# Ergonomic Evaluation Index System for Fighter Planes Cockpit Touch Screens

*Fei Wang [1], Wenzhe Tang [1], Zhiman Zhu[1], Ruoyu Hu[1], Xiuhang Xu[1], Chengqi Xue[1,*]*

*[1] School of Mechanical Engineering, Southeast University*

*Nanjing, 211189, China*

## ABSTRACT

The purpose of this paper is to establish and verify metrics to evaluate the usability of the cockpit touch screen, which includes operational performance, physiological performance, and the control-display resources of the touch screen and so on. In this study, 12 experts who specialized in the area of aviation efficiency evaluation were investigated to verify the effectiveness and reliability of the metrics framework from the aspects of the necessity, feasibility, identification of a single index, and the integrity and redundancy overlap of the overall metrics system. The results indicated that the metrics system includes a high reliability and validity. Overall, we conclude three first-level metrics including perception layer, interaction layer and impression layer, as well as seven second-level metrics and 42 third-level metrics. It can find the problem of design parameters during the cockpit touchscreen design, promote to improve the operational performance and survival probability of pilots.

**Keywords**: Cockpit, Touch Screen, Efficiency, Evaluation Indicators, Expert Investigation

## INTRODUCTION

The touch screen is an intuitive interactive method that can flexibly display other information related to the current task, so it has great potential in reducing the user's cognitive load and improving the user's situational awareness [1]. At the same time, touch screens are also faced with shortcomings such as accidental activation [2] and reduced visibility caused by environmental impacts in the application of aircraft cockpits.

The evaluation of the safety, performance and comfort of the touch screen interface is called the usability study of the touch screen. The study of usability originated in the field of software engineering in the 1970s, but was later widely applied to other fields, which led to the diversification of usability research [3]. Representative theories include the usability definition of ISO9241-11, the five dimensions proposed by Nielson, SUMI usability model, etc. With Sung H. Han creatively displays the subjective factors of industrial products through multiple dimensions [4], and takes the subjective experience of users into consideration. One of the representative ones is the MUG usability guide proposed by Microsoft.

Since the 20th century, many scholars have proposed corresponding ergonomic evaluation indicators for the application of touch screens in industrial equipment. Nout C. M. van Zon [5] uses Fitz's law to establish a touch accuracy and throughput model for navigation tasks based on touch interaction, and provides a quantitative evaluation method for accuracy and throughput. Neville A. Stanton [6] took four input devices (trackball, rotation controller, touch pad and touch screen) of the flight management system as the research object, and proposed evaluation indicators including time to complete a specific task, error rate, workload, and subjective Usability and physical comfort. Sung H. Han [7] proposed that usability includes the following two-dimension groups: objective performance and subjective impression, and finally established and defined 48 detailed usability dimensions. Qingyuan Bai et al. [8] systematically analyzed the ergonomic elements of the display and control system based on typical flight missions and specific requirements, and conducted research around 4 first-level indicators (general layout, display, control, and alarm system).

According to previous studies, the evaluation of the usability of the touch screen in the special environment of the aircraft cockpit is still lacking. Secondly, the usability theory system in the traditional software field is abstract and vague, and cannot be directly applied to the ergonomic evaluation indicators of different scenarios. It needs to be divided into different factors and evaluation indicators in detail to form a multi-level and multi-indicator Comprehensive evaluation system [9]. Based on the above situation, this research starts from theoretical analysis, sorts out the national standards and industry standards related to aircraft cockpit touch interaction, and combines the influencing factors of flight missions to construct a usability evaluation index system suitable for aircraft cockpit touch screens. And the validity of the evaluation model is verified, so as to provide an effective basis for design optimization in the design phase of the aircraft cockpit touch screen to help the crew complete the task in a healthy, safe, comfortable and efficient manner.

# THE ESTABLISHMENT OF ERGONOMIC EVALUATION INDEX SYSTEM FOR FLIGHTER PLANES COCKPIT TOUCH SCREENS

## Determination of Evaluation Index Set

In order to ensure the effectiveness and comprehensiveness of the index system, the determination of the ergonomics evaluation index for the aircraft cockpit touch screen interface needs to be based on the traditional ergonomics evaluation index and the predecessor's research on the ergonomics evaluation index in other fields. Specifications or guidance documents are used as a basic reference, and the pilot's actual operating experience is investigated. In an interview with a certain F-35A fighter pilot by Hush-Kit, the pilot of the F35 fighter explained that once he encounters turbulence or is performing tactical maneuvers during the flight, the false touch rate is as high as about 20%.

Touch screens are widely used in civil aircraft and fighter aircraft cockpits. There are industry norms and standards for the display of touch interactive software systems, touch screen hardware, and human factors as references. This article sorts out national standards, industry standards, and FAA advisory notices and other documents Regarding regulations related to aircraft cockpit touch screens, such as: MIL-STD-1472G, SAE ARP 60494 and other documents. Based on the combing of industry norms and standards, ergonomic evaluation indicators are extracted. At the same time, these specification documents can provide design improvement suggestions for comparison. Since each "sub-indicator" corresponds to multiple "standards and design criteria", it is possible to give targeted modification suggestions based on the evaluation results of a certain indicator. In the process of extracting indicators, follow the following principles: 1. Measurability: For regulations that can be quantified, try to extract relevant quantitative indicators. 2. Comprehensiveness: extract all indicators related to this design criterion as much as possible. 3. Reduce the repetition rate: The "standards and design principles" in different documents may have repetitive meaning, and they can be equivalent to the same index. Based on the above three principles, all ergonomic evaluation indicators are extracted from different documents and regulations to form an evaluation indicator set.

## Preliminary establishment of evaluation index system

Based on the ergonomic evaluation indicators extracted above, this article classifies the indicators and finally forms an indicator system. Studies have shown that for the establishment of usability indicators for different types of product interaction interfaces, the common method is to first take the internationally-used dimensions as the first-level dimensions, and then define the second- and third-level dimensions according to the specific application conditions to ensure the comprehensiveness of the evaluation indicators. Other classification methods such as Sung H. Han [7] divide the usability index system of consumer electronic products into Performace and impression dimesions. Sun Jianhua, Jiang Ting and others [10] divided the ergonomics evaluation index of the spacecraft software man-machine interface into three dimensions: the interface layer, the operation layer, and the demand layer. Sung H. Han and others have achieved considerable results in the usability index system. Taking consumer electronics as an example, they proposed a systematic usability evaluation system. He believes that consumer electronics is a system composed of hardware and software. Since the research object of this paper is the aircraft cockpit touch screen, it is also a system composed of hardware (touch screen) and software (interface). Based on the research of Sung H. Han and predecessors on the ergonomics evaluation index system, this paper finally established a three-level evaluation index model suitable for fighter aircraft. The first-level indicators of the framework include cognitive layer, interaction layer and impression layer. The cognitive layer focuses on the process from the pilot receiving information to completing the action, including perception/perception, learning/memory, and reaction/movement; the interactive layer includes display attributes, control attributes, and display control attributes, while the impression layer focuses on the pilot’s subjective use after operation Experience. The following details the 48 three-level indicators and their definitions of the evaluation index system.

The perception dimension refers to the pilot's intuitive perception, cognition, and operational performance of the touch screen at the physiological cognition level. The following Table1. lists 14 three-level indicators of perception dimensions.

Table 1: 14 three-level indicators of perception dimensions

| **Top layer** | **Middle layer** | **Bottom layer** | **Number** |
|---|---|---|---|
| Perception Layer | Perception/ Cognition | Direct | X1 |
| | | Clear | X2 |
| | | Simple | X3 |
| | | Discernible | X4 |
| | | Warning | X5 |
| | Memorization/ Learning | Familiar | X6 |
| | | Readable | X7 |
| | | Accessible | X8 |
| | | Understandable | X9 |
| | | Sensitive | X10 |

| | | | |
|---|---|---|---|
| | | Predicted | X11 |
| | | Memorable | X12 |
| | Action/Response | Smooth | X13 |
| | | False touch rate | X14 |

The interaction dimension refers to the factors that influence the pilot's operational ergonomics at the display layer, control layer, and display control integrated level. Table 2. lists 23 three-level indicators of interaction dimensions.

Table 2: 23 three-level indicators of interaction dimensions.

| **Top layer** | **Middle layer** | **Bottom layer** | **Number** |
|---|---|---|---|
| Interaction Layer | Display Attributes | Text | X15 |
| | | Image | X16 |
| | | Size | X17 |
| | | Brightness | X18 |
| | | Color | X19 |
| | | Contrast | X20 |
| | | Dynamic | X21 |
| | | Feedback | X22 |
| | | Operation confirmation | X23 |
| | | Warning information | X24 |
| | | Information quality | X25 |
| | | Information priority | X26 |
| | | Layout | X27 |
| | | Display consistency | X28 |
| | Control Attributes | Gesture | X29 |
| | | Reachable | X30 |
| | | Operation consistency | X31 |
| | | Control display ratio | X32 |
| Interaction Layer | Display-Control Attributes | Visual occlusion | X33 |
| | | Reset | X34 |
| | | Error recovery | X35 |
| | | Task matching | X36 |
| | | D-C consistency | X37 |

The impression dimension refers to the subjective overall experience after the pilot uses the touch screen to operate. The following Table 3. lists 11 three-level indicators of impression dimensions.

Table 3: 11 three-level indicators of impression dimensions.

| Top layer | Middle layer | Bottom layer | Number |
|---|---|---|---|
| Impression Layer | Overall Evaluation | Reliability | X38 |
| | | Effort level | X39 |
| | | Efficient | X40 |
| | | Useful | X41 |
| | | Learn | X42 |
| | | Convenient | X43 |
| | | Comfort | X44 |
| | | Efficiency | X45 |
| | | Personalized | X46 |
| | | Attractive | X47 |
| | | Satisfaction | X48 |

Therefore, this research has initially determined the evaluation index system of the aircraft cockpit touch screen to summarize its description. These ergonomic evaluation indexes are divided into three first-level indexes, the cognitive layer, the interaction layer and the impression layer. The cognitive layer includes sensory There are three secondary indicators of cognition and reaction. The cognitive layer includes three secondary indicators of display attributes, control attributes, and display control attributes. The impression layer includes one secondary indicator of use and feeling. In the end, a total of 48 three-level indicators have been determined.

## RELIABLITY VERIFICATION OF EVALUATION INDEX SYSTEM

The reliability verification of the ergonomic evaluation index system for the touch screen interface of the aircraft cockpit is mainly to check whether a single index is necessary and whether the overall index system is effective. Aiming at the 48 sub-ergonomic indexes extracted in the previous research work and the three-level evaluation index system, a single test and an overall test are carried out to verify the rationality of the index system.

The purpose of the single test is to verify the necessity, recognition, and overlap of a single index, and the purpose of the overall test is to verify the comprehensive redundancy and completeness of the index system. Necessity is the degree of necessity to test a single indicator, and recognition refers to the ability and effect of each indicator to distinguish certain aspects of the evaluation object's characteristics, value. The degree of overlap refers to whether there is overlap in content between the evaluation factors of the test index system, or whether the degree of overlap is within an acceptable range. Comprehensive redundancy is a measure of the redundancy of the index system, and completeness is whether the index system covers the characteristics of the evaluation object more comprehensively. Twelve aerospace ergonomics experts were invited to score individual indicators and indicator systems.

For a single test, the necessary degree coefficient is dispersion/concentration. A smaller dispersion value indicates that the opinions are more concentrated. Generally, if the dispersion value is less than 0.63, it indicates that the dispersion meets the requirements; greater concentration The value indicates the higher the importance of this indicator. Set the concentration to 2 to reach the general level, and limit the necessity coefficient to be less than 0.63/2=0.315. The experimental results show that the necessary degree coefficients of several indicators such as familiarity (X6), size (X17), color (X19), dynamic elements (X21), personalization (X46), and attractiveness (X47) are greater than the critical value of 0.315. Other indicators are less than 0.315. Moreover, the recognition coefficient of a single index is greater than 3, indicating that the bottom single evaluation index has a high degree of recognition. In terms of overlap, the interval of the overlap between the index and another index is [0,1], if the overlap is 0, there is no overlap at all, and if the overlap is 1, it means that the two indicators are completely overlapped. The overlap coefficients of each indicator and other indicators are all less than 0.5, and there is no serious overlap. For the overall test, the comprehensive overlap of the indicator system is the sum of the overlap coefficients of all indicators. Set the strict comprehensive overlap coefficient to 0.1, and limit the comprehensive overlap coefficient of the index system to be less than the number of underlying indicators multiplied by the strict comprehensive overlap coefficient. The results show that the comprehensive overlap degree is 3.5362, which is less than 4.8, and the structural redundancy of the comprehensive evaluation index system is at a low level, indicating that the requirements are met and the evaluation structure of the evaluation system has a high degree of independence. Finally, the entire evaluation index system is presented to the experts, and the experts are invited to evaluate the integrity of the index system. Experts score the integrity of the indicator system (1, 2, 3, 4, and 5 represent very incomplete, incomplete, general, complete, and very complete, respectively). The results show that the average value of the five-point scale is 4.667, indicating the indicator The system has high integrity.

After the experiment, the experts were interviewed about the necessity of indicators. The experts believed that the pilots would be fully trained to ensure the familiarity of the operation before using the cockpit touch screen, so the necessity of the familiarity (X6) is not high. In addition, the size (X17) and color (X19) of the aircraft display did not have enough opportunities for design improvement, and the dynamic elements (X21) were also very limited, so they were eliminated. Personalization (X46) and attractiveness (X47) are not suitable for professionally trained pilots because they may have formed professional operating habits and aesthetics, so they need to be eliminated.

# SUMMARY

This article discusses the ergonomics evaluation system for touch screen interfaces in the special environment of the aircraft cockpit. The purpose is to apply the interface usability evaluation theory to the ergonomics evaluation of the touch screen in the complex environment of the aircraft cockpit, expanding the traditional interface usability evaluation field. Using theoretical research on usability, a three-level aircraft cockpit touch screen ergonomic evaluation index system was established, with a total of 42 individual bottom-level indicators, and the reliability and effectiveness of the index system were verified from the individual indicators and the overall system. Under the guidance of usability indicators, designers can more clearly clarify the aspects of aircraft cockpit touch screen operations that need to be improved. At the same time, the indicator system can also provide guidance for interface optimization in related complex environments.

# REFERENCES

1. Rogers W A., Fisk A D., McLaughlin A C., et al. Touch a screen or turn a knob: Choosing the best device for the job. vol. 47(2), pp. 271-288. J. Human Factors. (2005)
2. Cockburn A., Masson D., Gutwin C., et al. Design and Evaluation of Braced Touch for Touchscreen Input Stabilisation. vol. 122, pp. 21-37. J. International Journal of Human-Computer Studies. (2018)
3. Qingfei M., Ju W., Zhenhua L. Intergrated research framework for moblie system usability. vol. 29(2), pp. 421-426. J. Application Research of Computers. (2012)
4. Han S H., Yun M H., Kwahk J., et al. Usability of consumer electronic products. vol. 28(3-4), pp. 143-151. J. International Journal of Industrial Ergonomics. (2001)
5. Zon N C M V., Borst C., Pool D M., et al. Touchscreens for Aircraft Navigation Tasks: Comparing Accuracy and Throughput of Three Flight Deck Interfaces Using Fitts' Law.vol. 62, pp. 897-908. J. Human Factors and Ergonomics Society. (2020)
6. Stanton N A., Harvey C., Plant K L., et al. To twist, roll, stroke or poke? A study of input devices for menu navigation in the cockpit. vol. 56(4), pp. 590-611. J. Ergonomics. (2013)
7. Sung H. Han., Myung Hwan Yun., Kwang-Jae Kim., Jiyoung Kwahk. Evaluation of product usability: development and validation of usability dimensions and design elements based on empirical models. vol. 26, pp. 477-488. J. International Journal of Industrial Ergonomics. (2000)
8. Bai Q., Bai Y., Wang X., Zhao X., Yu J. Ergonomics Index System of Airplane Cockpit Display and Control Resources. vol. 903, pp 827-832. J. Intelligent Human Systems Integration. (2019)
9. Mo C., Chengqi X., Haiyan W., et al. The Analysis of Virtual Product Based on the Usability. vol. 19(03), pp. 135-140. J. Industrial Engineering and Management. (2014)

10. Jianhua S., Ting J., et al. Research on Ergonomic Evaluation Indicator and Methodology for Interactive Interface of Spacecraft Software. vol. 26(2), pp.208-213. J. Manned Spaceflight. (2020)