Index

A
accessibility, qualities of usability, 4–6
accuracy statistics, performance data, 249–250
activity component, Bailey’s Human Performance Model, 7
actors, end users as, 118
ambiguity, moderator comfort with, 50
analyzing data. See data analysis
assessment (summative) tests, 34–35
for first-time users, 201
iterative testing in development lifecycle, 41–42
methodology for, 35
objectives of, 34–35
when to use, 34
assistance
how to assist participants, 211–212
when to assist participants, 211
associations, sources for participant selection, 137
attention span, of test moderators, 51
attitudes
discovering with pre-test questionnaires, 175–177
mental preparation for test sessions, 218
audio recordings, debriefing sessions, 236

B
background questionnaire, 162–164
administration of, 163–164
ease of use of, 163
focus of, 163
overview of, 162
participants filling out preliminary documents, 220
pilot testing, 163
purposes of, 162–163
Bailey’s Human Performance Model, 7
“bare attention”, test moderators practicing, 61
behavior
performance data, 165–166
rationales for, 208–209
behavioral measurements, of product usability, 13
benchmarks
as means of developing user profile, 119
for measuring usability, 4
profitability and, 22
test plans and, 80–82
validation tests and, 36
“best case” testing, 133
between subjects design, for test plan, 75
“big picture” view, of test moderators, 51–52
biometric data, gathering, 112
blueprint, test plan as, 66
body language, of test moderators, 203
branching questions, 198–199
BRD (business requirements documents), 118
bugs, assisting participants and, 212
business requirements documents (BRD), 118

card sorting, for findability of content or functionality, 18
catastrophe, validation tests as, 36
categorizing user profiles, 124
cause-and-effect relationships, in experimental method, 23
checkbox questions, 198
checklist 1 (approximately a week before test), 214–216
  checking equipment and test environment, 216
  conducting pilot test, 215
  freezing further development, 216
  making revisions, 215–216
  taking test yourself, 214
checklist 2 (one day before test), 216–217
  assembling test materials, 217
  checking equipment and test environment, 217
  checking product software and hardware, 217
  checking status of participants, 217
  checking video equipment, 216–217
checklist 3 (day of test), 217–225
  closing session, 224–225
  debriefing observers, 225
  debriefing participants, 224
  distributing/reading task scenarios, 224
  filling out post-test questionnaires, 224
  filling out preliminary documents, 220
  filling out pretest questionnaires, 220
  greeting participants, 219–220
  mental preparation of moderator, 218–219
  moving to test area and preparing for test, 220–221
  organizing data collection and observation sheets, 225
  overview of, 217–218
  providing adequate time between sessions, 225
  providing prerequisite training if part of test plan, 223–224
  reading orientation script, 220
  setting decorum for observers present, 221–223
  starting data collection, 224
  starting recordings, 221
  checklists, preparing for test sessions, 213–214
churches, sources for participant selection, 136–137
classical laboratory, 108–110
  advantages of, 109–110
  disadvantages of, 110
  overview of, 108–109
classifiers
  matrix test design and, 125
  participant selection and, 121–122
closing test sessions, 224–225
clubs, sources of participant selection, 136–137
code of ethics, 52
coding schemes, for note-taking, 171
college campuses, sources of participant selection, 139–140
common sense, usability design and, 9
communication
  orientation script as communication tool, 154
  skills of test moderators, 52
test plan as communication vehicle, 66, 74
community groups, sources of participant selection, 136–137
compensation, of test participants, 150–151
competitive edge, usability and, 23
compiling data, 246–247
  other measures for, 256
  overview of, 246–247
  while testing, 247
complaints, as reason for testing products, 69
compensation, of test participants, 150–151
consent forms
  participants filling out, 220
  recording, 173–174
courseware, as test moderators, 47–48
coping, as reason for testing products, 69
criticality
  prioritizing problems by, 261
  prioritizing tasks by, 86
cues
  avoiding in task scenarios, 184
customer support
  exploratory studies and, 31
  testing product software and hardware, 217
cutoff date, 39–41
  objectives of, 37
  of prototypes, 264
  when to use, 37
  compensation, of test participants, 150–151
  competitive edge, usability and, 23
  compiling data, 246–247
  other measures for, 256
  overview of, 246–247
  while testing, 247
  complaints, as reason for testing products, 69
  consent forms
  participants filling out, 220
  recording, 173–174
  consultants, as test moderators, 47–48
  content
  card sorting for finding, 18
  overview of, 192–193, 195
c oncicients
  pretest questionnaire, 129–130
  establishing prerequisite knowledge prior to product use, 181
criticality
  prioritizing problems by, 261
  prioritizing tasks by, 86
cues
  avoiding in task scenarios, 184
  moderator sensitivity to nonverbal, 208
customer support
  exploratory studies and, 31
  design and, 10
  overview of, 110
  while testing, 247
  complaints, as reason for testing products, 69
  consent forms
  participants filling out, 220
  recording, 173–174
  consultants, as test moderators, 47–48
  content
  card sorting for finding, 18
  overview of, 192–193, 195
context component, Bailey’s Human Performance Model, 7
counterbalancing technique
  avoiding biases with, 183
  within-subjects design and, 75–76
coworkers, sources of participant selection, 137–138
craigslist, sources of participant selection, 138–139
criterion tests
  for checking proficiency, 129
  establishing prerequisite knowledge prior to product use, 181
improving profitability and, 22

customers
increasing repeat sales, 22
sources of participant selection, 135

D
data
compiling, 246–247, 256
deciding what type to collect, 167–168
organizing raw data, 248–249
overview of, 245–246
performance data, 165–166
preference data, 166
summarizing performance data, 249
summarizing preference data, 254–256
summarizing scores by group or version, 256–258
task accuracy statistics, 249–250
task timings statistics, 250–254
data analysis
comparing differences between groups or product versions, 264–265
identifying tasks not meeting success criterion, 258–259
identifying user errors and difficulties, 260
inferential statistics and, 265–267
overview of, 258
prioritizing problems, 261–263
source of error analysis, 260–261
data collection
biometric data, 112
deciding what type of data to collect, 167–168
fully automated data loggers, 168–169
list of basic information covered, 168
manual data collection, 170
methods for, 168
online data collection, 169
organizing, 225
other methods for, 170–173
overview of, 165–167
research questions, 167
starting during test session, 224
test moderators overinvolvement with, 57
test plan and, 88
user-generated data collection, 169–170
data gather, lab setup and, 112
data loggers
fully automated, 168–169
summarizing collected data, 249
debriefing
audio recording of debriefing session, 236
“devil’s advocate” technique, 238–240
guidelines for, 231–233
guides, 199
locations for, 231
manual method, 235–236
observers, 223, 225, 241–242
overview of, 229
participants, 224, 230–231
reasons for, 229–230
replaying the test (retrospective review), 235
reviewing alternate product versions, 236
as source of preference data, 254
video method, 236
“What did you remember” technique, 236–238
decorum, setting for observers, 221–223
deliverables, 245
descriptive statistics, 249, 265
design
accessibility and, 5
design expertise compared with technical expertise, 11
generating user profiles and, 118
goals of usability testing and, 22
hard-to-use products and, 9
implementation not matching, 11–12
iterative development and, 14
not soliciting design ideas from participants, 234
participatory process for, 17
preparing for test session and, 214
test plans and, 74
designers, test plan as vehicle of communication with, 66
developers
neglecting human needs, 7–8
test plan as vehicle of communication for, 66
test plans and, 74
development lifecycle
assessment or summative tests and, 34–35
comparison tests and, 37–38
exploratory or formative studies and, 29–34
freezing development during test sessions, 216
involving users in, 13
iterative testing and, 39
nonintegrated approach to, 9–10
test 1: exploratory/comparison test, 39–41
test 2: assessment test, 41–42
test 3: verification test, 42–44
types of tests and, 27–29
user input included in, 14
validation or verification tests, 35–37
“devil’s advocate” technique, 238–240
example of, 239–240
how to implement, 238–239
overview of, 238
disaster insurance, validation tests as, 36
“discount” usability testing, 73
documentation
reasons for testing and, 69
requirements documents, 117–118
documentation (continued)
specification documents, 117–118
tasks lists and, 80
topics in post-test questionnaire, 195
user profile and, 122–123

early adopters, 8
early analysis/research. See exploratory
(formative) studies
ease of learning
informed design and, 22
measuring, 13
ease of use
measuring, 13
products and, 175
effectiveness
informed design and, 22
as performance criteria in validation testing, 36
qualities of usability, 4
efficiency
informed design and, 22
as performance criteria in validation testing, 36
qualities of usability, 4
specialization and, 9
electronic observation room, 107–108
advantages of, 107
disadvantages of, 108
overview of, 107
electronic observation room setup, 107–108
elements of usability testing, 25
empathy, characteristics of test moderators, 51
employment agencies, sources of participant selection, 141–142
enabling vs. leading, by test moderators, 57
end users. See also test participants; users as actors, 118
differentiated from purchasers, 116–117
end-to-end study, of product component integration, 36
environment
checking a week before test, 216
checking one day before test, 217
classic laboratory setup, 108–110
data gather/note taker and, 112
electronic observation room, 107–108
equipment, tools, and props, 111
experimental method and, 24
gathering biometric data and, 112
getting help, 112
large single-room setup, 105–107
limitations of usability testing and, 26
location selection, 94–96
modified single-room setup, 103–105
multiple geographic locations, 96–98
overview of, 93–94
portable test lab, 100–101, 110–111
product/technical experts, 113
simple single-room lab setup, 101–103
test observers, 113–114
in test plan, 87
timekeeper, 113
user sites, 98–99
environment (emotional), nonjudgmental, 219
equipment
checking a week before test, 216
checking one day before test, 217
lab setup and, 111
in test plan, 87
 ergonomic. See UCD (user-centered design)
errors
analyzing differences between groups or product versions, 264
conducting a source of error analysis, 260–261
identifying user errors and difficulties, 260
ethnographic research, 16
evidence-gathering, 35
expectations, explaining in orientation scripts, 160–161
experience design, 5
experimental design
reasons for not using classical approach, 24–25
usability testing and, 23–24
expert (heuristic) evaluations, 18–19
expertise
criterion tests for checking, 129
defining/measuring in participant selection, 119–121
design expertise compared with technical expertise, 11
ensuring minimum expertise of participants, 187–188
establishing prerequisite knowledge prior to product use, 181
techniques for rating participants, 179–180
test moderators acting too knowledgeable, 57
exploratory (formative) studies, 29–34
conducted as comparison tests, 34, 38
example of, 32–34
internal participants and, 133
iterative testing and, 39–41
methodology for, 30–32
objectives of, 29–30
when to use, 29
external consultants, as test moderator, 47–48
eye-tracking, gathering biometric data, 112
family, sources of participant selection, 133
feedback, creating test plans and, 65
field tests
in-session tips for, 99
locations for test lab, 98–99
fill-in questions, 198
first impressions, pre-test questionnaires and, 175–177
fixed test lab
classic laboratory setup, 108–110
electronic observation room setup, 107–108
large single-room setup, 105–107
modified single-room setup, 103–105
overview of, 101–111
simple single-room lab setup, 101–103
flexibility, of test moderators, 50–51
focal point, test plan as, 66–67
focus groups, for researching and evaluating concepts, 17
follow up studies, UCD techniques, 20
formative studies. See exploratory (formative) studies
formats
post-test questionnaire, 192
questions. See question formats
forms
for collecting data, 170, 172
for compiling data, 247
explaining in orientation scripts, 161
lab equipment and, 111
nondisclosure forms, 173–174
frequency
prioritizing problems by, 262
prioritizing tasks by, 85–86
friends, sources of participant selection, 133
frustration
allowing participants time to work through hindrances in testing, 55
assisting participants and, 212
eliminating design problems and, 22
measures of usability, 4
fully automated data loggers, 168–169. See also data loggers
functional specification documents, 117–118

g
geographic locations, for test lab, 96–98
goals, 21–23
design-related, 22
organizational, 16
overview of, 21
profit-related, 22–23
reviewing in test plan, 67–68
groups. See also user groups
analyzing differences between groups or product versions, 264–265
summarizing scores by group or version, 256–258
guidelines, for debriefing, 231–235
guidelines, for moderating test sessions
body language and tone of voice, 203
ensuring participants complete tasks before moving on, 210–211
how to assist participants, 212–213
impartiality, 202–203
making mistakes and, 210
not rescuing participants when they struggle, 209–210
objective, but relaxed approach, 209
overview of, 201–202
probing/interacting with participants appropriately, 206–209
“thinking-aloud” technique and, 204–206
treating participants as individuals, 203–204
when to assist participants, 211–212
guidelines, for observers, 154–155

H
hard-to-use products
design-related problems, 9
implementation not matching design, 11–12
machine or system focus instead of human orientation, 7–8
overview of, 6
specialization and lack of integration, 9–11
target audience expands and adapts, 8–9
hardware
checking one day before test, 217
exploratory studies and, 30–31
topics in post-test questionnaire, 196
helpers, lab
data gather/note taker, 112
overview of, 112
product/technical experts, 113
test observers, 113–114
timekeeper, 113
heuristic evaluations, UCD techniques, 18–19
horizontal representation, prototypes and, 31
hot spots, data analysis and, 245
human component, Bailey’s Human Performance Model, 7
Human Factors and Ergonomics Society, 52
human factors engineering. See UCD (user-centered design)
human factors specialists, as test moderator, 46
humor, moderator/participant interaction and, 209
hypothesis
in experimental method, 23
in usability testing, 25
I

identity, protecting privacy and personal information of participants, 151
impartiality, of test moderators, 202–203
implementation, not matching design in hard-to-use products, 11–12
in house list, sources of participant selection, 135
independent groups design for test plan, 75
testing multiple product versions and, 76–77
inferential statistics, 265–267
information, regarding users, 117
information-gathering, assessment tests and, 35
integration
specialization causing lack of, 9–11
testing techniques for ensuring, 201–202
internal participants, sources of participant selection, 132–133
international users, as factor in lab location, 97
intervention, when to intervene, 225
interviews. See also pre-test questionnaires
interviews, screening questionnaire, 145–146
introductions. See also orientation scripts
observers, 220
during orientation, 159
setting decorum for observers present at test session, 221–222
invisibility, usability and, 6
ISO (International Organization for Standardization)
SUS, 194
UCD and, 12
iterative testing
benefits of, 28
development cycles and, 14
overview of, 39
power of, 28
test 1: exploratory/comparison test, 39–41
test 2: assessment test, 41–42
test 3: verification test, 42–44

J

jargon, avoiding in task scenarios, 184
jumping to conclusions, test moderators and, 58

K

Keynote data logger, for remote tests, 169
knowledge, test moderators acting too knowledgeable, 57

L

lab setup options
classic laboratory, 108–110
electronic observation room, 107–108
large single-room, 105–107
modified single-room, 103–105
portable test lab, 100–101, 110–111
simple single-room, 101–103
LCUs (least competent users), 146–147
“learn as you go” perspective, in UCD, 15–16
learnability, qualities of usability, 4
learning
measuring ease of, 13
mediation skills, 60
taping sessions as learning tool, 59
test moderators as quick learners, 48–49
test moderators learning basic professional principles, 59
test moderators learning from watching, 59
least competent users (LCUs), 146–147
Likert scales, 197
limitations, of usability testing, 25–26
listening skills, of test moderators, 49–50
locations, for debriefing, 231
locations, for test lab
factors in selection of, 94–96
multiple geographic locations, 96–98
overview of, 94
user sites, 98–99
loggers. See data loggers
logistics
lab setup and, 95
test plan and, 87

M

machine focus, reasons for hard-to-use products, 7–8
machine states, for tasks, 79–80
malfunctions, assisting participants and, 212
management, role in UCD, 15
manual data collection, 170
manual debriefing method, 235–236
marketing research firms, sources of participant selection, 140–141
marketing specialists, as test moderator, 46
marketing studies, sources of participant selection, 118
materials
assembling one day before test, 217
background questionnaire, 162–164
data collection tools. See data collection
debriefing guide, 199
ensuring minimum expertise, 187–188
getting view of user after experiencing the product, 188–189
guidelines for observers, 154–155
nondisclosures, consent forms, recording waivers, 173–174
optional, 187
orientation scripts. See orientation scripts overview of, 153–154
post-test questionnaire. See post-test questionnaire
prerequisite training, 190–192
pre-test questionnaires. See pre-test questionnaires
prototypes and products, 181–182
question formats, 197–199
task scenarios. See task scenarios for tasks, 79–80
testing features for advanced users, 189–190
matrix test design, 125
mean time to complete, timing statistics, 251
median time to complete, timing statistics, 251–252
mediation skills, learning, 60
memory skills, of test moderators, 49
mental preparation, of test moderator on day of test, 218–219
mentors, test moderators working with, 59–60
methodology, data collection
fully automated data loggers, 168–169
manual data collection, 170
online data collection, 169
other methods, 170–173
overview of, 168
user-generated data collection, 169–170
methodology, test
assessment tests, 35
comparison tests, 38
experimental design, 23–24
exploratory studies, 30–32
reasons for not using classical approach, 24–25
restricting interaction with test moderator, 24
test plan and, 73–74
validation tests, 36
milestones, test plan as, 66–67
mistakes
acceptability in testing environment, 205–206
continuing despite, 210
not blaming participants, 213
modeling, exploratory studies and, 30–31
moderators
acting too knowledgeable, 57
allowing participants time to work through hindrances, 55
ambiguity, comfort with, 50
assessment tests and, 35
attention span of, 51
‘‘big picture’’ view of, 51–52
characteristics needed, 48
communication skills, 52
controlling observers during sessions, 223
degree of interaction with, 27–28
empathetic, 51
encouraging participants, 55–56
ensuring participants complete tasks before moving, 210–211
exploratory studies and, 31
external consultants as, 47–48
flexibility of, 50–51
getting most out of participants, 52–53
how to assist participants, 212–213
human factors specialists, 46
improving skills, 58–59
jumping to conclusions, 58
leading vs. enabling, 57
learning basic professional principles related to, 59
learning from watching, 59
learning mediation skills, 60
listening skills, 49–50
marketing specialists, 46
memory skills, 49
minimizing the differences between different moderators, 158
not rescuing participants when they struggle, 209–210
organization and coordination skills, 52
overinvolvement with data collection, 57
overview of, 45
practicing ‘‘bare attention’’, 61
practicing moderation skills, 60
probing/interacting with participants appropriately, 206–209
quick learners, 48–49
rapport with test participants, 49
relational problems with participants, 58
restricting interaction with, 24
retrospective review by participants, 54–55
rigidity with test plan, 58
role in selecting test format, 53
role outlined in test plan, 87–88
sit-by vs. remote observation, 53–54
taping sessions as learning tool, 59
team members as, 47
technical communicators, 47
test plan as vehicle of communication, 66
‘‘thinking-aloud’’ by participants, 54
treating participants as individuals, 203–204
troubleshooting typical problems, 56
UCD in background of, 48
validation tests and, 37
value of test plan to, 74
when to assist participants, 211–212
who should moderate, 45–46
working with mentors, 59–60
monitoring software, for compiling data, 248
Morae
- as fully automated data logger, 168
- for recording lab sessions, 112
- multidisciplinary team approach, in UCD, 14–15

N
- newspaper advertisements, sources of participant selection, 142–143
- “non-association” guideline, for products, 159
- nondisclosure forms
  - participants filling out, 220
  - as test material, 173–174
- nonjudgmental approach, 208
- nonverbal cues, moderator sensitivity to, 208
- note-taking
  - collecting data and, 165
  - compiling data and, 247
  - lab equipment and, 111
  - lab setup and, 112
  - shorthand or codes for, 171
- novice users, matching tasks to experience of participants, 184
- number of participants, 125–126

O
- objectives
  - assessment tests, 34–35
  - comparison tests, 37
  - exploratory studies, 29–30
  - organizational, 16
  - reviewing as part of session preparation, 218
  - reviewing purpose and goals in test plan, 67–68
  - validation tests, 35–36
- objectivity, moderating and, 45–46, 209
- observation sheets, organizing, 224, 225
- observers
  - debriefing at end of study, 243
  - debriefing following test sessions, 225
  - decorum of, 221–223
  - guidelines for, 154–155
  - inconspicuousness of, 222
  - introducing, 220
- lab setup and, 113–114
- manual data collection by, 170
- moderators controlling participation in debriefing, 234–235
- reasons for debriefing, 229–230, 241
- reducing amount of writing required of, 171
- role during sessions, 223
  - sit-by vs. remote, 53–54
- online data collection, 169
- open-ended interview, 143–144
- organizational skills, of test moderators, 52
- organizations
  - constraints on use of experimental method, 24
  - eliminating design problems and frustration, 22
  - goals and objectives, 16
  - “learn as you go” perspective, 15–16
  - management role, 15
  - multidisciplinary team approach, 14–15
  - overview of, 14
  - profitability, 22–23
  - usability labs and, 93
  - user input into development, 14
- orientation scripts
  - asking for questions, 161
  - describing test setup, 160
  - expectations and requirements explained, 160–161
  - forms explained, 161
  - introductions in, 159
  - offering refreshments, 159
  - overview of, 155–161
  - professional/friendly tone, 156
  - reading to participants, 159–160
  - session purpose explained, 159–160
  - shortness of, 156–157
  - writing it out, 158
- outliers, range statistics and, 252
- Ovo Logger, 168

P
- paper prototyping, UCD techniques, 18–19
- participants
  - allowing time to work through hindrances, 55
  - assessment tests and, 35
  - background questionnaire screening, 162–163
  - characteristics, in test plan, 72–73
  - checking status one day before test, 217
  - completing tasks before moving to next, 210–211
  - debriefing, 224, 230–231
  - establishing prerequisite knowledge prior to product use, 181
  - explaining what is expected, 160–161
  - exploratory studies and, 31
  - failure to show up, 226
  - filling out post-test questionnaires, 224
  - filling out preliminary documents, 220
  - greeting on day of test, 219–220
  - how to assist, 212–213
  - lab setup and, 95–96
  - learning if product valued by, 177–178
  - matching tasks to experience of, 184
  - moderators encouraging, 55–56
  - moderators getting most out of, 52–53
  - moderators not rescuing when they struggle, 209–210
moderators probing/interacting with appropriately, 206–209
moderators rapport with, 49
moderators relational problems with, 58
moderators treating as individuals, 203–204
orientation. See orientation scripts
qualifying for inclusion in test groups, 179–181
reading orientation script to, 157–158
reading task scenarios, 186–187
reading task scenarios to, 185
reasons for debriefing, 229–230
retrospective review by, 54–55
“thinking-aloud”, 54
validation tests and, 37
what not to say to, 227–228
when to assist, 211–212
participants, selecting. See also user profiles
answer sheet for screening questionnaire, 131
benchmarks as means of developing user profile, 119
categorizing user profiles, 124
categorization of users, 115–116
classifying user groups, 119
college campuses as source, 139–140
compensating participants, 150–151
completing screening questionnaire always or when fully qualified, 144
Craigslist as source, 138–139
documenting user profile, 122–123
employment agencies as source, 141–142
expertise, defining/measuring, 119–121
formulating screening questions, 128–131
in house list of customers as source, 135
identifying specific criteria for, 127–128
including least competent users (LCUs) in testing samples, 146–147
information regarding users, 117
internal participants as source, 132–133
marketing research firms or recruiting specialists as source, 140–141
matrix test design and, 125
newspaper advertisements as source, 142–143
not testing only “best” end users, 147–148
number of participants, 125–126
ordering screening questions, 129
overview of, 115
personal networks and coworkers as source, 137–138
protecting privacy and personal information of participants, 151
purchasers differentiated from end users, 116–117
qualified friends and family as source, 133
questionnaire vs. open-ended interview for screening, 143–144
requirements and classifiers in selection process, 121–122
requirements and specification documents and, 117–118
reviewing user profile to understand user backgrounds, 127
role of product manager (marketing) in, 118–119
role of product manager (R&D) in, 118
sales reps list of customers as source, 136
scheduling and confirming participants, 148–149
screening considerations, 143
screening interviews, 145–146
screening questionnaire, 126–127
societies and associations as source, 137
sources of participants, generally, 131–132
structured analyses or marketing studies and, 118
testing/revising screening questions, 131
tradeoffs, 148
user groups, clubs, churches, community groups as source, 136–137
visualizing/desciring, 116
Web site sign up as source, 133–134
participatory design process, UCD techniques, 17
passwords, protecting privacy and personal information of participants, 151
performance
background questionnaire focusing on, 163
measures in test plan, 88–89
performance data
accuracy statistics, 249–250
advantages of “thinking aloud” technique, 204
data collection and, 165–166
list of examples, 166
summarizing, 249–250
timings statistics, 250–254
personal information, protecting, 151
personal networks, participant selection and, 137–138
phone screener, background questionnaire compared with, 162
pilot testing
approximately a week before test session, 215
background questionnaire, 163
post-test questionnaire, 196–197
planning. See test plan
“playing dumb”, as technique for test
moderators, 57
portable test lab
advantages of, 100
disadvantages of, 100–101
overview of, 100
as recommended testing environment, 110–111
post-test questionnaire, 192–197
areas and topics for, 195–196
brevity and simplicity of, 196
debriefing and, 231–232
distributing before or after sessions, 193
filling out during test sessions, 224
marking areas to explore during debriefing, 233
overview of, 192
pilot testing, 196–197
research questions for, 193
reviewing, 232
sources of preference data, 254
subjective preferences and, 193–194
practice, test moderators, 60
preconceptions, 202
preference data
- advantages of “thinking aloud” technique, 204
data collection and, 166
list of examples, 167
post-test questionnaire and, 193–194, 254
summarizing, 254–256
preference measures, in test plan, 90
preliminary analysis, comprehensive analysis compared with, 245–246
prerequisite training
- comprehensiveness of, 190
- providing on day of test, 223–224
- purpose of, 188–190
questions regarding, 191–192
testing functionality and, 190
user learning as focus of, 191
pre-test questionnaires, 174–181
attitudes and first impressions discovered by, 175–177
filling out on day of test, 220
learning if participants value the product, 177–178
overview of, 174
prerequisite knowledge established by, 181
qualifying participants for inclusion in test groups, 179–181
principles, of UCD
- focusing on users and tasks, 13
- iteration in design/testing development cycles, 14
- measuring ease of learning and ease of use, 13
overview of, 13
prioritizing issues, from test sessions, 243
prioritizing problems
- by criticality, 261
data analysis, 261–263
- by frequency of occurrence, 262
- by severity, 262
prioritizing tasks, 85–87
- by criticality, 86
- by frequency, 85–86
overview of, 85
by readiness, 86–87
by vulnerability, 86
privacy, protecting privacy of participants, 151
problem solving
debriefing and, 234
moderator/participant interaction and, 209
prioritizing problems and, 261–263
product experts, lab setup and, 113
product manager (marketing), role in participant selection, 118–119
product manager (R&D), role in participant selection, 118
product requirements documents, 117–118
products
- complaints as reason for testing, 69
- ease of use, 175
- first impressions, 175–176
- learning if participants value, 177–178
- ‘non-association’ guideline, 159
- reviewing alternate versions in debriefing process, 236
- revisions during test process, 215–216
- as test materials, 181–182
- testing multiple versions, 76–77
- user opinion after experiencing, 188–189
- user satisfaction with, 194
product/technical experts, lab setup and, 113
professionalism
- code of ethics, 52
- orientation scripts and, 158
profit, goals of usability testing and, 22–23
props, lab setup and, 111
prototypes
- comparison tests and, 264
- exploratory studies and, 30–31
- exploratory/comparison test and, 39–40
- paper prototyping, 18–19
- as test materials, 181–182
public relations, lab setup and, 95
purchasers, differentiated from end users, 116–117
Q
qualifying participants, for inclusion in test groups, 179–181
qualitative approach
- performance data and preference data, 166
test plan and, 90
types of tests and, 27
validation tests and, 37
quantitative approach
- manual data collection and, 170
Index

Q–S 345

performance data and preference data, 166
types of tests and, 27
validation tests and, 37
question formats
branching questions, 198–199
checkbox questions, 198
fill-in questions, 198
Likert scales, 197
overview of, 197
semantic differentials, 197–198
questionnaires
background, 162–164
post-test. See post-test questionnaire
pre-test. See pre-test questionnaires
screening. See screening questionnaire
for user-generated data collection, 170
wrong questions in, 226
questions
neutral questions to ask test participants, 208
oral questions for debriefing participants, 192
orientation and, 161
random sampling
in experimental method, 23
in usability testing, 25
range (high and low) of completion times, timing statistics, 252
raw data, organizing, 248–249
readiness, prioritizing tasks by, 86–87
recordings
audio recording debriefing session, 236
checking recording equipment one day before test, 216–217
permissions, 173–174, 220
starting on day of test, 221
recruiting. See participants, selecting
recruiting specialists, 140–141
refreshments, offering during orientation, 159
relational problems, test moderators and test participants, 58
relaxed approach
greeting participants, 219–220
guidelines for moderating test sessions, 209
remote observation, styles of test moderation, 53–54
remote usability testing
data collection tools for, 169
overview of, 97
replaying the test (retrospective review), 54–55, 235
reports, in test plan, 90–91
requirements
explaining session requirements in orientation scripts, 160–161
participant selection and, 121–122
research questions
data collection and, 167
evaluations for Web site, 70–71
exploratory/comparison test and, 40–41
post-test questionnaire and, 193
in test plan, 69–72
unfocused and vague, 70
research tool, usability testing as, 21
resources, listing required resources in test plan, 66
retrospective review (replaying the test), 54–55, 235
revisions, approximately a week before test, 215–216
rigidity, test moderators, 58
risk minimization, usability and, 23
rolling issues lists, observers creating, 241–243
sales, increasing repeat sales, 22
sales reps, role in participant selection, 136
sample size
constraints on use of pure experimental method, 24
in experimental method, 24
in usability testing, 25
satisfaction
determining user satisfaction with a product, 194
informed design and, 22
qualities of usability, 4–5
SCC (successful completion criteria), 80
scheduling test participants, 148–149
screen representations, exploratory studies and, 30
screening questionnaire, 126–131
answer sheet for, 131
completing always or when fully qualified, 144
conducting interviews, 145–146
considerations regarding, 143
formatting for ease of use, 130–131
formulating questions, 128–129
identifying specific criteria, 127–128
vs. open-ended interview, 143–144
ordering questions, 129
overview of, 126–127
testing/revising, 131
SD (standard deviation), of completion time, 253–254
semantic differentials, 197–198
sequencing task scenarios, 183
service, improving profitability and, 22
sessions
checklist for a week before test. See checklist 1
(approximately a week before test)
checklist for day of test. See checklist 3 (day of
test)
checklist one day before test. See checklist 2
(one day before test)
guidelines for moderating. See guidelines, for
moderating test sessions
overview of, 201–202
scripts or checklists, 154
what not to say to participants, 227–228
when to deviate from test plan, 226–227
when to intervene, 225
severity, prioritizing problems by, 262
single-room lab setup, large, 105–107
advantages of, 106
disadvantages of, 107
overview of, 105–106
single-room lab setup, modified, 103–105
advantages of, 105
disadvantages of, 105
overview of, 103–105
single-room lab setup, simple, 101–103
advantages of, 102–103
disadvantages of, 103
overview of, 101–102
sit-by, styles of test moderation, 53–54
social security numbers, protecting privacy and
personal information of participants, 151
societies, participant selection and, 137
software
checking one day before test, 217
topics in post-test questionnaire, 195
specialization, reasons for hard-to-use products,
9–11
specification documents, participant selection
and, 117–118
spreadsheets, for organizing raw data, 248
standard deviation (SD), of completion time,
253–254
statistics
deciding which technique to use, 266
descriptive, 249, 265–267
expertise required for use of, 24
inferential, 265–267
task accuracy, 250
timing, 250–254
structured analyses, participant selection and,
118
subjective preferences, post-test questionnaire
and, 193–194
success criterion, identifying tasks not meeting,
256, 258–259
successful completion criteria (SCC), in test plan,
80
summarizing data
other measures for, 256
overview of, 249
performance data, 249, 265
preference data, 254–256, 265
scores by group or version, 256–258
summary sheets, transferring collected data to,
249
summative tests. See assessment (summative)
tests
surveys
sources of preference data, 254
UCD techniques, 17–18
SUS (System Usability Scale), for determining
user satisfaction with a product, 194
system focus, reasons for hard-to-use products,
7–8
System Usability Scale (SUS), for determining
user satisfaction with a product, 194
T
taping sessions
to gain awareness of use of voice tones, 203
as learning tool, 59
target audience. See also user profiles
expansion and adaptation as reason for
hard-to-use products, 8–9
informed design and, 22
limitations of usability testing and, 26
task scenarios, 182–187
avoiding jargon and cues, 184
distributing/reading during test session, 224
letting participants read, 186–187
matching to experience of participants, 184
overview of, 182
providing substantial amount of work in each,
184–185
reading to participants, 185
realistic and with motivations to complete, 183
sequencing, 183
timing issues, 227
when to deviate from test plan, 226
tasks
accuracy statistics, 249–250
describing in test plan, 79
development process focusing on, 13
data, example, 83–85
engineering, 82
materials and machine states in test plan, 79–80
materials for, 79–80
not meeting success criterion, 258–259
prioritizing, 85–87
timings statistics, 250–254
Index □ T–U 347

team members, as test moderator, 47

teams
  multidisciplinary team approach in UCD, 14–15
test plan as vehicle of communication between, 66
technical communicators, as test moderator, 47
technical expertise, vs. design expertise, 11
technical experts, lab setup and, 113
techniques, usability
  card sorting, for findability of content or functionality, 18
  ethnographic research, 16
  expert (heuristic) evaluations, 18–19
  focus groups for research and evaluation, 17
  follow up studies, 20
  overview of, 16
  paper prototyping, 18–19
  participatory design, 17
  surveys, 17–18
  testing, 19–20
  walk-throughs, 18
terminal test, 176–177
test groups, qualifying participants for inclusion in, 179–181
test labs. See lab setup options
test materials. See materials
test observers. See observers
test participants. See participants
test plan
  as blueprint, 66
  as communication vehicle, 66
  data collection in, 88
  environment, equipment, and logistics, 87
  example of, 91
  example task in, 83–85
  as focal point and milestone in testing process, 66–67
  independent groups design, 75
  logistics of testing at user site, 98–99
  materials and machine states for tasks, 79–80
  methodology for, 73–74
  moderator’s role, 87–88
  for multiple product versions, 76–77
  for multiple user groups, 77–78
  overview of, 65
  participant characteristics and, 72–73
  parts of, 67
  performance measures, 88–89
  preference measures, 90
  prioritizing tasks, 85–87
  purpose and goals of, 67–68
  qualitative data, 90
  reasons for creating, 65–66
  reports, 90–91
research questions, 69–72
resource requirements, 66
SCC (successful completion criteria), 80
task descriptions, 79
task list, 82
timing and benchmarks, 80–82
when not to test, 68–69
when to deviate from, 226–227
when to test, 69
within-subjects design, 75–76
test sessions. See sessions
test setup, describing, 160
test types
  assessment tests, 34–35
  comparison tests, 37–38
  exploratory studies, 29–34
  overview of, 27–29
  validation tests, 35–37
The Observer data logger, 168
‘‘thinking-aloud’’ technique, 204–206
advantages of, 54, 204–205
disadvantages of, 54, 205
enhancing, 205–206
organizing raw data and, 248
overview of, 204
timekeeper, lab setup and, 113
timing, in test plan, 80–82
timing statistics, 250–254
  mean time to complete, 251
  median time to complete, 251–252
  range (high and low) of completion times, 252
  SD (standard deviation) of completion time, 253–254
tools, lab setup and, 111
topics, for post-test questionnaire, 195–196

U
UCD (user-centered design), 12–16
accessibility and, 5
background of test moderators in, 48
card sorting, 18
ease of learning and ease of use, 13
ethnographic research and, 16
expert (heuristic) evaluations, 18–19
focus groups, 17
follow up studies, 20
goals and objectives, 16
iterative development, 14
‘‘learn as you go’’ perspective, 15–16
management’s role, 15
multidisciplinary team approach, 14–15
organizations practicing, 14
overview of, 12–13
paper prototyping, 18–19
participatory design process, 17
UCD (user-centered design), (continued)
surveys, 17–18
techniques, 16
testing and, 19–20
user input included in development, 14
users and tasks as focus of, 13
walk-throughs, 18
usability engineering. See UCD (user-centered
design)
Usability Professionals’ Association, 52
usability testing
basic elements of, 25
defined, 21
goals of, 21–23
limitations of, 25–26
methodology for, 23–25
Usability Testing Environment (UTE), 169
usefulness, qualities of usability, 4
user expertise questionnaire, 180
user groups. See also groups
classifying in participant selection, 119
categorizing, 124
classification of users, 115–116
documenting, 122–123
determination of, 119–121
design and implementation not always
matching, 11–12
exploratory studies and design of, 29
user expertise
benchmarks as means of developing, 119
classifying, 124
categorizing, 124
classification of users, 115–116
documenting, 122–123
determination of, 119–121
design and implementation not always
matching, 11–12
exploratory studies and design of, 29
verification tests. See validation (verification)
tests
versions
analyzing differences between product
versions, 264–265
summarizing scores by, 256–258
video equipment, checking one day before test,
216–217
video method, for debriefing, 236
voice tone, guidelines for moderating test
sessions, 203
vulnerability, prioritizing tasks by, 86
waivers, as test material, 173–174
walk-throughs
participants walking through a product with
moderators, 31
UCD techniques, 18
Web site sign up, participant selection and,
133–134
Web sites, topics in post-test questionnaire,
195–196
“what did you remember” technique, 236–238
within-subjects design, 75–76