Development of Usability Requirements at NASA

Anikó Sándor, Johnson Space Center
Kritina L. Holden, Johnson Space Center

The International Organization for Standardization [ISO] published a number of standards for the operational evaluation of usability. The ISO 9241-11 standard defines usability as "the extent to which a product can be used by specified users to achieve specified goals," and recommends evaluating usability in terms of measures of effectiveness, efficiency, and satisfaction. Measures of effectiveness relate the goals or sub-goals of the user to the accuracy and completeness with which these goals can be achieved. Measures of efficiency relate the level of effectiveness achieved to the expenditure of resources. Finally, satisfaction measures the extent to which users are free from discomfort as well as their attitudes toward the use of the system. All three factors are important in evaluating the usability of a system.

Usability is a key element of the human-centered design approach. Human-centered design is a design philosophy and a process that takes into account human capabilities and limitations at each stage of the design process. Applying a human-centered design process to system development contributes to crew health and safety by increasing system usability. Insufficient integration of human concerns with the vehicle design may result in inadequate interfaces and deficiencies in commonality, consistency, and usability that translate into less-than-optimal operations, higher training costs, and an increased risk to mission objectives. A usable system will provide increased effectiveness, efficiency, and satisfaction. Furthermore, good usability reduces errors, training time, and overall life cycle costs, and it is essential to ensure crew safety and mission success.

The Need for a Usability Requirement at NASA

At NASA, the development of space vehicles and related hardware and software are driven by verifiable requirements. In spite of the need for requirements and the recognized importance of usability, it has taken many years to add a usability requirement to NASA program documentation. First, writing a usability requirement is difficult because the requirement must be easy to operationalize. For example, how can usability be measured and, therefore, verified? Furthermore, can usability be represented by a single, simple metric? Second, the verification methodology needs to be achievable with the time and cost limitations typically encountered in the space industry. For example, the number of interfaces tested, number of test sessions, and number of participants tested must be limited to maintain reasonable cost. Third, the requirement must be broadly applicable to system, subsystem, and component levels. Finally, the requirement should encourage a process approach. Good usability cannot be accomplished with a one-time evaluation but only with iterative usability evaluations from the beginning to the end of the life cycle of a system.

General Practices in the Area of Usability

NASA reviewed the documentation of a number of companies and agencies such as the Department of Defense and the Federal Aviation Administration to understand their approach to usability. NASA found that the common practice is to follow human factors guidelines and standards, such as the Military Standards, MIL-STD-1472; include human factors and usability professionals in the system development life cycle; and evaluate designs through expert evaluations and user testing. However, NASA did not find any practice among those agencies that includes objective, verifiable usability requirements and criterion of verification. Furthermore, industry often has multiple criteria for its products. For example, a voice recognition system should have a 98% accuracy rate to be usable, but a cell phone may have 85% success rates on expert tasks, and 95% on common tasks. The space agency concluded that these practices and approaches were not appropriate for the safety-critical space industry.

Development of the Usability Requirement at NASA

In developing a usability requirement for NASA, the metric of choice was errors. Obviously, there are many other metrics for usability. However, errors are objective, easy to understand by all stakeholders, and highly related to all factors of usability. Virtually all tasks in the space domain are driven by procedures, and NASA's approach was to put the emphasis on identification of usability problems by calculating error rates per procedure ***step*** and

Table 1. *An Example Scenario Illustrating the Error Count on a Procedure with 10 Steps Completed by a Group of 10 Participants*

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6	Step 7	Step 8	Step 9	Step 10
Participant 1	0	1	0	0	0	0	0	1	0	0
Participant 2	0	1	0	0	0	0	0	0	1	0
Participant 3	0	1	0	0	1	0	0	0	0	0
Participant 4	0	1	0	1	0	0	0	0	0	0
Participant 5	1	1	0	0	0	0	0	0	0	0
Participant 6	0	1	0	0	0	0	1	0	0	1
Participant 7	0	1	0	0	0	0	0	0	0	0
Participant 8	0	0	0	0	0	0	0	0	0	0
Participant 9	0	1	1	1	1	1	1	1	1	1
Participant 10	0	1	0	0	0	0	0	0	0	0

per ***participant***, rather than calculating an overall error rate. Furthermore, NASA focused on errors that were design-induced to ensure the space agency captured errors that are design-related, and not necessarily caused by other factors, such as the variability of the human. Because NASA wanted to reduce the complexity and subjectivity of the method, it decided not to focus on the severity of the errors. Early experiences during development of the usability requirement showed that using severity ratings increases the subjectivity of the decisions, and there are already many other documented requirements that mitigate severe errors.

In a usability test it is important to detect steps with many errors. High error rates on a step suggest usability issues related to the execution of that step (e.g., design problem with the hardware, software, or instructions). On the other hand, it is equally important to detect participants who commit many errors. These participants may be outliers due to things such as different training background. Table 1 shows an example of the error counts for a sample of 10 participants executing a procedure composed of 10 steps. Note that on Step 2, most participants committed errors, and Participant 9 committed errors on most of the steps. This is a situation in which Step 2 merits human factors attention and potential design changes. Participant 9's background and test circumstances should be investigated further to identify any differences that resulted in the high error rate.

In developing the usability requirement, NASA took into account both of these aspects, and thus developed criteria for both (5% and 10%, respectively), based on testing and computer simulations. As a result, the requirement ensures that the number of design-induced errors is minimized for every task step and every participant. With proper error definition, analyses indicate that this type of requirement is very stringent for the types of tasks anticipated for spacecraft verification. The final approved requirement wording documented in the Human-Systems Integration Requirements document is as follows:

"The system shall provide crew interfaces that result in a maximum of 5% erroneous task steps per participant, where each erroneous task step is committed by 10% or fewer participants."

Achieving a community-approved usability requirement in a NASA document is a significant accomplishment because this indicates a first step toward full inclusion of usability in the NASA system development life cycle. NASA is at the forefront of evolving the usability practice by having a verifiable usability requirement using objective measures for usability verification.