

Analysis on the influence factors of operational test effectiveness

Yun Bo¹, Hongbo Bai² and Guojun Wei²

¹Graduate School, Space Engineering University, Beijing, China

²Department of Space Support, Space Engineering University, Beijing, China

yunmails@163.com

Abstract. Operational test is the concrete embodiment of actual combat guidance in the field of weapon equipment test and evaluation. Based on the analysis of the operational test process, this paper summarizes the operational test requirements and components. Based on the interaction between them, 18 factors of 5 categories affecting the effectiveness of operational tests were analysed, provides a reference for promoting the follow-up operational test to be carried out more scientifically.

1. Introduction

The operational test ^[1] relies on the imaginary combat background. According to the combat mission of the tested equipment, the combat personnel control the equipment under test and simulate the execution of relevant combat tasks. Through collecting, collating and analyzing the data generated in this process, the operational effectiveness and applicability of equipment are objectively evaluated. Therefore, operational test is a verification process, and the verification process can be divided into three levels, and each level can be completed through three verification steps. By examining the interaction between the 4 aspects of operational test capability requirements and the 5 components, 18 factors that are not conducive to the effective implementation of operational test can be analyzed.

2. Hierarchical verification process

2.1. Mission level

It is the verification via comprehensive evaluating, and mainly used to confirm whether the weapons and equipment can effectively fulfil the established combat capability after being fielded to complete its combat mission. For example, if the TT observation system were deployed in the land border area, there would be no hiding place for the threat.

2.2. Task level

It is the verification via operational effectiveness and applicability evaluating. It is mainly used to prove that when weapons and equipment perform combat missions derived from combat missions, the completion degree of the mission and the applicability of the equipment to the mission can support weapons and equipment to achieve combat missions. Mission-level verification decomposes combat missions into measurable performance and functional parameters. For example, if an advanced XX system is configured, the threat will be continuously tracked. Continuous tracking here is a measurable parameter in terms of performance or function.



2.3. *Index level*

From the point of view of statistics, it carries out hypothesis testing [2] on whether the parameter values of various aspects of weapons and equipment meet the requirements of the index.

2.4. *Progressive verification steps*

In order to ensure that the conclusions drawn from the operational test truly reflect the inherent effectiveness and applicability of weapons and equipment, the above three levels of verification must be carried out according to the following steps:

Step 1: confirm whether X has occurred;

Step 2: confirm the occurrence of Y;

Step 3: confirm that Y is due to X.

Where, X refers to events related to the function and performance of the tested weapon and equipment, such as the operation of the infrared module of TT; Y is an event related to operational effectiveness or serviceability, such as the discovery of a suspect target.

3. **Operational test capability requirements**

According to the above hierarchical verification process and progressive verification steps of operational tests, it can be concluded that successful operational tests must have four capabilities, namely:

3.1. *Ability to control equipment*

This is the most basic capability that a combat test should have. It requires appropriate operational training before the test, and the operation of the tested equipment in accordance with the specified tactical combat methods during the test.

It corresponds to step one of progressive verification. The direct evidence that the combat test has this capability is that the tested equipment operates normally in the combat test. The risk that is not conducive to the operational test proving the capability is that the equipment is not used in accordance with tactical methods or the equipment is not in operation.

3.2. *Ability to measure parameters*

In other words, the combat test should be able to measure the difference of combat effects according to the different application scenarios of weapons and equipment in a quantitative way.

It corresponds to step two of the progressive verification. The direct evidence that the combat test has this capability is that when the combat configuration of weapons and equipment changes, the quantified combat effect also changes. The risk that is not conducive to proving that a combat test has this capability is that the combat test introduces so much interference information that it is difficult to judge whether the combat effect has changed or not.

3.3. *Ability to judge cause and effect*

That is, the operational test should be carried out through the experimental design and task, so that the causes of operational effects can be separated. In other words, it can be proved through the test that the tested weapon and equipment is the reason for the combat effectiveness.

It corresponds to step three of the progressive verification. The direct evidence that the combat test has this capability is that the data analysis can infer that the tested weapon and equipment is the only cause of the combat effectiveness. The risk in favor of proving that a combat test has this capability is an alternative explanation for the operational effectiveness.

3.4. *Ability to promote results*

This is the core capability that combat test should have, that is, the results of combat test should not only apply to the specific weapons and equipment of the subject, but also be able to be extended to other weapons and equipment of the same type produced in large quantities.

The first three aspects of capability are necessary for the combat test as a complete verification process. The ability of the fourth aspect is the foundation of operational test with practical significance. Only when operational test has this ability, can the test conclusion be used as a reference for the decision of weapon and equipment assembly. Thus, this fourth capability can be viewed as an external requirement for conducting operational tests.

The direct evidence that the combat test has this capability is that the tested equipment is representative of production, that is, it can reflect the average combat capability after mass production of this type of weapon and equipment, and the process of combat test is highly similar to combat. The risk that is not conducive to proving the capability of the combat test is that the parameters of the tested equipment are not solidified, or the operational mission simulated by the combat test is not coherent enough, or is interfered by too many non-operational factors.

4. Components of operational tests

Compared with DoE^[2] (Design of Experiment) based on statistics, the data generation process of combat test can be divided into five components.

4.1. Treatment -- equipment of the subject

Treatments in experimental design refer to the factors applied or observed by the researcher according to the research purpose, which can act on the subject and cause direct or indirect effects. In operational tests, treatments are the equipment. It is the "independent variable" of the process of generating data.

4.2. Subject -- combat unit

The subject in experimental design refers to the basic unit^[3] that receives treatment and serves as the object of observation. In a combat test, the subject is the smallest combat force operating weapons and equipment. For example, the operator operating the sensor is targeted at the subject of component-level weapons and equipment, the special combat personnel equipped with individual combat system are the subjects of system level weapons and equipment, and the new armored infantry battalion is the subject of system-level weaponry.

4.3. Effect -- parameter indicators

Effect^[4] in experimental design refers to the reaction or result of treatment factors on the subject. In operational tests, effects are usually expressed in the form of various parameters. It is the "dependent variable" of the process of generating data.

4.4. Non-treatment factors -- test conditions

Non-treatment factors^[4] in experimental design refer to other factors that influence the effect besides treatment factors. In the operational test, the non-processing factors are the background conditions under which the combat unit manipulates the tested equipment to simulate the execution of the combat mission, which usually include the operational considerations, terrain and geomorphology, weather, electromagnetic environment, simulated threats and other factors.

4.5. Data analysis -- data analysis

It refers to the three steps of the above verification and the realization of the three-level verification process through the measured effect. In operational test, data analysis usually needs to refer to the given target threshold or compare with the baseline combat force^[5] to investigate the effect of the equipment under test.

5. Influencing factors related to the ability requirements of operating equipment

It is the logical starting point for the research of weapon and equipment through combat test to be able to control the weapon and equipment of the subject and produce combat effect.

5.1. The equipment to be tested cannot perform its functions normally

It mainly refers to the fact that a certain function of the equipment to be tested cannot operate normally and relevant data cannot be collected for the test, thus it is difficult to organize subsequent verification, evaluation and prediction procedures.

5.2. The combat personnel's skill level of operating equipment does not meet the operational requirements.

It mainly refers to the failure to organize the equipment operation training before the test according to the operational requirements and achieves the required skill level.

5.3. Measurement method is insensitive to response.

It mainly refers to the adoption of a certain parameter indicator, but the existing measurement methods are difficult to distinguish the response with similar level of this indicator. That is, the ability of operational tests to identify changes in results under these conditions is very limited.

5.4. There is no opportunity to implement new ability.

It mainly refers to that under certain test conditions, the simulated operation may not have the opportunity to implement a certain function under test, so as to collect relevant data, which makes it difficult to organize the subsequent verification, evaluation and prediction procedures.

6. Influencing factors related to ability requirements of measurement parameters

If the weapon and equipment of the subject can be manipulated to produce combat effect, the measurement method should be configured to measure the effect and statistical method should be used to determine whether the result has statistical significance.

6.1. The functions of multiple devices (of the same type) of the subjects were different

It mainly refers to the fact that the test samples [2] contain multiple weapons and equipment of the same type, and the functional indexes of these weapons and equipment are obviously different. In terms of data, there is a large variance, and the effectiveness of experimental results is limited.

6.2. Inconsistent proficiency of combat personnel

It mainly refers to the fact that although the warfighters have been trained, their control level of weapons and equipment is obviously different, which makes the results of the test very noisy. That is to say, it is difficult to judge whether the result is good or bad because of the characteristics of weapons and equipment, or because of the personnel's operational level, so the effectiveness of the test results is limited.

6.3. The correctness of data collection is inconsistent

It mainly refers to the error rate of the measurement method aimed at the parameter index in the implementation process, which makes the test data deviate from the truth value. That is to say, the test has non-sampling error, which makes the reliability of the test results lower.

6.4. Drift exists in the test conditions

When the subject of the test is repeated for many times, the level of some factors of the test conditions changes, which also introduces noise to the test results, and the harm is the same as (B).

6.5. Test power is too low

It mainly refers to that the data analysis and design of the experiment did not fully consider the pre-test information and did not set the experimental conditions with statistical representativeness in the subjects, so as to make efficient use of the experimental resources and draw the correct conclusions of the experiment.

6.6. *Violating the statistical hypothesis*

It mainly refers to the fact that some statistical technology was used in the data analysis of the test, but the actual data did not conform to the hypothesis and premise of the statistical technology, so the technology was wrongly used, which reduced the reliability of the test results.

7. **Influencing factors related to the ability to judge cause and effect**

If the first two capabilities can produce combat effects, and the effects can be effectively measured, then it is necessary to explain the reasons for the effects, and focus on determining whether the expected combat effects are generated by the use of the tested equipment.

7.1. *The performance of the equipment in different subjects was different*

This means that multiple test subjects of the test examine the same function, but the combat effectiveness of manipulating the function will become better or worse over time. In this case, the effect of the sequence of trials would prevent other explanations.

7.2. *In each simulated operation, the proficiency of the combat personnel is inconsistent*

It mainly refers to that the combat personnel's proficiency in handling weapons and equipment will increase with the time of participation. In this case, the warfighter's learning effect on weapon handling hinders other explanations. Or the operational test is organized as a comparison test between the tested equipment and the baseline equipment, but the operational control level of the two groups of equipment is different. In this case, the difference in personnel proficiency also hinders the explanation of other causes.

7.3. *In each simulation operation, the correctness of the collected data is inconsistent*

It mainly refers to that the data collection personnel have participated in several test subjects, and the accuracy and accuracy of the data collection may also be higher due to the more skilled collection work in the later stage, resulting in the inconsistency in the degree of deviation from the truth value and the difference in the degree of dispersion of the data at different time points. Thus, the interpretation of the results of combat tests is reduced. Or in the comparison test between the tested equipment and the baseline equipment, the correctness of data collection by the data acquisition personnel is inconsistent, which hinders the interpretation of other causes.

7.4. *With the implementation of the simulated combat task, the test conditions changed*

The mainly refers to that over time, the weather conditions of simulated missions, the combat capability of the blue army, and other test conditions may become better or worse, making it difficult to explain the cause of the effect. In addition, if there is a difference between the above test conditions of the tested equipment and the baseline equipment in the comparison test of combat test, it will also be difficult to explain the cause of the effect.

8. **Influencing factors related to the ability to promote results**

When it is confirmed that the effect of simulated combat is indeed generated by the tested equipment, the most important question is whether the equipment's verified function can be used in future combat practice.

8.1. *The tested function cannot represent the combat mission*

It mainly means that the function of the subject only reflects part of the combat mission, or cannot reflect the actual situation of future operations, or even the function of investigation may not appear in future operations.

8.2. *Combat personnel are not representative of combat units*

Basically, the training level of the warfighter may not be up to the required intensity, or the skill level of the tested warfighter may be lower or significantly higher than that of the warfighter who will use the equipment in the future.

8.3. *Measurement scales do not reflect important effects*

This mainly refers to the fact that the measurement method of the experiment is not high resolution, unable to identify the difference in response of similar level, or excessive qualitative measurement method is adopted. As a result, the measured effects are of limited representativeness to combat effects.

8.4. *The authenticity of operational scenarios is limited*

It mainly refers to that the operational plan does not reflect the actual situation of the battlefield, for example, the warfighters are very familiar with the operational scene, the operational action of the blue team is different from the actual situation, and the operational configuration and actual situation are different.

References

- [1] DoD, Test and Evaluation Management Guide, 6th ed., Washinton D. C., 2013.
- [2] Gravetter, F.J., Wallnau, L.B., Essentials of Statistics for the Behavioral Sciences, 8th ed. Cengage Learning Asia Pte. Ltd, Singapore, 2014.
- [3] Azarian, M., Keimasi, M., and Pecht, M., "Non-Destructive Techniques for Detection of Defects in Multilayer Ceramic Capacitors," Components for Military and Space Electronics Conference, Los Angeles, CA. February 2006, pp. 125-130.
- [4] Boydston, A., and Lewis, W., "Qualification and Reliability of Complex Electronic Rotorcraft Systems," the AHS Specialists Meeting on Systems Engineering, Harbor, CT., October 2009, pp. 15-16.
- [5] Briand, L.C., Wuest, J., Ikononovski, S., and Lounis, H., "Investigating quality factors in object-oriented designs: An industrial case study," International Conference on Software Engineering, Aug 1999, pp. 345-354.