Department of the Navy
Small Business Innovation Research (SBIR)

Guidebook to Experiment Analysis & Reporting

Guidance for Small Business Innovators on Using Analysis and Planning to Achieve Experimentation Objectives and Create Reports

Version 1

July 2022

Record of Changes

Revision Number	Date Published	Summary of Changes
1	7/28/2022	New Version

Contents

1 Introduction 4
1.1 Purpose & Scope 4
2 Why Experiment? 4
2.1 Experiments & Demonstrations 5
2.2 Experiments and Test & Evaluation 5
3 Three Step Process 6
4 What is Important? 8
4.1 Understanding Capabilities 9
4.2 Problem Statement 10
5 Creating Objectives 10
6 Developing a Data Collection and Analysis Plan 12
7 Collecting the Data/Execution 14
8 Analyzing Results 15
9 Reporting 15
10 Closing 16
11 Appendices 17
11.1 Data Collection and Analysis Plan Excerpt 17
11.2 Reporting Templates 21
12 References & Further Reading 32
13 Glossary 33
14 Acronyms 34

1 Introduction

Experimentation is a key process used by naval acquisition and engineering commands to shepherd warfighters into the future by discovering and examining innovative, technically advanced solutions to solve warfighter gaps. One acquisition strategy is to target the Navy Small Business Innovation Research (SBIR) programs. These programs are used to deliver technology solutions to today's warfighters. SBIR's mission is accomplished by applying the agility, dedication, and ingenuity of small business entrepreneurs to the Research and Development (R&D) needs of the Navy. Experimentation and analyses provide a feedback loop to the Program Executive Office (PEO) in support of strategic vision and direction, helping to overcome knowledge and requirement barriers.

While the experimentation route can appear long and fraught with complexity, there are agencies and mentors available to enable and assist the community through these processes. One such agency is the recently established DoN SBIR Experimentation Cell (SEC), available to facilitate, mentor, and train both the small business and government personnel for participation in experimentation events. A key goal of this effort is to highlight the importance of data collection and reporting to gain value from the experiment.

1.1 Purpose & Scope

This is an introductory guide to analysis and reporting, consolidating approaches, best practices, and recommendations. Additionally, it provides the reader with discretionary advice that can be tailored to the individual experiment. Concepts are discussed and explored in conjunction with best practices, with the latter distinguished accordingly. The definition and usage of terms used throughout this guide are in the context of experimentation analysis and are based on empirical research and the authors' experience.[1]

As a dynamic document, this guide will undergo cyclical reviews and revisions to ensure the latest updates are implemented. This work is not intended to be used in lieu of official policy or training, but instead like a simplified study guide. Consult the latest version of the documents listed in References for the most updated information as well as additional exploration materials. It is important to remember analysis and reporting are two key pieces in the larger experimentation process. To learn more about experimentation basics and how each piece connects, consult the Experimentation 101 Guidebook.

2 Why Experiment?

At the most basic level, the purpose of an experiment is to gain knowledge. For an experiment to be determined successful, important questions need to be answered and formally documented. For a Department of the Navy (DoN) experiment, these questions often concern production readiness and maintainability, but most importantly, the impact to fleet and missions. The collection of data and analysis that follows an experiment enables the expansion of knowledge and the answering of those questions. Reporting on what was learned enables that knowledge to be shared and support decision making both for the SBIR and the warfighter.

Conducting an experiment is not a trivial thing. It often has many moving parts, schedules that must be coordinated, logistical/engineering questions to be answered and more. Taking all that into account, one might ask if it is worthwhile to do them at all. The simple answer is yes, because the data collected can hold great value and use.

[1] Further notes on terminology: The verbs "evaluate" and "assess," for example, may get used interchangeably depending on where and how they are being used. To the casual observer, evaluations and assessments might sound like synonyms. In both cases, metrics are collected, and something is measured. However, there are two key differences, the sense of exploration and a standard. With an evaluation, the measurements are graded against an established standard which details the metrics that are to be gathered. In an assessment, processes are broken down and examined with data collected on the components. With no standard to be graded against, assessments allow for exploration into the unknown.

The purpose of this document is to help small businesses understand the purpose of an experiment, and the value they can get out of one. Navigating the analysis and reporting processes effectively can help to obtain beneficial data and use it.

The key to achieving value at the end of an experiment or demonstration is planning. Envision yourself presenting your program to the sponsor or Technical Point of Contact (TPOC). What do they need to know that will help gain funding? What are their questions? What will it take to convince them your program is worth investing in? Will it fulfill the warfighter gap? They are investing in an experiment that will take time/money/energy to make happen. Planning what you are going to deliver for the investment and justify the expense will help in answering these questions.

2.1 Experiments & Demonstrations

While experiments and demonstrations can have comparable qualities, there are two key differences that make them distinct—the "who" and the "new." Both experiments and demonstrations involve putting a technology through its paces and provide an opportunity to collect data. The biggest differences are associated with the following questions: Who will be learning something? Is this a combination of factors that has been tried before, or is this new ground entirely?

An experiment is looking into a new area, with both the organizers on the team and observers learning new things. Typically, an experiment is testing a new combination of factors that have not been tried before with the intent to expand the knowledge base on the topic.

With a demonstration, the activities have been tried before and the outcome is generally known by the organizers. It is likely new for the observers, and they will get to learn how the processes/technology works. However, because the combination of factors has previously been conducted, very little new information is added to the body of knowledge on the topic. For more definitions on terms used throughout, consult the Glossary section.

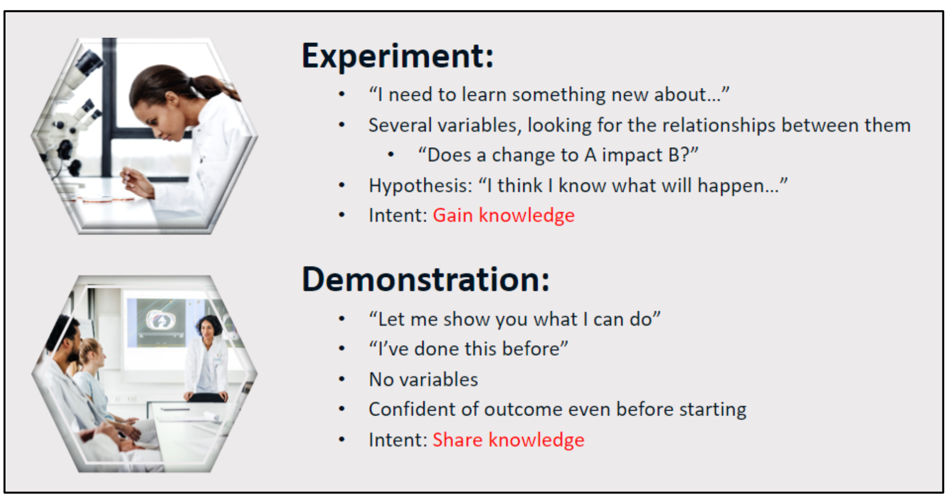

Figure 1 What's the Difference?

2.2 Experiments and Test & Evaluation

Experiments and Test & Evaluation (T&E) events have much in common. They both focus on specified objectives, develop data collection plans, analyze and report results. However, while they seem very similar, their purpose and definition of success/failure make them quite different. The purpose of an experiment is to expand knowledge and study the technology with respect to a warfighter gap. The purpose for T&E is to determine if the technology is able to fill a warfighter gap by meeting very specific standards. Those standards are what separates the two event types. The parameters associated with a warfighter gap an experiment is studying might have known values/needs. These parameters may be defined by the objectives they are attempting to address (for example, specific measurements or targets) but are not as rigid as established acquisition requirements. In an experiment, as long as knowledge is gained, an experiment can be judged a success even if the technology is not able to fill a gap. However, for a T&E event, the requirements are pass/fail standards, and while gained knowledge is always useful, not reaching the standard can cause a T&E event to be judged a failure.

3 Three Step Process

The analytical process for a successful experiment can be viewed as three main steps:

1) The planning beforehand to identify key questions that need to be answered and the associated data that needs to be gathered.
2) Collecting the data during the experiment (to ensure potentially perishable data is not lost).
3) Analysis and publication of the data afterwards to turn that data into knowledge.

These processes enable the transition from Questions to Data to Information to Knowledge/Answers. This document should not be looked at as a set of consequential instructions, so much as advice that can help you think through the process and past potential pitfalls.

The three step process is similar to the construction of a ship, with steps of design, build, and use. Just as with building a ship, the design step is critical to everything that follows, the importance of planning ahead for an experiment cannot be overstated. Throughout the analytic process, focus should always be on the end goal— the answers to the key questions. Because planning is so critical to the overall result, this guide will focus largely on step one, with overviews of steps two and three.

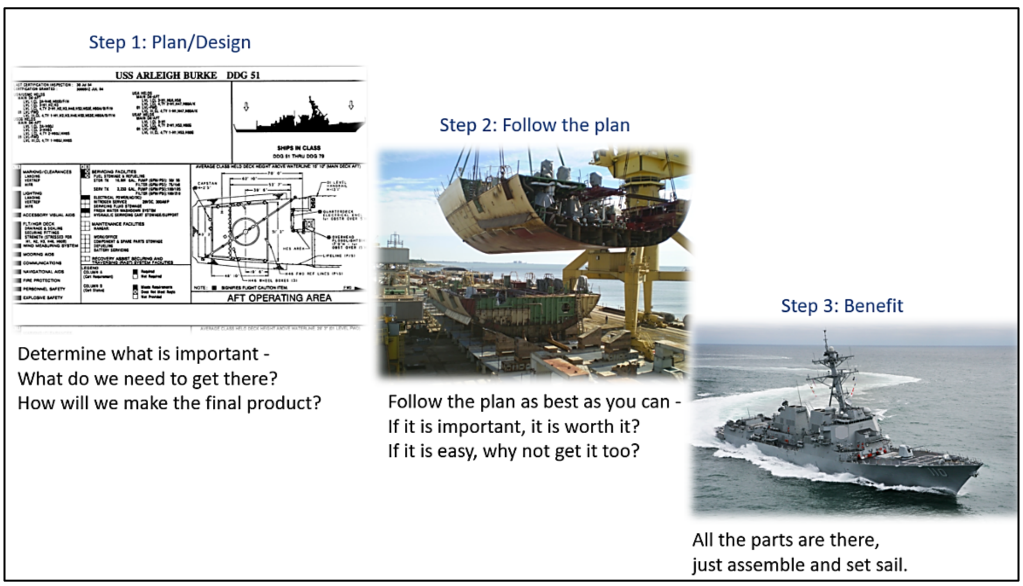

Figure 2 Analysis Process

The Data Collection and Analysis Plan (DCAP) flows through all three steps as the outline of the data within the experiment, detailing the who/what/when/where/how and why of the data. Who is gathering the data, what data needs to be gathered, when, where, and how it will be gathered, and most importantly, why? As one portion of the overall experimentation process, the DCAP is the data collector's script, but is not necessarily set in stone. It is a plan, and if during the execution something turns out to be harder to gather than anticipated, maybe it needs to be dropped from the plan. Conversely, an unanticipated opportunity might surface, which could produce more information than originally planned. The key is the balance between the "why" you need it, the cost/effort of getting it, and the opportunity cost of not collecting.

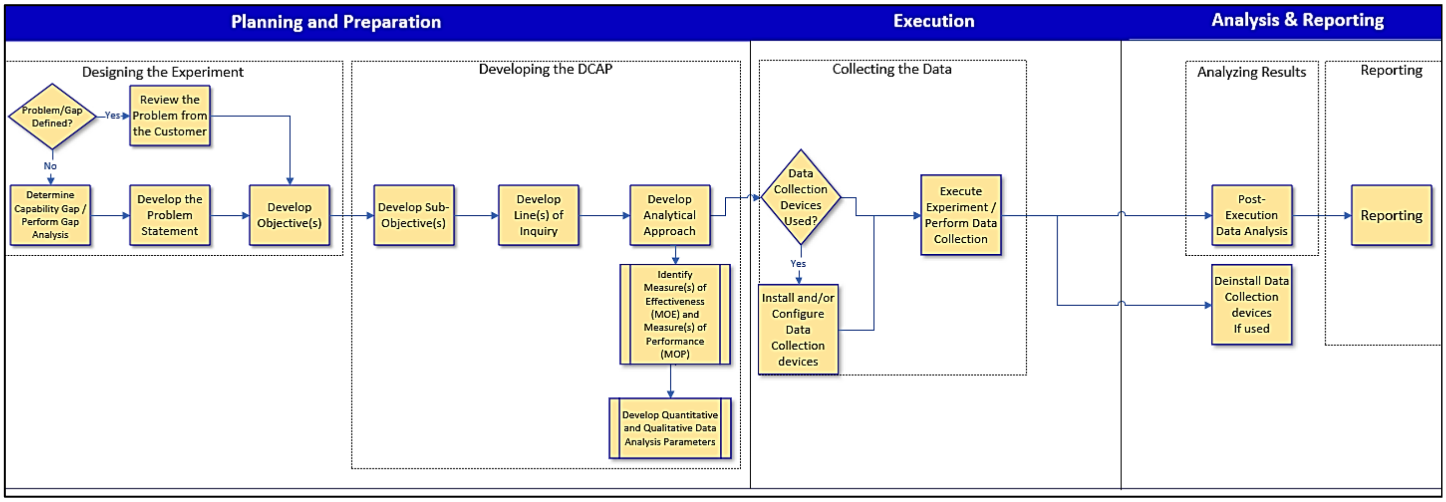

Figure 3 Functional Areas

Experimentation itself fits into the larger cycle of taking an idea from conception to implementation, as shown in Figure 3. The graphic demonstrates the continuous process of utilizing recommendations from the initial experiment as a future roadmap. The results can feed into the planning for the next demonstration or experiment, which will go through a similar process of meeting requirements and objective planning. Upon execution, the data that is reported can move the technology closer to transition, or it can be helpful in developing future concepts and revising tactics.

The most important thing to remember throughout the process is that analysis does not take place in a vacuum, and it is more than just math. You cannot simply look at numbers and extract truth, it is critical to look at the systems, the processes, and the context the number are coming from. The most important row of information in a data table are the labels at the top. As such, having a subject matter expert (SME) available to work with the analyst is critical.

Note of Best Practices

Plan changes can be caused by outside circumstances, so it is wise to be flexible while focusing on what is important for the initiative. Imagine planning for four days of events, but weather is closing in, and you will only get two. You can no longer do all the scheduled events, which subset will you execute? By working with the decision makers during the planning process, the team can know what is the most important and where to focus. This will ensure the capture of the most important runs. Just as an actor needs to know more than just the words in the script, they need to know the character's emotional state and motivations. The team gathering the data needs to know what is important and why so that if/when adjustments need to be made, they can be the right ones.

Initial data can be processed even before the experiment is finished. This early analysis can help extract additional information out of the later iterations by suggesting ways they should be adjusted or focused. No need to wait until all the data has been gathered before any analysis can be started.

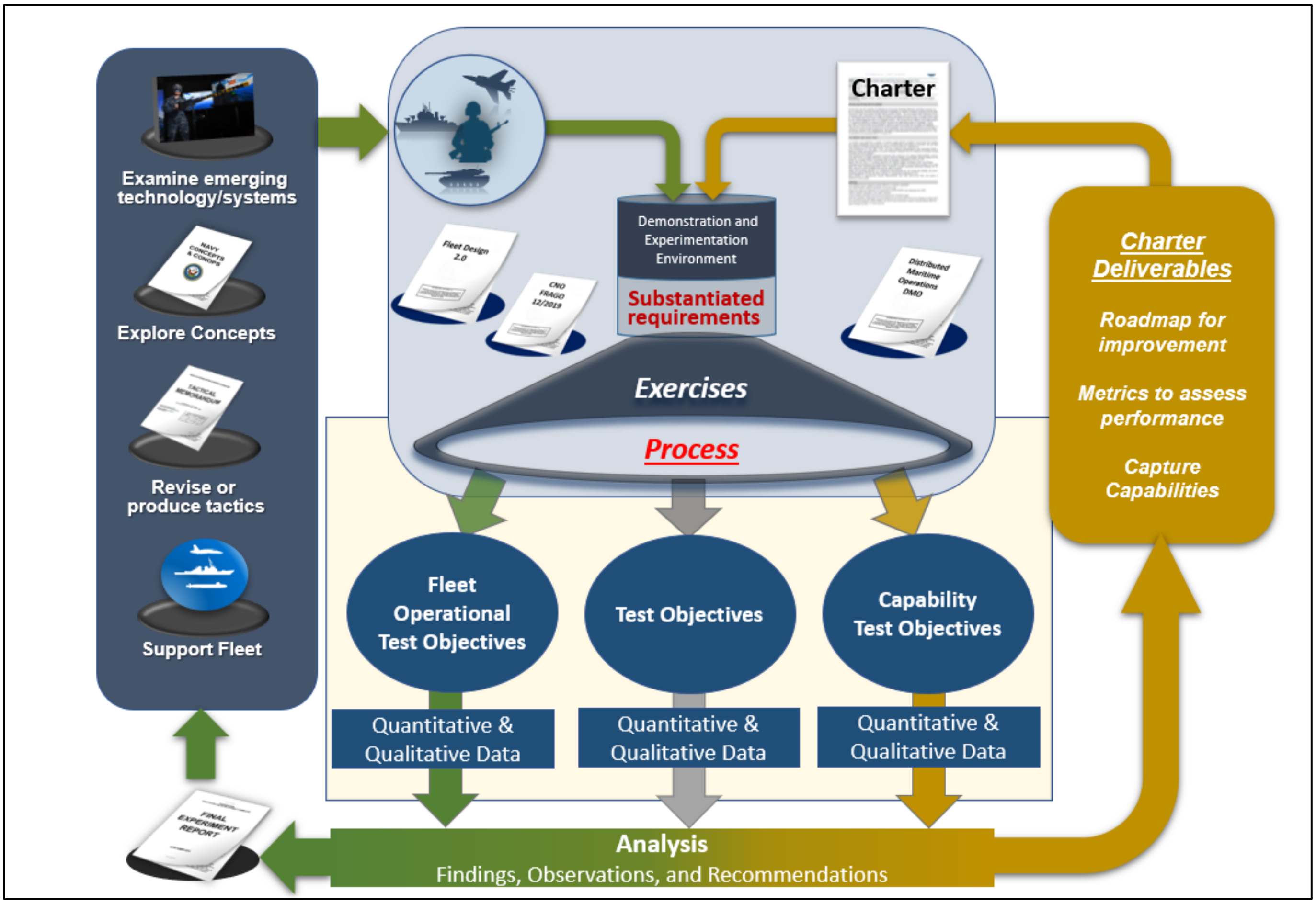

Figure 4 Objective Planning

4 What is Important?

During experiment design and development, it is critical to first identify the objectives. Figure 4 above demonstrates how objectives feed into the overall process, leading to the production of data and the consequential analysis and deliverables. Consider everything that goes into building a ship, whether it be a small littoral craft or an aircraft carrier. Before the first line of a blueprint is drawn, the general purpose of a ship is established (approximate size, tonnage, basic weapon systems, primary mission, etc.) Is it going to be a surface vessel or a submarine – what are the needs of the end user? The same is true for experiment design, the process must start with the customer's needs defining the capability gap and establishing a basic problem statement (what is this technology going to fix?) This will lead to the objectives and ensuring they align with stakeholder's needs. This is accomplished by identifying what makes you different and the information the decision maker needs to know to advance to the next stage of the acquisition process. In the overall process laid out by Figure 3, the experiment design subprocess is magnified below.

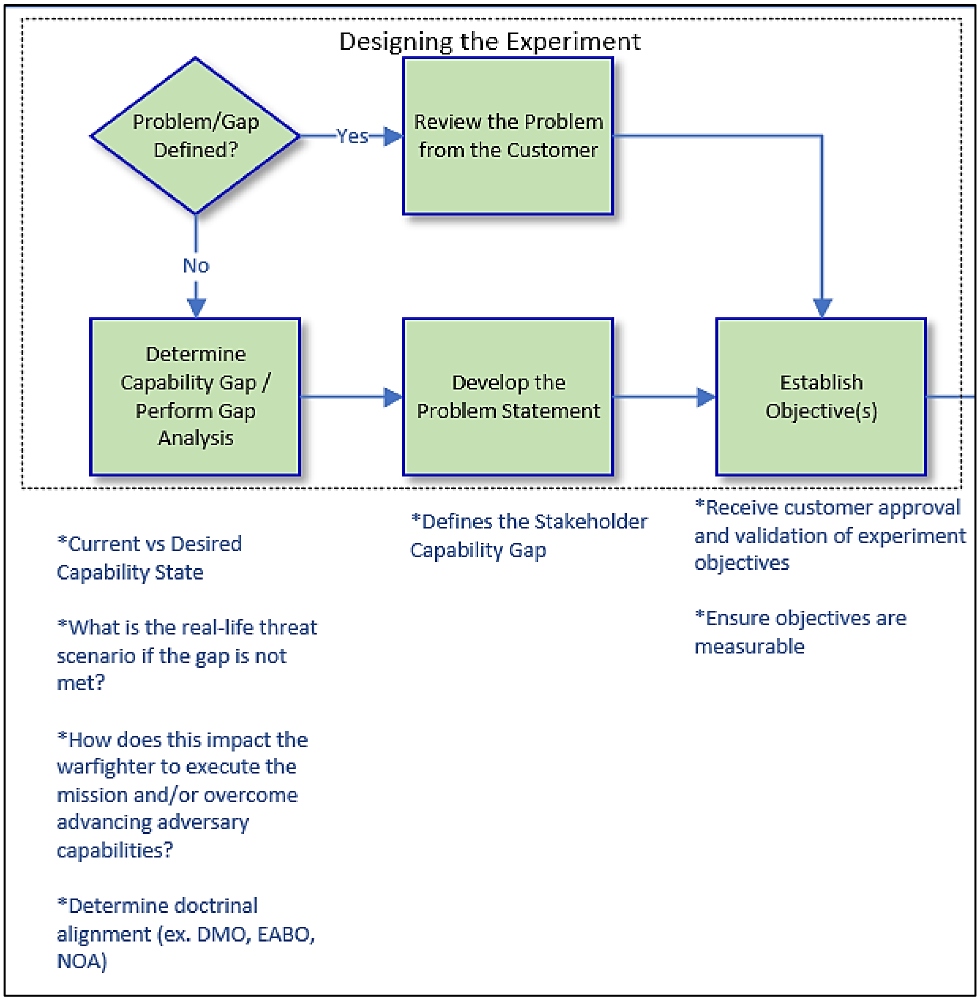

Figure 5 Experiment Design Subprocess

4.1 Understanding Capabilities

Even if you do not know the specific gap your technology will fill, having a firm understanding of your capability can shed light on why the DoN should consider it. How are you improving or replacing what the Navy already has? As a sensor, do you have a wider field of view? As an algorithm, can you process inputs faster and enable quicker/better decisions? As a weapon, do you have more range? As an aircraft, do you have longer dwell time or cheaper maintenance? Examine how your technology stands out and makes things better for the Navy. These differences will give you places to start looking for objectives. Much like building a ship, designing an experiment and gaining value from it is a multi-step process, as demonstrated in Figure 2.

Experiments are designed to collect data that will predict the impact on warfighter gaps if the technology is adopted. This data will enable the Navy to make acquisition-related decisions. Part of designing an experiment is looking at the big picture, defining the data and objectives that must be collected in the DCAP to assess the performance of the technology against the gap. Developing close relationships with key warfighter and sponsor stakeholders can lead to understanding operational needs and a comprehensive knowledge of the stimuli that drive them (e.g., threats, challenges, and opportunities) for current-, near-, mid-, and far-term capabilities. These factors influence future research and development (R&D), funding, and the resource requirements necessary to advance and transition new capabilities to the affected stakeholder(s).

Note of Best Practices

The DoN spends significant amounts of resources to identify, assess and prioritize capability gaps. Qualitative and/or quantitative analysis is used to assess the magnitude and helps create a basis of measurement to assess experiment outcomes, findings, and recommendations by investigating current tactics, techniques, procedures (TTP), doctrine, technology, and their cumulative impact on the warfighter to fulfill current/future mission requirements. Discovering a gap during a tabletop exercise might lead to a loss on paper, but the consequences can be significantly higher if one is discovered under combat conditions. As new technology or TTPs are developed and presented to DoN, they must be evaluated by asking the following: How well do they fill the gap? Is the gap permanently fixed or is this a temporary patch or workaround? How does this new information integrate or interface with existing capabilities? Experimentation assists to answer these questions and can be designed to address numerous other inquiries prior to Fleet implementation.

4.2 Problem Statement

The questions asked depend on the capability gap and the proposed solution. As you formulate these questions, you will also be identifying the experiment dependencies. What do you need to succeed? Are there outside resources or permissions that are critical for the success of your experiment? Identifying them in advance will help in establishing dependencies. Some common categories for dependencies are the constraints, limitations, and assumptions (CLA). CLAs are vital to a successful study. They tie together a study by identifying what must (or must not) and can (or cannot) be accomplished. They frame the study space and set the stage for the study team's methodology development. A constraint is a restriction imposed by the study sponsor that limits the study team's options in conducting the study. A limitation is an inability of the study team to fully meet the study objectives or fully investigate the study issues. An assumption is a statement related to the study that is taken as true in the absence of facts, often to accommodate a limitation.

The following sections of this guide will discuss steps of the data collection, analysis and reporting process (DCA&R) and how they can be used, along with different types of exercises and how experimentation may fit into the exercise. Additionally, an examination will be made of the basics of the processes and ideas to help think through the important steps that make up the planning, preparation, execution, and analysis/close-out phases. Because every experiment is unique, this guide should not be considered a step-by-step recipe. Rather, it is a framework with ideas to identify the right questions and how to find the right answers.

5 Creating Objectives

As stated above, the main goal for an experiment is to add to the knowledge base. To support that, objectives need to be identified for each experiment that can answer key questions and support the decisions. Each objective should focus on something important that is distinctive and measurable. The key to unlocking good objectives is asking questions like: *What is the next milestone? What is the next key decision? Why is that important? What key piece of information can unlock the next door? How will it impact the mission? When you report on the experiment, how will you highlight achievements? What will you tell the funding source they got for their money?* A good place to start answering these questions is the source documents that established the program. What did the Navy state was going to be important? What metrics were established for the source selection?

An experiment's objective can be very broad or very specific, have quantitative or qualitative answers, possibly with multiple subobjectives. A general objective for a vehicle might be to "assess the performance characteristics…" with several characteristics such as speed, turning radius, battery life and range listed as subobjectives or metrics to be gathered. For a sensor, a specific objective might look like:

"Measure the range at which a vessel with a cross section of XX can be detected." Depending on how and who you ask to "Assess the impact X had on mission Y", you could get either quantitative data such as hours saved, or qualitative inputs with a few sentences on what their expert opinion is on the impact.

Note of Best Practices

The way questions are asked can impact the way a test is developed and in turn, the data gathered. For example, consider testing a rope and wanting to know if it could hold 200 pounds. You could ask that question directly, "*Can the rope hold 200 pounds?*" This could lead to an experiment where one end of a rope is attached to a 200 lb. weight and the other end is lifted. If the rope snaps, the answer would be no. If the weight is lifted, the answer is yes. You could repeat the experiment with different pieces of rope, but the data that is gathered will be binary in nature, a yes or no.

Now change the question to "*How much can the rope hold?*" The experimenter attaches the rope to a fixed point and the weight on the other end is increased until failure. The second test would provide "better" data then the first. For each run, if the failure occurred after 200 pounds, it would be a yes, and if it was less the answer would be no. But not only would you receive the information the first test gave, but extra information is yielded as well. How much more than 200 can it hold? What is the "safety margin"? Or when analyzing the failures, how close to 200 did you get? This additional data could be very helpful in the development of a program.

The second version, while "better," has more moving parts, so is it worth it? Should we always go for the extra information? Not always. It depends on the return on investment. It is possible that the first test is easy and low-cost, but the second would be difficult and expensive. If the decision maker only needs binary (yes/no) data, then the extra expense might not be worth it. If the extra information would have value or if the two versions cost about the same, it makes sense to obtain it.

In general, there are four "levels" of data as shown in the table[2]—Nominal, Ordinal, Interval, and Ratio. Nominal data can be binned by category and while it can be sorted alphabetically, it does not enable very much analysis. Consider a list of ship names. They can be studied and examined, but other than being able to say that aircraft carriers are named after presidents and attack submarines are named after cities, the names themselves do not provide very much information. Within an experiment, nominal data is typically categorical, such as the daily weather conditions being clear vs. cloudy. They could also be of a pass or fail nature. Analysis can be done on this data, but it has limits. If the system is not sensitive to weather conditions, then a simple reading might be enough. If weather is important, getting higher order data can be important to record.

	Nominal	Ordinal	Interval	Ratio
Categorical, with labeled variables	✓	✓	✓	✓
Ranks categories in order		✓	✓	✓
Known, has a scale			✓	✓
Has a true zero/absence of what is being measured				✓

Table 1 Levels of Measurement

Ordinal data has a recognizable sequence, but not a scale. The hull numbers of ships can easily be sorted and tell you more than just the order the keels were laid. But they do not provide insight on innovations that were implemented and the differences between them. CVN 76, 77 and 78 were laid in sequence, but while 76 and 77 are both Nimitz-class and thus share many similar characteristics, 78 is a Ford class and has significant differences. Ordinal data can also come in the form of asking for options, which is favored most, or how many stars something should be rated. While that might appear to be higher order because it gives the impression of scale, it is in fact ordinal.

[2] Career Foundry

Interval data has a scale, but no true zero. The scale enables many types of analysis such as averaging and standard deviations. But the lack of a zero can be a limiting factor. 30º, 60º and 90º are all 30º apart, but 60º is not twice as warm as 30º and 90º is not three times that. Ratio data has both a true scale and zero. It looks at multiple capability parameters such as a vessel's speed, fuel capacity, endurance, or a sensor's range or sensitivity.

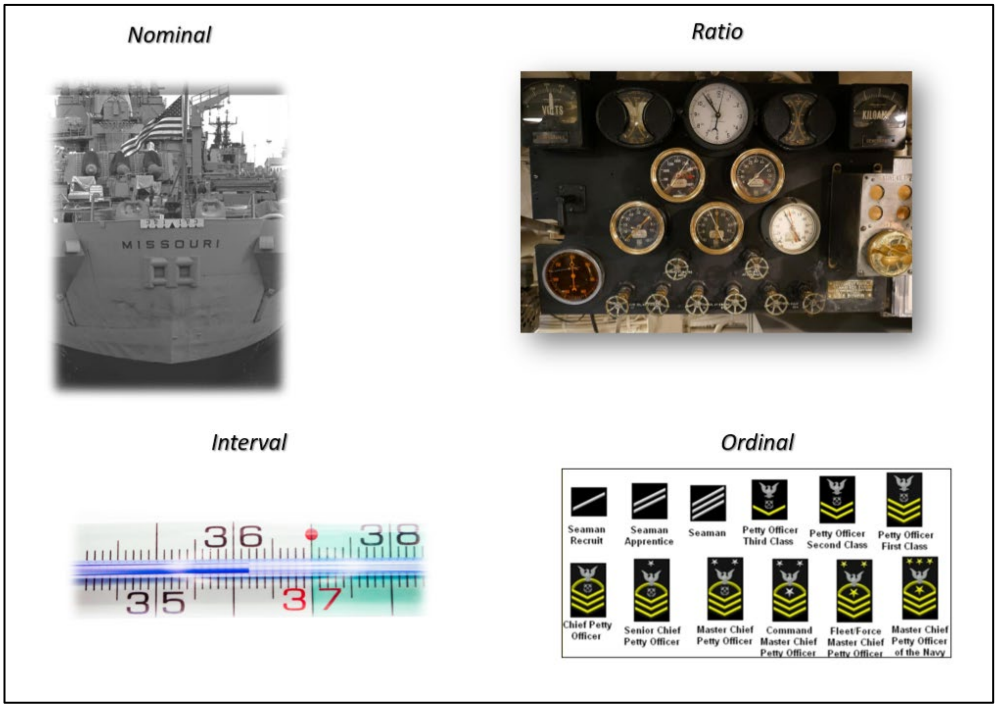

Figure 6 Types of Data

6 Developing a Data Collection and Analysis Plan

The experiment objectives are the first things that need to be established within the DCAP. They are typically closely linked with the capability or knowledge gap. Objectives will be prominently featured as the problem statement and form the foundation that gives the rest of the DCAP purpose. If the experiment is more demonstrative in nature, there may not be a true problem statement, but instead key highlighted features. Next, the hypothesis is developed, asking the question of what cause-and-effect relationship is being investigated. This will in turn lead to the metrics and data that are required to answer the questions and establish the cause/effect relationships.

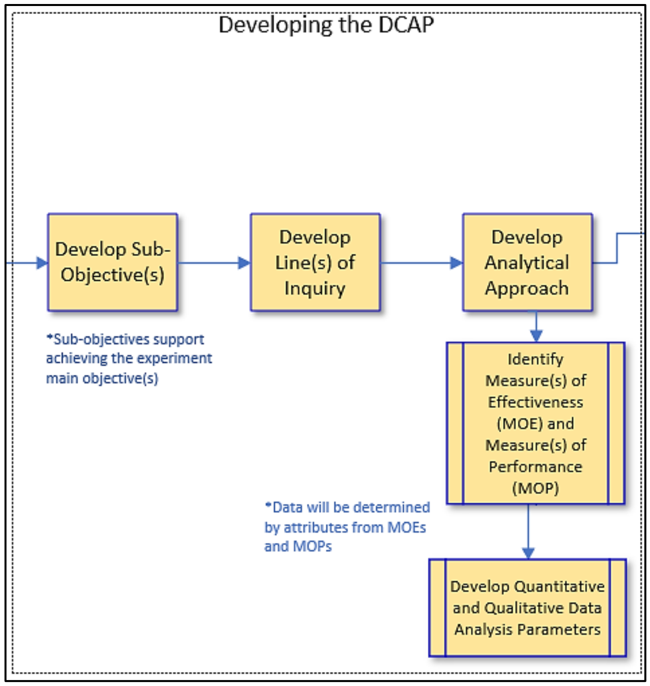

Figure 7 DCAP Development Subprocess

Properly designed and executed experiments are focused on gaining new knowledge, the understanding of systems/operational concepts within environments and situations of relevance. The processes (planning and execution) of an experiment can be compared to the building of a ship as described earlier. There are three main phases to the construction: design, construction and finally, use. An experiment is similar as it is planned, executed, and studied.

Within shipbuilding, there are many steps that must take place prior to laying the keel, such as consulting with the users or crew to ensure requirements and expectations are established, along with the drafting of blueprints. Developing stakeholder relationships paired with detailed planning are just as essential to ensure the experiment design and analytical strategy align with experiment objectives. From this point, available resources and potential alternatives can be investigated. Just as a ship is more likely to succeed if the preliminary steps are done well, experiments entail an iterative process where the level of success is dependent upon sub-elements and planning intricacies. The planning process is a continuous cycle that must be monitored throughout experiment design and execution.

As the supporting structure for experimentation, the DCAP enables analytical rigor specific to the experiment objective(s) and associated line(s) of inquiry (LOI). Developing a predetermined research method is a critical strategy to observe, collect, assess, and report on experiment data. Research design and methods are different, but closely related, because good research design ensures that the data obtained will more effectively answer the important questions. Each DCAP must be tailored to the experiment and entail a holistic approach towards data collection and analysis procedures.

The DCAP comprises the main elements of the experiment, including the problem statement, objectives/sub-objectives, LOIs, CLAs, critical questions, measures of performance (MOPs) and measures of effectiveness (MOEs). Furthermore, it provides the analytical rigor required to collect and assess quantitative and qualitative data. Just like a blueprint for a ship, the DCAP provides detail and clarity to see the individual parts/components, as well as the big picture. Each step in the experiment will add a piece of knowledge helping to fill the requirements. Developing a research method is a critical strategy to observe, collect, and report on data obtained during an experiment. Research design and methods are different, but closely related, because good research design ensures that the data obtained will more effectively answer the research question. For any type of experiment, the DCAP explains the end-to-end structure. Reference Section 12.1.

7 Collecting the Data/Execution

Once the pieces are in place and the event is ready, it is time to follow the plan and collect the data. In general, data collectors should try to gather as much data as possible, working with a balance between the importance/utility of the data and the effort required to gather it. This balance should be accounted for in the DCAP by consulting SMEs on the availability/ease of access and decision makers on the anticipated value that data will bring. Refer to Figure 3 to review where this stage falls in the overall process.

There are three general categories of data: Automated, Observation, and Interactive. Each of these data types have their own factors, but two of the most important factors are common between them—data collection opportunities are fleeting and the details are important. It is much easier to record gather good data at the time of execution then to try to recreate them later.

Automated data includes elements that are generated by systems and recorded without any manual intervention. Consider examples like data logs, GPS tracks, telemetry information, or any other data that is generated by an automated system and can be downloaded as a file. Automated data files should be archived on an established schedule. Every data system has its own cycle rates and storage capacities. It is critical that data files be downloaded well in advance of being lost. If data files are not archived in a timely manner, they might get lost or overwritten. If at all possible, files should be stored in at least two locations to ensure that there is a backup available, should a primary storage device get corrupted.

Observation data involves a data collector manually recording data elements as they occur. The collector does not interact with the experiment or the participants, the collector is simply making an observation, such as recording the timestamp as decisions are made or events transpire. Observational data should be recorded in detail as soon as possible. It is a mistake to rely on observer memories alone. This type of data can be the most fleeting and hardest to recreate after the fact. If an experiment involves an aviation component, data on the local windspeed might be useful for post-event analysis. Recording localized windspeed and direction at the time of the experiment can be relatively easy if an appropriate censor is available, but getting accurate data after the fact can be much harder.

Interactive data involves a data collector interacting directly with the experiment or the participants to generate the data and record it at the same time, such as interviews and surveys. Interactive data collection can also be achieved by issuing user surveys or forms. This method can be utilized in addition to observation, however it is important to allow an appropriate timeframe. Formulate a set of questions ahead of time relevant to the topic and targeted to accessible end users. The questions will need to be presented in a way that is easily understood in order to promote detailed, objective, accurate answers. If interviews are not conducted in a timely manner, the participants could forget some of the details they observed. Records need to be accurate and factual, with special attention to constants, multiple measurements and other varying factors.

Note of Best Practices

There are many types of events that can be large- or small-scale, live or simulated, and with varying thresholds of entry levels. For specific events and descriptions, reference the Experimentation 101 Guidebook. In relation to designing a significant experiment, they can be categorized simply as either scripted or unscripted.

Scripted events are characterized by a sequence of orchestrated activities or a scenario; any inclusion of new technology needs to fit in to the established story. If you are looking for mission impact, qualitative/non-quantitative data, this type of event can be ideal. Scripted events typically have numerous SMEs that can provide opinions on the impact the new technology can have on the mission. However, because each part must fit into the overall story and follow the scenario timeline, it can limit the number of iterations and parameters that can be tested.

An unscripted event will typically have fewer moving parts and lack some of the realism that a scripted event can have. What it does have is significantly more flexibility. Because the schedule is built around the experiments and not a specific scenario, the experiment designers have more control over the parameters and replications. Sequences can be repeated with controlled variations enabling the collection of quantitative data.

8 Analyzing Results

Once an event has concluded, the final step in the analysis process is to examine/interpret the data and see what can be learned. The DCAP should not only contain details for the data collection plans and what will be collected, but also what will be done with it. It is critical to enlist the aid of SMEs to help interpret the results, leading to the validation (or invalidation) of the hypothesis. While the DCAP is important, it is encouraged to not limit the analysis just within the plan. The data may contain additional information that was not anticipated and should be checked for potential insights.

Once the data is in hand, it is ready to be processed. From this point, the data needs to be checked and cleaned for any anomalies, and basic summary statistics calculated. Organizing it into tables, or even creating something as simple as a histogram might provide insight on the makeup of the data and the relationships between the variables. Building graphs and creating other visuals can also help interpret the data. General exploring can be useful, but the focus should be on following the analysis plan that was created and determining if the expected results/relationships are supported.

The purpose for analysis is to sift through the data and cultivate it from its raw form of information to knowledge and hopefully gain some insights along the way. Analysis is about finding the significance or implications of the experimental findings. It is about determining the relationship between variables, understanding the impact they have on each other and potentially recommending areas for future research. It is important to relate results to the aims of the experiment through summarization and explanation and to lead cross-initiative analysis of the generated data. Oversight of specific initiative analyses for internal consistency and connection can contribute to overall program goals. Thus, it is important to maintain documentation, previous event results, and associated relevant data.

When a study is conducted at the end of the experiment, many numbers remain: the input values (settings) and the output values (results). The challenge is to extract from the data a meaningful summary of the behavior observed and a conclusion regarding the influence of the experimental treatment (independent variable) on participant behavior. Statistics can provide an objective approach to performing this process. However, there is more to data than just the mean and the variance.

Note of Best Practices

One key pitfall that must be avoided is thinking that correlation equals causation. Correlation is when two sets of data seem be closely associated with each other. An example of the pitfall would be the link between commissary sales and sunburn. A quick look at that data in the summer would show that the values are highly correlated when ice cream is involved. Every month that has high ice cream sales has many sunburn cases and months with low ice cream sales do not. The correlation says they must be linked; therefore, ice cream causes sunburn. This is of course not true, it just so happens that during summer months when it is hot, the crew purchases more ice cream and goes outside in the sun. Summer/hot weather is the cause for both, and there is no real connection between ice cream and sunburn. This type of issue is why it is critical to have SME availability during the analysis phase to ensure that the interpretation of the results is appropriate.

9 Reporting

The results are ultimately the pot of gold at the end of the rainbow, so to speak. The experiment is complete, the data collected and processed, now it is time to impress the decision makers. You will want to present these findings and show the decision maker the fruits of the labor. The format of the final report is as varied as the technologies being tested, no matter if it is a PowerPoint presentation, Excel spreadsheet or a PDF. What matters is the bottom line of the findings. What did you learn? How much of an improvement is this new development? What did the SMEs say? What is the expected impact on the mission set? There is nothing like hard data to convince a skeptical reader that you have a solution to one of their problems. Once you demonstrate that to the audience, rewards and progress will follow.

The results could come in the form of any type of statistic or summary metric. The numbers are certainly important as the hard data that was obtained during the experiment, but the report should focus on the meaning behind them. The decision maker will want to know the "so what?" This question will demonstrate the impact the technology can have on the systems it interacts with, as well as the mission it will be called on for. It will establish the value and potential next steps after the experiment.

There are several commonly used formats and, while not universal, will be summarized here for awareness. Events such as Fleet Experimentation (FLEX) and Trident Warrior (TW) are considered a gold standard when examining report structure. Each venue has its own unique structure, expectation, requirements and depth so it is important to note that each format is not a "one-size-fits-all" solution and may be different than the examples listed here. Overall, what disseminates the different types of reports is usually the level of detail and expediency of review and release.

Quicklook reports are a specific type of report that tends to be modular and contains sectioned information on the data collected as well as background information. An accurate, high-level Quicklook report considers efficiency in publication, main points, and summaries. Goals should be stated clearly and answered by the time of event conclusion. Utilize the DCAP for applying the data into this report. Final Reports can contain a higher level of detail. These reports usually have more time allotted as they involve more in-depth research and analysis. Factors such as cause and effect and future implications can be explored in final reports.

The following examples are more of a format as they are based on usage and/or a specific audience. There are report forms that examine key areas of Doctrine, Organization, Training, Materiel, Leadership, Personnel, Facilities, and Policy (DOTMLPF-P). This style of report is a way to section various military areas after the exercise and examine the value. DOTMLPF-P reports follow an established structure and will evaluate how an individual initiative will impact areas such as training, maintenance, etc. A data analysis report commonly uses qualitative and quantitative data to analyze as well as evaluate a strategy/process while facilitating decisions based on this experimental data.

For more examples and templates, consult the Reporting Templates section. While the format or type of the report can vary, you will also want to ensure that the intended recipients have access to it. Asking these questions ahead of time ensures that the report is written correctly and then submitted through the proper channels.

10 Closing

Because of the unique factors for every technology/venue, there is no clear step-by-step recipe for building an experiment or the associated analysis plan. Ideally, the concepts that have been shared here will facilitate gaining value out of an experiment. Experiments can require significant resources. Taking the technology out in the field and playing with it can be fun and useful in demonstrating what you can do. However, there is nothing like hard data to convince a skeptic that you have exactly what they need. A successful demonstration might also generate some good pictures. While a picture might be worth a thousand words, good numbers could be worth many thousands of dollars. The value to be extracted out of an experiment is in the data. The way you display and analyze that data for key players can be crucial in transitioning a technology to lasting impact.

11 Appendices

11.1 Data Collection and Analysis Plan Excerpt

The below appendix is an excerpt or outline of what a DCAP could potentially look like. Please note this is only for example purposes. Each DCAP must be tailored to each individual technology and initiative. Some venues might have a specific set of topics that they want to make sure are captured, so there may be a set outline. Other venues might leave the DCAP development fully up to the experiment designer.

The areas covered here are potential areas and are not “required”, but they are included as ideas to consider. The most important thing to be captured in this document is the “why.” Why are you conducting the experiment? What do you hope to gain out of it? Remember, this is not a math text that talks about statistical tests, it should describe the data that you wish to capture and descriptions for how and why. A very detailed DCAP can even include a schedule for when each event in the experiment will take place. The more detail that can be included, the better.

I. Introduction

- **Summary of Tasking**:
- **Problem Statement**:
- **Hypothesis**:
- **Primary Analysis Objectives**:
- **Event Description**:

  a. The primary mission area is [describe here]

II. Analysis Objectives and Critical Questions (CQ)

1. **Analysis Objective 1:** This objective focuses on [describe here]

   CQ 1.1:
   Metrics: Capability (Y/N); Accuracy (Likert Scale 1-5)

   - Assumptions:
   - Data Types:
   - Who Collects: Preferred:
     Alternate:
   - When Collected:
   - Where Collected:

   CQ 1.2: [Repeat as needed]
   Metrics: Capability (Y/N); Relevance (Y/N); Completeness (Y/N); Availability (Y/N):

   - Assumptions:
   - Data Types:
   - Who Collects: Preferred:
     Alternate:
   - When Collected:
   - Where Collected:

2. **Analysis Objective 2:** This objective will focus [describe here]

   CQ 2.1:
   Metrics: Capability (Y/N) to Timely (Likert Scale 1-5)
   - Assumptions:
   - Data Types:
   - Who Collects:
   - When Collected:

3. **Analysis Objective 3:** This objective will focus on [describe here]

CQ 3.1:
Metrics: Usable (Likert Scale 1-5) to….

- Assumptions:
- Data Types:
- Who Collects:
- When Collected:

4. **Analysis Objective 4:** This objective will focus on [describe here]

[Note format if including graphics/figures]

CQ 4.1:
Metrics: Robust (Likert Scale 1-5)

- Assumptions:
- Data Types:
- Who Collects:
- When Collected:

5. **Operational Environment Data**: Environmental data that will be used to categorize and provide context during the analysis that will be conducted with the other data sets.

5.1: What were the observed weather conditions?

- Loading / Unloading Area
  - Sea State
  - Sea Direction
  - Wave Type (rolling, cresting, chop, calm, etc)
  - Wind Direction
  - Wind Speed - sustained
  - Visibility (unlimited, 100 meters, etc)
  - Weather (clear, rain, fog, etc)
  - Temperature
- Operating Area
  - Sea State
  - Sea Direction
  - Wave Type (rolling, cresting, chop, calm, etc)
  - Wind Direction
  - Wind Speed - sustained
  - Visibility (unlimited, 100 meters, etc)
  - Weather (clear, rain, fog, haze, etc.)
  - Temperature

Metrics: Capable (Y/N); Effective (Y/N) ; Robust (Y/N)

- Assumptions:
- Data Types:
- Who Collects:
- When Collected:

5.2: What were the reported weather conditions from other sources? (Buoys, military base, etc)

- Data Types:
- Who Collects:
- When Collected:

6. **Ground Truth Data**

III. EXPERIMENT DESIGN

1. **Experiment Design**:
   Experiment Location: The experiment will be conducted in [list location, include a map image below if desired]

2. **Specific experiment vignette/serials to be conducted:**
   a) [Describe here]

IV. DATA COLLECTION METHODOLOGIES

Data will be collected in three methods:

1. **Automated collection**:
2. **Observation logs**:
3. **Survey collection**:

V. ANALYTICAL CONSIDERATIONS

Q 1.3 [List considerations here and repeat as needed]

VI. ANALYSIS PRODUCTS

1. Quicklook Report: Produced by [enter organization]
2. Post-event Analysis Workshop Schedule: There is no post-event scheduled. The data collected will be compiled and analyzed by [enter organization] and [enter organization]
3. [Sponsor] will provide any lessons learned from the execution of the exercise.
   a. Analysis Summary Reference Document: Produced by [enter organization]
4. Final Report: Produced by [enter organization]

11.2 Reporting Templates

Many reports may focus on the training elements instead of the experiments and in some cases upon conclusion, the organizers consider everything complete. They may list the experiments in an After Action Report, but their focus remains on the units trained instead of reporting any experiment results. As this guidebook has emphasized, there is significant value in analyzing the data generated by a well-run experiment.

To enable merging individual reports into a larger overall report, some events have specific templates. In these cases, the event organizers will provide their template and/or examples to be followed. Other events, while interested in seeing the experiment results and wanting to share information as much as permissions allow, give teams a freer hand in the structure of their report. In these cases, it is recommended to return to the basics and focus on decision maker needs. Consider the types of questions asked at the beginning and the types of findings the data can support.

The following documents are suggested structures, examples, and templates for various reporting styles at the unclassified level. Types of reports shown include Limited Technology Experiment (LTE), Technical, FLEX, Lessons Learned, and Military Utility Assessment (MUA) reports. It is essential not only to verify ahead of time what time of deliverable is expected, but to choose the right format that is best suited for the data and associated results. Each event will have a variety of functions and focus. Some seek to train, improve processes, showcase new technologies, or brainstorm for solutions. Your report should reflect that purpose and establish value to the stakeholder.

A MUA is an assessment of the operational utility of a system and commonly completed at the end of a Joint Capability Technology Demonstration (JCTD). MUAs often contain the critical elements of information such as description of the capability gap(s); associated tasks, conditions, and operational performance standards/metrics; and how the material and non-material approaches and analyses from the JCTD addressed these factors.[3]

A technical report can serve as a standard template for a variety of topics and can be used if the venue does not require a pre-set format. This type of report is flexible and can take the form of an After Action or Quicklook report. An LTE report is very similar to a technical report in that it contains similar front matter structure, but it is more focused on the individual experiments. It is more condensed than a larger FLEX report. FLEX reports are used for fleet experiments and are highly detailed with data points and observations.

By definition, a lesson learned is information that contributes to the military's corporate body of knowledge, producing increased process efficiency and improved execution of future operations. Additionally, it should provide value added to existing federal policy, doctrine, TTPs, training, systems or equipment.[4] Lessons Learned reports can be a variety of different formats, even Excel spreadsheets, but all should contain detailed problem areas, responses, and corrective actions. These reports are process-driven and focused on future improvement, rather than the technology tested.

[3] AcqNotes, 2021
[4] CINCPACFLTINST 3500.37A

LTE Report Structure

Executive Summary

Introduction

- Objective
- Background & Hypothesis

Experiment #1

- Method
  - Participants
  - Design
  - Procedure
- Data Analysis
- Results
  - Preference
  - Questions & Answers
  - Information Products
- Discussion

Experiment #2

- Method
  - Participants
  - Design
  - Procedure
- Data Analysis
- Results
  - Preference
  - Questions & Answers
  - Web Accesses
- Discussion

General Discussion

- Summary
- Further Research
- Recommendations

References

Appendix A: Usage/Activity Survey

- Results and Correspondence to Experimental Design

Appendix B: Rater Instructions

Appendix C: Results Tables

Technical Report Template

Abstract

The abstract is a one or two paragraph summary of the work. It stands alone with no reference to figures, charts, or tables in the text. The line spacing default is double-spacing for academic reports. Other manuscripts may require different line spacing options. The abstract should not exceed one page of double-spaced text.

Table of Contents

This section should begin on a new page. Microsoft Word® has an automated table of contents (TOC) feature under the "References" menu. For this feature to work, you must use the preset heading styles (which you can modify). However, you can also generate a table of contents manually. The TOC is the last element of the report to be completed. If changes are made after the TOC is created, be sure to update to TOC as well.

List of Figures

This section should begin on a new page. Microsoft Word has an automated feature under the "References" menu called "Insert Table of Figures." For this feature to work, all figures must have a caption. This should be one of the last pages to be completed. Some examples follow:

1. Test matrix 5
2. Design prototype 6

List of Tables

Continue to use the automated feature under "Insert Table of Figures." Change the caption label from “figure” to “table.” For this feature to work, all tables must have a caption. This should be one of the last pages to be completed. Some examples follow:

1. Model parameters 3
2. Data under nominal conditions 10

I. Introduction

This section should begin on a new page. This should use "Heading 1" style font settings for the heading, then “Normal” style for the content.

II. Background

This section does not begin on a new page and should use "Heading 1" style font settings for the heading, then “Normal” style for the content.

a. First Subheading

Subheadings are sections beneath headings. These sections should use "Heading 2" style font settings. Subheadings and sub-subheadings are not mandatory. However, if there is one subheading, there must be at least a second subheading. Otherwise, there is no reason for the subdivisions under the primary headings.

b. Second Subheading

This paragraph is repeated. Subheadings are sections beneath headings. These sections should use "Heading 2" style font settings. Subheadings and sub-subheadings are not mandatory. However, if there is one subheading, there must be at least a second subheading. Otherwise, there is no reason for the subdivisions under the primary headings.

i. First Sub-subheading

This paragraph is repeated. Subheadings are sections beneath headings. These sections should use "Heading 3" style font settings.

Subheadings and sub-subheadings are not mandatory. However, if there is one subheading, there must be at least a second subheading. Otherwise, there is no reason for the subdivisions under the primary headings.

ii. Second Sub-subheading

Avoid any further divisions under the sub-subheading. Otherwise, the number of divisions becomes distracting and difficult to follow.

III. Methodology

This section does not begin on a new page and should use "Heading 1" style font settings for the heading, then “Normal” style for the content.

IV. Results and discussion

This section does not begin on a new page and should use "Heading 1" style font settings for the heading, then “Normal” style for the content.

V. Conclusions and recommendations

This section does not begin on a new page and should use "Heading 1" style font settings for the heading, then “Normal” style for the content.

VI. Acknowledgements

This section allows authors to acknowledge contributors and other sources that are not appropriate to list in the references section. Example:

This work was conducted under Grant No. 12345, administered by X. The authors are also particularly grateful to Dr. Jane Smith for her insight into the nature of Y.

VII. References

This is the last section of the report, prior to any appendices. The references should not be double-spaced, but single-spaced. For a technical report, use the CSE style.

[1] Reference 1 information.
[2] Reference 2 information.
[3] Reference 3 information.

VIII. Appendix A: Place the title of appendix here

Provide appropriate appendices as necessary. Each appendix should begin on a new page.

Major Event Example

[Sample] Table of Contents

Section I EXPERIMENT DESCRIPTION

1.0 Introduction

1.1 Fleet Battle Experiments Purpose and History

1.2 General Description

2.0 Initiative Descriptions

Section II PRINCIPAL RESULTS

3.0 Principal Results

3.1 Summary of Findings

3.2 Initiatives' Context

3.3 Experimentation Status and Recommendations

Section III RECONSTRUCTION

4.0 Experiment Reconstruction

4.1 Scenario and Timeline

4.2 Actual Setting

4.3 Joint Forces: Live and Computer Simulated Forces

4.4 Operations Overview

Section IV KEY OBSERVATIONS

5.0 Initiative Key Observations

5.1 Experiment Objectives

5.2 Analysis Specifics

5.2.1 Experiment Design

5.3 Baseline Model

5.3.1 Background

5.4 Experiment Design, Data Collection, and Analysis Methods

5.5 Sub-Initiative Observations

5.5.1 Key Attributes

5.6 Input Parameters

5.7 Model Execution

5.7.1 Sample Results

5.8 Maritime Planning Process Key Observations Summary

5.8.1 Structure

5.8.2 Organization

5.8.3 Management

5.8.4 Feedback

5.8.5 Optimization of Resources

5.8.6 Situational Awareness

5.9 General Conclusions

Section V SUMMARIES

6.0 Lessons Learned

6.1 Value Added

6.2 Appropriate Missions

6.3 Conditions and Design Features

6.4 Suitability

6.5 Characteristics

7.0 Joint Interoperability (USN/USAF), Engagement, Timelines, etc.

7.1 Objective

7.2 Analytical Questions

7.3 Findings

8.0 Associated Analyses

Section VI APPENDICES

Appendix 1 Master Scenario Event List

Appendix 2 Participants

Appendix 3 Data Collection

Appendix 4 Initiatives, Data, and Analysis

Appendix 5 Collaborative Tools

Appendix 6 Knowledge Management Supported Analysis

Appendix 7 Observations, Comments, and Suggestions

Appendix 8 Network Analyses

Appendix 9 Human Factors

Appendix 10 Operational Sequence Diagrams

Appendix 11 Acronyms

After Action Structure (Mission-Focused)

[Mission Title]

Overall Classification: **Date:**
Organization: **Event Name:**

Event Summary:

Mission and Intent:

Objectives: [Tasks and Objectives]

Locations: [Include Coordinates]

Participants:

Organizations:

Subject Matter Experts:

Linked Observations:

Observation Summary:

Observation Details:

Linked Issues:

Issue Summary:

Issue Details:

Comments:

Milestones: [Include Description, Location, Start/End Dates, Status in table format]

Conclusion:

Point of Contact & Signature:

Attachments:

Lessons Learned Template

[EVENT NAME] Lessons Learned
[Location, Date]

Background:
Team members [participant names] visited [event, location] to review [technology, performer, etc.] Include project summary and event description here as well.

Objectives:
1. Overarching objectives can be listed here.

Schedule:
MM/DD/YY: Events of the day
MM/DD/YY: Events of the day

Lessons Learned:

- *Observation*: Description of project issue and any associated signs. Were there any signs of impending failure?

- *Recommendation*: Corrective action and response for future performance. Corrective actions usually include a response to either **mitigate**, **avoid**, **transfer**, or **accept**.

What are the key potential risks and how will these be addressed in the future? Include any applicable deadlines.

Discussion:
Implications, future goals and additional comments can be listed at the conclusion of the document.

Please note that the template above is just one example of a commonly used Lessons Learned format. Refer to graphic[5] below for an overview of the process.

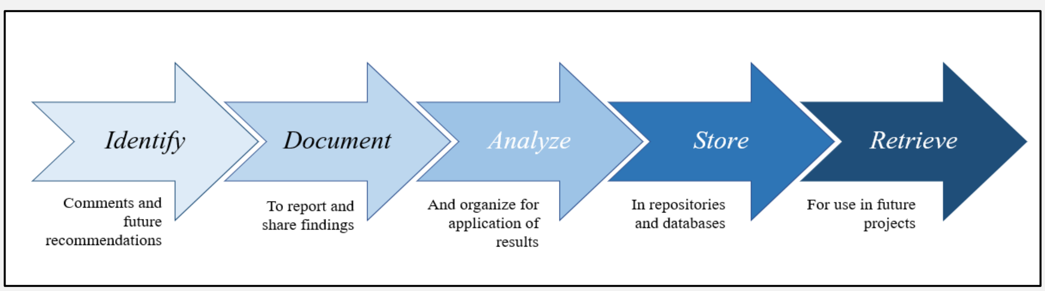

Figure 8 Lessons Learned Process

[5] Rowe, S. F. & Sikes, S. (2006)

MUA Template

Military Assessment Report[6]
Date

[Technology In Focus]

Ref: [Reference to the reviewed report], Referee: [Name]
Interview: [Name of researcher]

Introduction
[A summary on the technology's present and predicted states. The description should relate to TRL assessments. The assessments made in the reviewed report are held for true.]

Identified Possibilities and Constraints
[A summary on the technology's possibilities and constraints. The assessments made in the reviewed report are held for true.]

Suggested Military Use
The following military use for [technology in focus] are suggested in the reviewed report:

- [A listing of suggestions from the reviewed report]

Assumptions
The concept scenarios are based on the following assumptions:

- [Assumption necessary to support the scenario described later, regarding maturity of underpinning technologies or other developments in society]
- [Next assumption]
- [Continue as necessary]

Concept Systems and Scenarios In [YYYY]

Concept system 1 [name of concept system]

[A suitable description of a concept system exploiting the technology in focus]

Scenario 1 [name of scenario 1]

[A suitable description of a scenario describing beneficial use of the technology in focus, either from your military actor's perspective or from a potential adversary perspective. Ideally the scenario includes a description of a challenging military situation on suitable command level and an operational environment.]

Concept system 2 [name of concept system 2]

[A suitable description of a second concept system exploiting the technology in focus]

Scenario 2 [name of scenario 2]

[Another scenario describing other aspects of the technology in focus]

SWOT- Analysis

[6] Template adapted from Andersson et al., 2019

The following strengths, weaknesses, opportunities and threats with [the concept for a technical system in the name of] scenario were identified at the seminar:

Strengths:
- …

Weaknesses:
- …

Opportunities:
- …

Threats:
- …

Assessment of Capability Impact

[Describe the impact the technology is likely to have on warfighter capabilities. Focus on the military effectiveness dimension, which is the capabilities that will benefit, and why. Consider Effect, Mobility, Sustainability, Command and Control, Protection, and Intelligence/Information.]

Assessment of Footprint

The following list is a compilation of anticipated footprints on capability development if the technology is to be used as described.

Influencing factor	**Footprint**
Doctrine	[Ex. An overview of logistic concepts will be necessary…]
Organization	[Ex. Could have great impact on logistic organization due to …]
Training	[Ex. A new organization might render …]
Personnel	[Ex. …will render new demands for education and training. …]
Materiel	[Ex. ... manufacturing could cause health risks …]
Facilities	[Ex. …facilities … must be modified to reduce health hazards.]
Leadership	[Ex. …will require change management to…]
Interoperability	[Ex. common standards for quality assessment and test required. …]

Assessment of the Need for Military R&D

[The assessment group´s view on actions needed from the military R&D should they wish to facilitate the introduction of the technology into service.]

Conclusions On Military Utility and Recommendations

[The final discussion of the assessment group is documented, and the military utility of the future technology is assessed significant, moderate, negligible or uncertain. The discussion is followed by a recommendation to R&D policymakers.]

The example above is an example of the types of topics an MUA may cover. Refer to the figure[7] below for an overview of where the MUA fits in the overall development process.

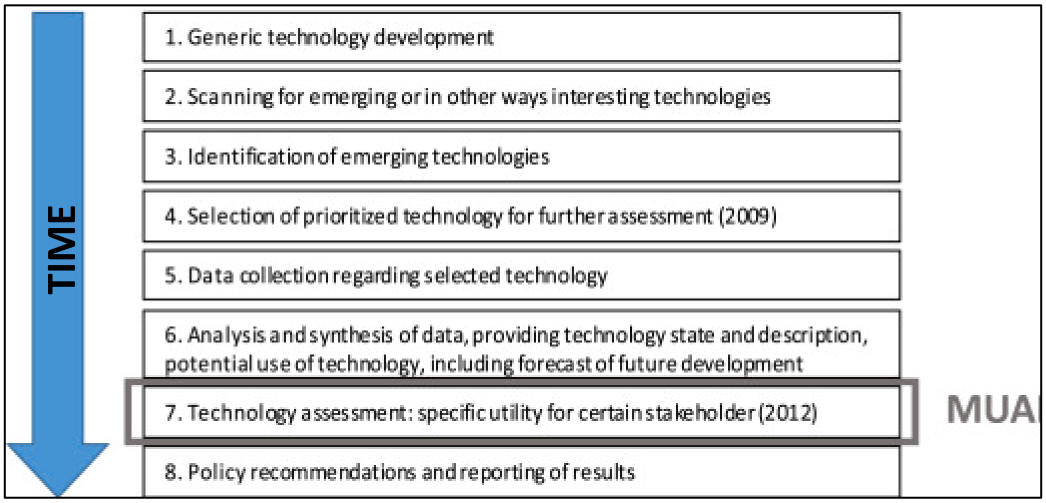

Figure 9 Sequence of Technology Forecasting

12 References & Further Reading

1. Alberts, D. S., &; Hayes, R. E. (2002). *Code of Best Practice for Experimentation*. Washington, D.C.: DoD Command and Control Research Program. Retrieved from http://dodccrp.org/files/Alberts_Experimentation.pdf
2. *Department of Defense Experimentation Guidebook*, Office of the Under Secretary of Defense for Research and Engineering, 2019.
3. *Experiment Planning Guide (EPG)*, Naval Warfare Development Command (NWDC), Revision 2, 1 July 2017.
4. Gallup, S., Schacher, G., & Jensen, J. (2003, April). *Fleet Battle Experiment Final Reconstruction and Analysis Report*. DTIC. Retrieved March 11, 2022, from https://apps.dtic.mil/sti/pdfs/ADA415972.pdf
5. Hillier, W. (2021, August 20). *What is interval data?* CareerFoundry. Retrieved January 10, 2022, from https://careerfoundry.com/en/blog/data-analytics/what-is-interval-data/
6. *Initial capabilities document (ICD)*. AcqNotes. (2021, December 21). Retrieved March 30, 2022, from https://acqnotes.com/acqnote/acquisitions/initial-capabilities-document-icd
7. *Military Utility Assessment (MUA)*. AcqNotes. (2021, August 10). Retrieved March 7, 2022, from https://acqnotes.com/acqnote/acquisitions/military-utility-assessment-mua
8. Meyers, T. J. (2007). *A Risk- and Fuzzy Set-Based Methodology for Advanced Concept Technology Demonstration Military Utility Assessment Design*. ODU Digital Commons. Retrieved March 18, 2022, from https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1101&context=emse_etds
9. Morey, C (2008, Jan). *Constraints, Limitations, and Assumptions Guide*. TRADOC Analysis Center. Retrieved January 3, 2022, from https://orsagouge.pbworks.com/f/CLA+Final.ppt#:~:text=Constraints%2C%20limitations

[7] Silfverskiöld et al., 2021

%2C%20and%20assumptions%20(,the%20study%20team's%20methodology%20development.

10. Navy Lessons Learned Database. Joint Lessons Learned Information System. Retrieved July 6, 2022, from https://www.jllis.mil (CAC-enabled)
11. Oonk, H. M. (2003, March). *Knowledge Desk limited objective experiment (LOE)*. DTIC. Retrieved March 7, 2022, from https://apps.dtic.mil/sti/pdfs/ADA413070.pdf
12. *RAND's standards for high-quality research and analysis*. RAND Corporation. (n.d.). Retrieved February 3, 2022, from https://www.rand.org/about/standards.html
13. Rowe, S. F. & Sikes, S. (2006). *Lessons learned: taking it to the next level*. Paper presented at PMI® Global Congress 2006—North America, Seattle, WA. Newtown Square, PA: Project Management Institute.
14. Silfverskiöld, S., Andersson, K., &; Lundmark, M. (2021, September 16). *Does the method for military utility assessment of future technologies provide utility?* Technology in Society. Retrieved March 17, 2022, from https://www.sciencedirect.com/science/article/pii/S0160791X21002116
15. Silfverskiöld, S., Andersson, K., &; Lundmark, M. (2019, June 26). *The Military Utility Assessment Method for Future Technologies*. DivaPortal. Retrieved March 30, 2022, from https://www.diva-portal.org/smash/get/diva2:1333451/FULLTEXT01.pdf
16. *Technical Report Template*. Kent State University. (n.d.). Retrieved March 4, 2022, from https://www.kent.edu/cae/microsoft-word-document-technical-report-template
17. TRADOC. (2009, August 28). *Initial Capabilities Document (ICD) Writer's Guide*. AcqNotes. Retrieved March 30, 2022, from https://www.acqnotes.com/Attachments/Initial%20Capabilities%20Document%20(ICD)%20Writers%20Guide.pdf

13 Glossary

A

Assessment: Estimation of the nature, quality, or ability of a subject.

C

Constraints, Limitations, and Assumptions (CLA): A framework for both the experiment team and sponsor to understand the conditions under which the experiment was conducted, and the results are applicable.

Critical Questions (CQ): Questions that can accompany and clarify objectives in a DCAP. Optimally 3-4 CQs are listed per objective.

D

Data Collection and Analysis Plan (DCAP): The outline of the experiment, listing the key components and details of the data to be collected.

Demonstration: An activity performed to show how something works, with expected results.

Doctrine, Organization, Training, Materiel, Leadership, Personnel, Facilities, and Policy (DOTMLPF-P): Tool that allows the ability to analyze organizational capabilities for decision-making.

E

Experiment: Structured procedure undertaken to make a discovery or test a hypothesis.

Evaluation: Determination of a subject's merit, worth and significance, using criteria governed by a set of standards.

H

Hypothesis: A specific, tentative prediction, question, or statement about the relationship between two or more variables.

I

Initiative: A stakeholder's attempt to achieve a goal or solve a problem to fill warfighter gaps. Separately, in FLEX events, the initiative is an experiment that has been nominated to or accepted into a formal DoN experimentation process and has been vetted in some way.

M

Measures of Effectiveness (MOE): Factors that gauge purpose accomplishment.
Measures of Performance (MOP): Factors that calculate task completion.
Master Scenario Events List (MSEL): Chronological timeline of scripted actions and events to be inserted into an experiment or exercise.

O

Objective: Goal to accomplish in the experimental procedure.

L

Line of Inquiry (LOI): A set of questions to answer through reading and research.

Q

Quicklook: Format may differ by venue, but is generally a heading-based report with summaries of background, efforts, results, lessons learned, etc.

S

Subject Matter Expert (SME): An authority on a respective area of expertise.

T

Technical Point(s) of Contact (TPOCs): The program managers of the contract, with a focus on meeting contractual requirements.
Tactics, Techniques, and Procedures (TTPs): Behaviors, methods, tools and strategies.

14 Acronyms

C

CLA Constraints, Limitations, and Assumptions
COA Certificate of Analysis
CQ Critical Questions
CSE Council of Science Editors
CVN Carrier, Aircraft, Nuclear

D
DCAP Data Collection and Analysis Plan
DCA&R Data Collection, Analysis & Reporting
DoN Department of the Navy
DOTMLPF-P Doctrine, Organization, Training, Materiel, Leadership, Personnel, Facilities, and Policy

F
FLEX Fleet Experimentation
I
ICD Initial Capabilities Document

M
MOE Measures of Effectiveness
MOP Measures of Performance
MUA Military Utility Assessments

J
JCTD Joint Capability Technical Demonstration

L
Lbs. Pounds
LOI Line of Inquiry
LTE Limited Technology Experiment

P
PDF Portable Document Format
PEO Program Executive Office

R
R&D Research and Development

S
SBIR Small Business Innovation Research
SEC SBIR Experimentation Cell
SME Subject Matter Expert

T
TPOC Technical Point(s) of Contact
TTP Tactics, Techniques, and Procedures
TW Trident Warrior

V
Vs. Versus