

How well does web-based testing work?

Results from a survey of NetAssess™ users

Alan Mead

Aon Consulting

alan_d_mead@aoncons.com

Why use the Internet or web?

- Lots of Internet hype... corresponding HR interest in using web technology
- Increasing requests for web-based version of 16PF™ questionnaire
 - Convenience
 - Technological edge
 - Perceived larger markets
 - Everyone's doing it
 - Reduce price
 - Experimentation
- Although research sparse, generally positive results (e.g., cognitive web-based experiments were generally replicated well in lab settings; Musch & Reips, 2000)

Kinds of web-based tests

Factor	Low Level	High Level
1 Inclusiveness	Relies on optional technologies; bleeding-edge technology; or requires add-ons ★	Uses basic/old technologies, does not require plug-ins
2 Computer resources	Uses no CPU, RAM, or disk space ★	Requires significant CPU, RAM, or disk space
3 Network resources	No streaming media, few small images, simple HTML ★	Requires streaming media, heavily graphical, complex/verbose HTML, large style sheets
4 Security	Security not an issue	Security significant issue ★
5 Interactive responsiveness	All interaction crosses the Internet to server ★	All interaction occurs on client computer

NetAssess

- No test-user interface (yet)
- Test-takers access secure server with Explorer or Navigator (or Opera or...) using a single-use username/password pair
- Automatically emails assessment report or data
 - Reports are PDF files
 - Optionally placed in an encrypted zip file
- Built using apache+mod_ssl, Linux, Samba, Perl, Adobe Acrobat, some custom programming
- Gluing Linux web server to Windows report generator was hardest part

Participants

- 200 NetAssess users sent email soliciting participation and describing inducements; 68 responded with useable data
- Response rate nominally 38%
 - But actual rate higher due to duplicated addresses and bounces
 - Actual response rate possibly 40% - 50%
- Modal respondent:
 - Male (65%), 40-49 (53%), lived in US (91%)
 - Most were management consultants or organizational psychologists (57%)
 - Assessment purposes included: Coaching (56%), Selection (54%), Development (49%), and Career Counseling (34%)

About the survey

- Administered on-line from an IPAT server
- Average respondent spent about 20 minutes
- Four sections:
 - About the test-user (demographics, etc.)
 - Satisfaction with NetAssess (including problems and desired features)
 - Perceptions of test-takers
 - Computer familiarity

Satisfaction

- 81% “satisfied” or “quite satisfied”
- 70% predicted that they would use NetAssess more over next 2 years
- Drivers:
 - Remote administration (94%)
 - Quick reporting (54%)
 - Saves time/money (35%)
 - More convenient for test-taker (10%)
 - Easy ordering (33%)
 - “Tech Image” (32%)
 - Web site (21%)
- Only variable separating satisfied from dissatisfied was self-reported rate of technical difficulties (5.4% vs. 26.25%)

Problem nature and frequency

- About 7% of administrations have some problem
- About 9% of test-takers need technical support
- Providing technical support is a sensitive issue:
 - Test-taker's are IPAT's customers' clients
 - IPAT has no idea who they are or why they are being assessed
 - Compared to supporting test users, this could generate mountains of calls
- NetAssess displays the name, email, and phone number of the test-user (not IPAT)
- So who *really* does technical support?
 - IPAT: 14%, Test-User: 65%, no need: 21%

Ideal Test-Taker

- Several people said “anyone”
- Others suggested these themes:
 - Have access to Internet
 - Computer literate
 - Able and willing to follow instructions
 - “White-collar”
 - No reason to “game” test

When not to use web-tests?

Characteristics discouraging web-based testing

Reason	Count	Percent
Test taker does not like, or is not skilled with, computers or Internet	16	30.8
Younger, poorer, less educated test takers	10	19.2
“high-stakes and unproctored”	7	13.5
Test taker has no/poor Internet access	6	11.5
Test takers in any way hostile to testing	6	11.5
Test takers have privacy concerns	2	3.8
Test takers with reading difficulties	2	3.8

Note: Fifty-two individuals responded in some manner.

Trace measures

- Web servers keep detailed logs of testing sessions, providing some “trace measures’ of test administration:
 - Tuesday is the most popular day for testing but 13% of assessments happen over the weekend
 - About 30% of test-takers respond outside regular business hours; about 15% respond between 11PM and 6AM CST
 - Although most test-takers proceed rapidly through assessment, a few seem to take hours (possibly due to interruptions)

Limitations/Future Directions

- 68 people is not a huge sample
- All perceptions based on one implementation (NetAssess)
- All in sample self-selected into “NetAssess users” group... would non-users respond differently?

Conclusions

- Drivers of satisfaction include:
 - Perceptions of reliability (currently)
 - Ease of use
 - Relationship with customer/test-user
- It's surprising what are **not** drivers: Aesthetics; Complexity of system; Cost (within reason); Computer comfort
- We testing “works” most of the time

Conclusions (cont.)

- Security is a small or non-issue for most test users and test-takers
 - Possibly, they trust “us” as having “solved’ that problem
 - Maybe using an SSL-enabled web server is important
- Web-based testing may not yet be appropriate for all jobs and all test-takers everyone

Misc. design recommendations

- The system should be up, accessible, and quick
- Even the most reliable system needs to present pages and save work of completed pages in case the test-taker gets disconnected
- But plan the numbers of items in a page to minimize server hits; each hit is perceived as slow by test-takers
- I would recommend using an SSL web server
- Some NetAssess users found PDF files too hard to use; would prefer text or Word
- Some NetAssess users would prefer a user interface; others like NetAssess' simplicity
- Several noted that the single-use username/passwords were hard to manage, transmit, and test-taker's mis-typed them
- An on-line manual is not particularly important to user satisfaction