UX Evaluation: Reporting Results

27

Highlights

- Reporting different kinds of data:
    - Reporting informal summative results.
    - Reporting qualitative results.
- Report audiences.
- Report content.
- Report mechanics.
- Report tone.

27.1 INTRODUCTION

27.1.1 You Are Here

We begin each process chapter with a "you are here" picture of the chapter topic in the context of The Wheel, overall UX design lifecycle template (Fig. 27-1). This chapter is about reporting UX evaluation results and applies more or less regardless of the evaluation method or data collection technique.

27.1.2 Importance of Quality Communication

Evaluation reports are often required to communicate across discontinuities of time, location, and people. Redesign activities are often separated from UX evaluation by delays in time that can cause information loss due to human memory limitations. This is further aggravated if the people doing the redesign are not the ones who conducted the evaluation.

Finally, evaluation and redesign can occur at different physical locations, rendering all information not well communicated to be unrecoverable. UX evaluation reports with inadequate contextual information or incomplete UX
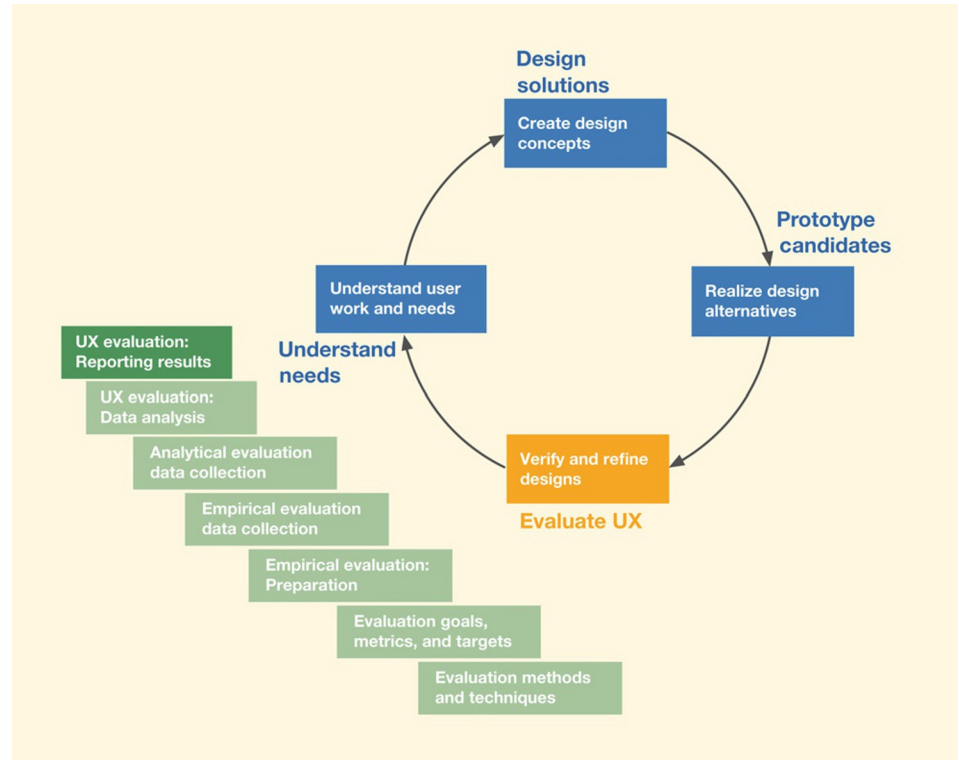
Fig. 27-1

You are here, in the chapter on the reporting subactivity within the Evaluate UX lifecycle activity, in the context of the overall Wheel lifecycle process.

problem descriptions will be too vague for designers who were not present for the evaluation.

To the project team, the report for an evaluation within an iteration is a redesign proposal. Hornbæk and Frøkjær (2005) show the need for usability evaluation reports that summarize and convey UX information, not just lists of problem descriptions by themselves.

All the effort and cost you invested thus far in UX evaluation can be wasted at the last minute if you don't follow up now to:

- Inform the team and project management about the UX problems in the current design.
- Persuade them of the need to invest even more in fixing those problems.

27.1.3 Participant Anonymity

We remind you, before we get into the details, that regardless of the kind of evaluation or reporting you are doing, you must preserve participant anonymity. You should have promised this on your informed consent form, and you have an

ethical, and perhaps a legal, obligation to protect it religiously thereafter. The necessity for preserving participant anonymity extends especially to evaluation reporting.

27.2 REPORTING DIFFERENT KINDS OF DATA

27.2.1 Reporting Informal Summative Results

If we collect and analyze quantitative data in UX practice, it will always be informal summative evaluation. When you refer to quantitative UX data in a report, it should never be associated with anything that even remotely looks like a statistical claim. This is a nonnegotiable professional and ethical requirement.

Beyond this caveat, because any quantitative data collected during UX evaluation are for the UX team only, our coverage of the reporting of informal summative results will be limited.

27.2.1.1 What if you need to convince the team to fix the problems?

What good is doing the UX evaluation if no one is convinced the problems you found are "real" and, as a result, the design does not get changed? It may be part of the job of UX engineers to convince others in the project team to take action about poor UX, as revealed by UX evaluation. This part of the role is especially important in large organizations where people who collect data are not necessarily the same people (or even people who have a close working relationship with them) as those who make the decisions about design changes.

In your team evaluation report, you may want to explain why some of the changes are necessary. If statistical "proof" is requested, however, you will not be able to provide it. At its extreme, this kind of request could be indicative of an unfortunate management or organizational problem showing a lack of trust that the UX team can do its job.

27.2.2 Reporting Qualitative Results—The UX Problems

All UX practitioners should be able to write clear and effective reports about problems found but, in their "CUE-4" studies, Dumas, Molich, and Jeffries (2004) found that many cannot. They observed a large variation in reporting quality over teams of usability specialists, and that most reports were inadequate by their standards.

If you use rapid evaluation methods for data collection, it is especially important to communicate effectively about the analysis and results because this kind of data can otherwise be dismissed easily "as unreliable or inadequate to

> ### Informal Summative Evaluation
> A quantitative summative UX evaluation method that is not statistically rigorous and does not produce statistically significant results (Section 21.1.5.2).

inform design decisions" (Nayak, Mrazek, and Smith, 1995). Even in empirical evaluation, though, the primary type of data from formative evaluation is qualitative, and raw qualitative data must be skillfully distilled and interpreted to avoid the impression of being too "soft" and subjective.

27.2.2.1 Common Industry Format for reporting

We don't include formal summative evaluation in typical UX practice, but the US National Institute of Standards & Technology (NIST) did initially produce a Common Industry Format (CIF) for reporting formal summative UX evaluation results.

Following this initial effort, the group—under the direction of Mary Theofanos, Whitney Quesenbery, and others—organized two workshops in 2005 (Theofanos, Quesenbery, Snyder, Dayton, and Lewis, 2005), these aimed at a CIF for formative evaluation reports (Quesenbery, 2005; Theofanos and Quesenbery, 2005).

In this work, they recognized that, because most evaluations conducted by usability practitioners are formative, there was a need for an extension of the original CIF project to identify best practices for reporting formative results. They concluded that requirements for content, format, presentation style, and level of detail depended heavily on the audience, the business context, and the evaluation techniques used.

While their working definition of "formative testing" was based on having representative users, here we use the slightly broader term "formative evaluation" to include usability inspections and other methods for collecting formative usability and user experience data.

27.3 REPORT AUDIENCES

As Theofanos and Quesenbery (2005) say, choices about content, format, vocabulary, and tone are all about the relationship between the author and the audience. Authors of the 2005 UPA Workshop Report on formative evaluation reporting (Theofanos et al., 2005) stressed different reporting requirements for different business contexts and audiences.

27.3.1 Reporting to Inform Your Project Team
27.3.1.1 Convey UX problem results clearly

The primary audience for a report of UX problem details is your own project team—the designers and implementers who will fix the problems. The key goal is to convey results and product implications clearly and meaningfully to your

---

Formal Summative Evaluation

A formal, statistically rigorous summative (quantitative) empirical UX evaluation that produces statistically significant results (Section 21.1.5.1).

Inspection (UX)

An analytical evaluation method in which a UX expert evaluates an interaction design by looking at it or trying it out, sometimes in the context of a set of abstracted design guidelines. Expert evaluators are both participant surrogates and observers, asking themselves questions about what would cause users problems and giving an expert opinion predicting UX problems (Section 25.4).

workmates, informing them about UX flaws in the design and/or informally measured shortcomings in user performance with the purpose of understanding what needs to be done to improve the design in the next iteration. Your report can usually be short and to the point, with little need for embellishments.

27.3.1.2 Meet with UX team and software developers in person

You should try to meet with the UX team and software developers to present the results personally, so you can explain points and answer questions.

Start with a "boilerplate" summary of the basics, including evaluation goals, methods, and UX targets and benchmark tasks used. Screen shots illustrating actual problem encounters are always good for selling your points about problems.

Your audience will expect you to prioritize your redesign recommendations, and cost-importance analysis is a good way to do this. Assuming that your team has technical savvy, use tables to summarize your findings; don't make them plow through a volume of text for the essence. If your development schedule is short, and things are already moving fast, keep the report and your problem list short.

27.3.2 Explaining UX Evaluation to Stakeholders

Because your goal is to persuade stakeholders of the need to invest time and cost into taking action to fix problems discovered, you must include your audience in the process and reasoning that led from raw data to conclusions, so that your recommendations don't appear to be pulled out of the air. To include them in your process, you must explain the process.

This kind of audience requires a different kind of report from all the others. It is more like an evaluation report contained within a more general presentation about UX evaluation. First, you have to establish your credentials and credibility and gain their engagement.

The goals for reporting to this kind of audience include (more or less in this order):

- Engender awareness and appreciation.
- Teach concepts.
- Sell buy-in.
- Present results.

Building rapport. Start on the first goal by building rapport and empathy. You want to get them to appreciate the need for usability and a good user experience to

Cost-Importance Analysis

An approach to prioritizing the time and effort of fixing UX problems found in UX evaluation based on priority ratios, which are calculated by dividing the importance of making a change by the cost (Section 26.4).

appreciate the value of these things to them and their organization. This is basically a motivation for UX based on a business case.

Educating the audience. The next goal of your presentation is teaching, explaining terminology and concepts. Help them understand how to view evaluation results as a positive thing, an opportunity and a means to improve.

Persuading and selling UX. You want to get their buy-in to the idea of doing UX. You want them to want to include a UX component in their overall development process (and budgets and schedules).

27.3.3 Reporting to Inform and/or Influence Management

Reports to management have to be short and sweet. Be concise and get to the point. Start with an executive summary. Briefly explain the process, the evaluation goals, the methods, and the UX targets and benchmark tasks used. Because this can be counted as at least a partly "internal" audience, you can share high-level aspects of informal quantitative testing (e.g., user performance and satisfaction scores), but just trends observed, not numbers and no "claims," and remember not to call it a "study."

Define your priorities and relate them directly to business goals. This is easier if you used UX targets driven by UX goals, based on business and product goals. You need an "explicit connection between the business or test goals and the results" (Theofanos and Quesenbery, 2005; Theofanos et al., 2005). The team's key goal for reporting to this audience is to influence and convince them that this is part of the process and that the process is working.

Use cost-importance analysis to focus on UX problems that can be fixed within the number of people hours allocated in the budget, but paint a complete picture of your findings. Screen shots illustrating actual problem encounters might be useful in engaging them in the whole evaluation scene.

27.3.4 Reporting to Customer or Client

If your team did not create the design you evaluated, it is best to not start by hitting the client square on with what is wrong with the system. This audience needs first to understand the whole concept of designing for UX and the methods you use and how they help improve the product. You have to tell them tactfully that their baby is ugly, but can be fixed.

Explain that finding UX problems in evaluation is simply part of a process focused on *fixing* problems. Be selective and describe a few examples. Clients and customers will not want to hear that there is a whole list of problems with the design of their system. For clients, UX problems are best described with scenarios

and screen shots that tell stories of how design flaws affected users, and how your UX engineering process finds and fixes those problems.

27.4 REPORT CONTENT

In this section, we cover the different types of content that could go into a UX evaluation report.

27.4.1 Individual Problem Reporting Content

Many researchers and practitioners have suggested various content items that might prove useful for problem diagnosis and redesign. The idea is to provide all the essential facts a designer will need to understand and fix the problem. Of course, at this point, the evaluators would have had to collect sufficient data to be able to provide all this information. The basic information needed includes:

- The problem description.
- A best judgment of the causes of the problem in the design.
- An estimate of its severity or impact.
- Suggested solutions.

In the first of these items, be sure to describe each problem as a problem, not as a solution. Because the problems were experienced by users doing tasks, describe them in that context—users and tasks and the effects of the problems on users. This means saying, for example, "Users could not figure out what to do next because they did not notice the buttons," instead of "We need flashing red buttons."

The second item, the engineering judgment of the causes of the problem in the UX design, is an essential part of the diagnosis of a UX problem and perhaps the most important part of the report. Because the flaw in the design is what needs to be fixed, you should connect it with the appropriate design guidelines and/or heuristic violations as much as possible in terms of UX issues and principles.

Next is an estimate of severity or importance in terms of the impact on users. To be convincing, this must be well reasoned. Finally, to help designers act to fix the problems, recommend one or more possible design solutions, along with cost estimates and tradeoffs for each, especially if a solution has a downside. To justify the fixes, make compelling arguments for improved design and positive impact on users.

*Pyramid of
User Needs*

An abstract representation as a pyramid shape with the bottom layer as ecological needs, the middle layer as interaction needs, and the top layer as emotional needs (Section 12.3.1).

Emotional Impact

An affective component of user experience that influences user feelings. Includes such effects as enjoyment, pleasure, fun, satisfaction, aesthetics, coolness, engagement, and novelty and can involve deeper emotional factors such as self-expression, self-identity, a feeling of contribution to the world, and pride of ownership (Section 1.4.4).

*Cost-Importance
Analysis*

An approach to prioritizing the time and effort of fixing UX problems found in UX evaluation based on priority ratios, which are calculated by dividing the importance of making a change by the cost (Section 26.4).

There are many other kinds of information that can be useful in a UX problem report content, including an indication of how many times each UX problem was encountered (by each user and by all users) to help convey its importance.

27.4.2 Give Some Coverage of the Ecological and Emotional Layers of the Needs Pyramid

Most of the UX problems you will find and report will occur at the interaction level. But you should also give some attention to the ecological and emotional layers. If there are problems with the design at the ecological level, especially problems with the conceptual design, be sure to stress the importance of this level of design. Unless the design is successful at the ecological level, evaluation at the other levels won't ever make it a good design.

Special discussion should be directed to reporting emotional impact problems, as those problems can be the most important for product improvement and marketing advantage, but these problems and their solutions can also be the most elusive. Emotional impact problems should be flagged as a somewhat different kind of problem with different kinds of recommendations for solutions.

Provide a holistic summary of the overall emotional impact on participants. Report specific positive and negative highlights with examples from particular episodes or incidents. If possible, try to inspire by comparing with products and systems having high emotional impact ratings.

27.4.3 Include Cost-Importance Data

Usually, cost-importance analysis is considered part of the nitty-gritty engineering details that would be beyond the interest or understanding of those outside the UX team and its process. However, cost-importance analysis, especially the prioritization process, can be of interest to those who have to fix the problems and those who have to pay for it.

Importance ratings and supporting rationale can be helpful in convincing designers to fix at least the most urgent problems. The cost-importance table, plus any discussion supporting the choice of table entries, will tell the story.

27.5 REPORT MECHANICS

27.5.1 Consistency Rules

Consistency in reporting UX problems is important for all audiences. Evaluation reports are, above all, a means of communication, and understanding is hampered by wildly varying vocabulary, differences among diagnoses and

descriptions of the same kinds of problems, the language and style of expression in UX problem descriptions, and level of description contained (e.g., describing surface observables versus the use of abstraction to get at the nature of the underlying problem). Establishing your own standards for reporting results helps control broad variation in content, structure, and quality of UX problem reports.

Abstraction

The process of removing extraneous details of something and focusing on its irreducible constructs to identify what is really going on, ignoring everything else (Section 14.2.8.2).

27.5.2 Reporting Vocabulary
27.5.2.1 Precision and specificity

You are communicating with others to accomplish an outcome. To get the audience to share the vision of that outcome or to even understand what outcome you want, you need to communicate effectively; perhaps the first rule for effective communication is to be precise and specific. It takes more effort to write effective reports.

Sloppy terminology, vague directions, and lazy hand-waving are likely to be met with indifference, and the designers and other practitioners are less likely to understand the problems and solutions we propose in the report. This kind of effect of a problem report on our audience usually results in their being unconvinced that there is a real problem.

So, instead of saying a dialogue box message text is hard to understand and recommending that someone write it more clearly, you should, in fact, make your own best effort at rewording to clarify the text and say why your version is better. The criterion for effectiveness is whether the designer who receives your problem report will be able to make better design choices (Dumas et al., 2004).

27.5.2.2 Jargon

As UX professionals, we, like most others in technical disciplines, have our own jargon. But, as UX professionals, we must also know that our UX problem reports are like "error messages" to designers, and that guidelines for error message design apply to our reports as well. And one of those guidelines about messages is to avoid jargon.

So, while we might not put jargon in our UX designs, we might well be tempted to use our own technical language in our reports about UX. Yes, our audience is supposed to include UX professionals and UX designers, but you cannot be sure how much they share your specialized vocabulary. Spell things out in plain natural language.

27.5.3 Report Tone

The British are too polite to be honest, but the Dutch are too honest to be polite.

– Candid Dutch saying

All your audiences deserve respect in evaluation reports. In reporting to customers and clients, most UX professionals appreciate the need to temper reports with restraint. But even your own team should be addressed with courtesy.

27.5.3.1 Respect feelings

Whether the design you evaluated was created by your own team or by others, treat the work with respect. Don't attack. Don't demean. Don't insult. Your goal is not to get designers angry or resentful, but to get them to act on the report and fix the problems. As Dumas, Molich, and Jeffries put it: "Express your annoyance tactfully."

Some evaluators believe that design flaws must be stated in strong language to convey the message. But designers say they are insulted by emotional rants and that, "Being blunt is not helpful; it is simply rude" (Dumas et al., 2004).

27.5.3.2 Accentuate the positive and avoid blaming

Most UX professionals do realize that they should start with good things to say about the system being evaluated. However, even when encouraged to be positive, some practitioners in studies (Dumas et al., 2004) proved to be reticent in this regard. This may be because their usual audience is the project team who just want to know what the problems are so that they can start fixing them.

However, even if the report is mainly critical, it is best to start with *something* positive. Include information about places where participants did not have problems, where they were successful in task completion, and where users expressed great satisfaction or joy of use. Stories of good things happening can start things off with very positive feelings. The rest is, then: "We are on a roll: How can we make it even better?"

Present reports about design flaws as opportunities for design improvement, not as a criticism. A good way to do this is to remind them that the *goal* of formative evaluation is to find problems so that you can fix them. Therefore, a report containing information about problems found is an indication of success in the process. Congratulations, team; your process is working!

27.5.4 Reporting on Large Amounts of Qualitative Data

If you are reporting on a large amount of evaluation, about a large number of UX problems, you need to be well organized. If you ramble and jump around among different kinds of problems without an integrated perspective, it will be like a hodgepodge to your audience, and you will lose them, along with their support for making changes based on your evaluation.

One possible approach is to use an affinity diagram technique (a hierarchical diagram of evaluation notes, a bottom-up technique for organizing lots of

disparate pieces of evaluation data, Section 8.7.1). We showed how to use an affinity diagram to organize work activity data, and you can use the same technique here to organize all your UX problem data for reporting. Post notes about each problem at the detailed level, and group them according to commonalities and categories, for example, with respect to task structure, organization of functionality, or other system structure.

27.5.5 Your Personal Presence in Reporting

Don't just write up a report and send it out, hoping that will do the job. If possible, you should be there to make a presentation when you deliver the report. The difference your personal presence at the time of reporting can make in reaching your goals, especially in influencing and convincing, is inestimable. Nothing beats face-to-face communication to set the desired tone and expectations. There is no substitute for being there to answer questions and head off costly misunderstandings. If the audience is distributed geographically, this is a good time to use videoconferencing or at least a teleconference.

Exercise 27-1: UX Evaluation Reporting for Your System

Goal: Write a report of the formative UX evaluation you did on the system of your choice.

Activities: Report on your informal summative evaluation results using a table showing UX targets, benchmark tasks, questionnaires, and so on used to gather data, along with target values and observed values.

Add brief statements about whether or not each UX target was met.

Write a full report on a selected subset (about half a dozen) of UX problems found in the qualitative part of your formative UX evaluation. Follow the guidelines in this chapter regarding content, tone, and format, being sure to include redesign proposals for each problem.

Report on the results of your cost-importance analysis, including problem resolutions, for all the problems you reported previously and, if appropriate, some others for context.

Deliverables: Your formative evaluation report.

Schedule: We expect this exercise to take about an hour.