Technical Guidelines Development Committee
22 March 2007 Plenary Meeting

Human Factors and Privacy:
Progress Report

Technical Guidelines Development Committee
22 March 2007 Plenary Meeting

Overview

- Changes and issues in the VVSG HFP section
- Issues requiring further analysis
- Usability benchmark development progress
- Next research steps

Technical Guidelines Development Committee
22 March 2007 Plenary Meeting

**There are three significant changes
from the December meeting.**

- I will refer to the requirements using the Chapter 12 numbering from the version you have in your binders.

Require availability of choice of font size and contrast on all VEBD-V machines, not just the accessible-VS.

- **12.2.4-E** Available Font Sizes

  A voting station that uses an electronic image display shall be capable of showing all information in at least two font sizes, (a) 3.0-4.0 mm and (b) 6.3-9.0 mm, under control of the voter. The system shall allow the voter to adjust font size throughout the voting session while preserving the current ballot choices.

- **12.2.4-J** High Contrast for Electronic Displays
  The voting station shall be capable of showing all information in high contrast either by default or under the control of the voter. The system shall allow the voter to adjust contrast throughout the voting session while preserving the current ballot choices. High contrast is a figure-to-ground ambient contrast ratio for text and informational graphics of at least 6:1.

➔ removal of the reqs from the Acc-VS as redundant

➔ "throughout the voting session"?

4

Technical Guidelines Development Committee
22 March 2007 Plenary Meeting

General adjustability throughout voting session

- A number of requirements have clarified that, in general, when voter can control or adjust some aspect of the voting station, that can be done throughout the voting session without loss of information.   New:

- **12.3.2-F** Synchronized Audio and Video
  The voting station shall provide synchronized audio output to convey the same information as that which is displayed on the screen. There shall be a means by which the voter can disable either the audio or video output, resulting in a video-only or audio-only presentation, respectively. **The system shall allow the voter to switch among the three modes (synchronized audio/video, video-only, or audio-only) throughout the voting session while preserving the current ballot choices.**

- **12.2.6-A.1** Voter Control of Language
  The system shall allow the voter to select among the available languages throughout the voting session while preserving the current ballot choices.

Technical Guidelines Development Committee
22 March 2007 Plenary Meeting

The safety requirement now refers to UL 60950 (actual technical specs), not OSHA regulation.

- **12.2.8.3-A** Compliance with Federal Regulations –
  All equipment associated with the voting system shall be certified in accordance with the requirements of UL 60950, Safety of Information Technology Equipment by a certification organization accredited by the Department of Labor, Occupational Safety and Health Administration's Nationally Recognized Testing Laboratory program. The certification organization's scope of accreditation shall include UL 60950.

Technical Guidelines Development Committee
22 March 2007 Plenary Meeting

There are issues that require further analysis.

Vendor Usability Testing with Voters: Specializing the CIF

- **12.2.1.2-A** Usability Testing by Vendor

➔ Required to report in the ISO/IEC_25062:2006 Common Industry Format for Usability Test Reports: to tailor specifically to voting systems, see also:

- **12.2.6-A.3** Usability Testing for Alternative Languages
- **12.2.8.1-B** Usability Testing by Vendor for poll workers
- **12.3.2-A** Usability Testing by Vendor for voters with partial vision
- **12.3.3-A** Usability Testing by Vendor for voters who are blind
- **12.3.4-A** Usability Testing by Vendor for voters who lack fine motor skills

- More information: http://www.nist.gov/iusr

What is the CIF?

- **C**ommon **I**ndustry **F**ormat for Usability Test Reports
- Details not WHAT to test but how to report it
- Focus on SUMMATIVE usability test
- Purpose is to have a common format to make it easier for consumer organizations to review and compare results

- CIF became an ANSI standard in December, 2001 (ANSI/NCITS 354-2001) and an ISO Standard (ISO/IEC 25062:2006), May 2005

5.3.2 Test Objectives

The following information shall be provided:

a) The objectives for the test and any areas of specific interest.

    NOTE: Possible objectives include testing user performance of work ta: using the product.

b) Functions and components with which the user directly and indire

The following information should be provided:

a) Reason for focusing on a product subset, if the whole product wa

5.4 Method

Sufficient information shall be provided to allow an independent tester to testing.

5.4.1 Participants

The following information shall be provided:

a) The total number of participants tested.

    NOTE: In order to generate valid summative statistical analyses, it is no

b

5.4.4 Usability Metrics

As defined in Section 4.1, usability is measured by three types of metrics: effectiveness, efficiency, and satisfaction.

The following information shall be provided:

a) Metrics for effectiveness.

b) Metrics for efficiency.

c) Metrics for satisfaction.

Effectiveness and efficiency results shall be reported, even when they are difficult to interpret within the specified context of use. In this case, the report shall specify why the supplier does not consider the metrics meaningful.

The CIF Requires A Description of:

- Intended and actual users, tasks, and environments
- Simulated working conditions
  - Don't think aloud
  - Assistance only if normally available
- At least 8 participants
- Objective measures of effectiveness, efficiency, (satisfaction data)
  - Measures for effectiveness may include:
    1. *Completion Rate*
    2. *Number of Errors*
    3. *Number of Assists*
  - Measures for efficiency may include:
    1. *Task time*
    2. *Completion Rate/Mean Time-On-Task*

Technical Guidelines Development Committee
22 March 2007 Plenary Meeting

Performance Metrics

- **12.2.1.1** Overall Performance Metrics
  The requirements of this section set benchmarks for the usability of the voting session.

➔ Will discuss status in detail later in this presentation

End-to-end accessibility evaluation

- Glossary definition: **end-to-end:** (1) (Security) Supporting both voter verification and election verification. (2) (Generically) Covering the entire elections process, from election definition through the reporting of final results

- Accessibility testing of components and usability testing with voters of the Acc-VS are not sufficient to ensure that the entire voting process is accessible, does not violate any of the VVSG requirements (e.g. privacy), and does not "break".

➔ Goal is to create a place for a test method to ensure that we've looked at how the whole process fits together
➔ We need a requirement for a system to support end-to-end process accessibility, which will then be demonstrated by end-to-end, comprehensive accessibility evaluation
➔ The 2nd part of the requirement: vendor shall document the process by which the system supports end-to-end accessibility.

12

Technical Guidelines Development Committee
22 March 2007 Plenary Meeting

We drafted wording to address the voter verification and accessibility issue. (To be discussed in detail Friday)

- **12.3.1-D** Accessibility of Paper-based Vote Verification – If the Acc-VS generates a paper record (or some other durable, human-readable record) for the purpose of allowing voters to verify their ballot choices, then the system should provide a mechanism that can read that record and generate an audio representation of its contents. The use of this mechanism should be accessible to voters with dexterity disabilities.

Technical Guidelines Development Committee
22 March 2007 Plenary Meeting

Dexterity

- **12.3.4-C** Ballot Submission
  If the voting station supports ballot submission for non-disabled voters, then it shall also provide features that enable voters who lack fine motor control or the use of their hands to perform this submission.

➔ We recognize that privacy is an important part of accessibility. This req. from VVSG 05, supports privacy, but also independence. This has implications for the software independence and accessibility for EBMs and PCOS. More discussion on this tomorrow. Also needs discussion with the EAC.

We have completed the first phase of our usability performance benchmark research

- Our goal for the VVSG is to have quantitative performance benchmark requirements; with conformance determined by running usability tests with "typical" voters

- To achieve this:
  - Develop a test protocol and metrics ✔
  - Show the test is valid ✔
  - Show the test is reliable (reproducible, repeatable)
  - Test a number of commercial machines to set a performance baseline
  - Determine
    - most cost effective means of running the conformance tests with large enough numbers of voter participants to ensure statistically significant results
    - which metrics to use for conformance
    - statistical treatment of these metrics
    - benchmark to be applied

15

Test Protocol

- Recruit participants with specified demographics
- Use standard ballot
  - Medium complexity, 20 contests and referenda
  - Asked vendors to implement ballot to give best performance on their system
- Follow script
  - Participants are told how to vote (28 entries)
  - No assistance or training given
- Measure errors and time to vote
  - Errors: differences from the 27 votes the participants were asked to cast and did they cast the ballot
- Administer questionnaire: modified Survey of User Satisfaction (SUS), 10 statements, 5 point Lickert scale, e.g.,
  - I felt confident that I used this voting machine correctly.
  - I think that I would need support to be able to use this voting machine.
  - I thought this voting machine was easy to use.

The test protocol is valid.

- We ran the test protocol
  - 47 participants, 21-30, HS through college: we wanted people who would perform reasonably well
  - 2 different types of voting systems.

- The test produced the types of errors we expected
- The test detected differences between the machines
  - The differences matched our predictions based on expert usability review and other research results
- The differences were statistically significant for the errors (as we expected)
  - Time on task did not show statistical significance
  - Most were confident about voting correctly with high SUS scores (no statistically significant differences between systems)

Metrics and how to count errors for benchmarks are under discussion.

- Need to determine how to count the errors for the effectiveness benchmark
  - Binary:
    - Ballot choices are correct or not.
    - Ballot was cast or not.
    - Binary Success rate #correct/total #participants
  - Number of errors for each contest
  - Number of individuals making an error
  - Error rate: #errors /
    (#participants * voting opportunities per participant)
  - Need tight confidence intervals
- Time on a task and satisfaction as benchmarks
  - Set a lower bound that we expect all voting systems will meet.   A lower score would indicate a large flaw in the system.

Technical Guidelines Development Committee
22 March 2007 Plenary Meeting

Experiments to determine reliability are underway

- **Test Repeatability** – can test results be repeated with same test administrators and same participant pool.
- **Test Reproducibility** – can tests results be reproduced in different geographic regions and with different test administrators
- Series of tests
  - Larger set of participants: mix of age range, F/M, Socio-economic status, geographic region
  - 200-400 participants
- To set benchmark: tests on a wider range of commercial systems.
- Future: Test of usability for voters with disabilities
  - Wide range
  - Different interaction
  - Do the benchmarks generalize

Technical Guidelines Development Committee
22 March 2007 Plenary Meeting

Next Research Steps

- Harmonizing with D4D ballot guidance, when accepted by the EAC.

- Additional voting-specific plain language guidance

- Color guidance

- Analysis of how/when to use icons and pictures for cognitive disabilities

- Usability of documentation guidance

- Accessibility performance benchmarks