/formidable/13/D-32982-Remley-et-al.pdf.