Extracting "Broken Expectations" from Call Center Records: *Why* and *How*

Aylin Koca
Faculty of Industrial Design
Eindhoven University of Technology
P.O. Box 513
Eindhoven, NL-5600 MB
The Netherlands
A.Koca@tue.nl

Aarnout C. Brombacher
Faculty of Industrial Design
Eindhoven University of Technology
P.O. Box 513
Eindhoven, NL-5600 MB
The Netherlands
A.C.Brombacher@tue.nl

CHI 2008, April 5 – 10, 2008, Florence, Italy.
ACM 978-1-60558-012-8/08/04.

Abstract
Currently, despite the explicit industrial consideration to improve the appeal and usability of *technically sound* electronics products, users increasingly seem to have dissatisfactory experiences in interacting with them. These unforeseen experiences (attributable to specifications omissions, usability/learnability problems, or specific usage context) lead to a large and increasing share of *unknown* field complaints. To correct and prevent such complaints or user reports, we promote effective exploitation of call centers: Valuable usage data is retrievable from the field by adopting a user-centered failure classification model, which we developed. We also report on the supporting results of a test from applying our model to a set of call center data.

Keywords
Business value of usability, usage-failure classification, soft reliability

ACM Classification Keywords
H5.2. User Interfaces: Evaluation/methodology.

Introduction
Relevant user and context information is of topmost importance in designing for easy- and pleasurable-to-use products. Premeditated tests before official product release are bound to be partial in revealing the many

use cases observed in the field. The reality is about actual consumers, in whichever country and specific use context, experiencing the product over time. If *early* feedback from actual consumers can be captured fast enough, any gap between the product and consumer expectations can be bridged more effectively. Therefore, call centers, where consumers make contact with to e.g. get guidance about problems, report contentment, or complaints about products, are promising sources of feedback information.

Our investigations at several call centers of a multinational consumer electronics company [8] showed that the data collected from the field is mainly logistics driven, i.e. aiming to capture information about broken hardware and software, but not "broken expectations" of users [1]. Hence, the current way of identifying and logging user feedback misses out the user-centered view that is valuable input for the stakeholders in product development (e.g. designers, developers, managers). We observed that, in a special in-house effort to rapidly make use of field feedback data in the case of really new products, time-costly and expert-dependent (i.e. subjective) post-processing of call center data is carried out. However, such product-specific processing is rather costly, still misses out the user-centered view and is not scalable across projects.

In the context of a large multidisciplinary research project with actively collaborating industrial partners and with support from the government, we developed a user-centered failure classification model (Figure 3) that addresses the aforementioned issues [7, 8]. This model has already been tested numerous times with large sets of real field- and test-data of various high-tech electronics products. In this paper, we delimit our scope to sharing some founding ideas of our research and to reporting intriguing findings from the application of our model on a particular set of call center data. More specifically, we discuss -in order-:

- the kind of user reports at e.g. call centers that are in fact *not* due to a malfunction of a product, and the outcome of current practices to deal with them,
- the relevant part of our model as an improvement for the treatment of such reports and for their effective exploitation to support design activities,
- some of the test results (from applying our model) that identify the types of such reports, and
- the overall potential of call center data for gathering relevant usage information.

Soft Reliability in a Nutshell

Our ongoing research on soft reliability [6, 7, 8] deals with the so called "broken expectations" of users [1], i.e. field cases where products are being returned, or being sought redress for, that in fact *technically function well according to specifications*. Such cases may be attributable to product-specifications omissions, usability and learnability problems, customer misunderstandings, or specific usage context.

As shown in Figure 1, soft reliability depends on the conformance of the actual product to individual user's requirements, whereas hard reliability depends on the conformance of the actual product to its formal technical specifications. Since there is *one* set of technical specifications per product, but *many* users with *individual* user expectations, hard reliability can be

defined as a one-to-one relationship, while soft reliability as one-to-many. Moreover, each user has dynamically changing (*explicit* and *latent*) expectations as designated by the Kano model [5] (Fig. 1), making soft reliability management especially challenging.

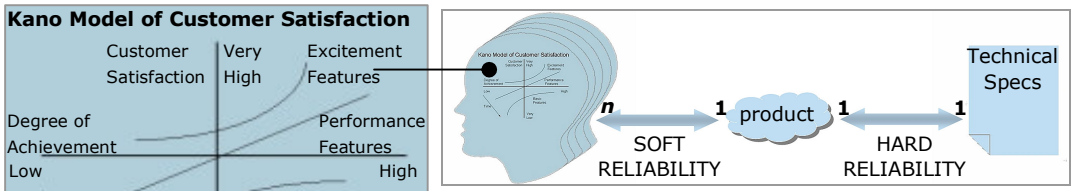

Figure 1. Soft reliability is about "making the right product", while hard reliability is about "making the product right". One does not entail the other.

Current logistics-driven field feedback collection is not tailored to capture incidents about soft reliability. User reports are typically tested for hard reliability problems. Therefore, such incidents lead to large and growing numbers of "product assistance" calls at helpdesks [8], "no fault found" labeled products at service/repair centers, returned products at dealers that function well, and to top it all, damaged brand image of companies.

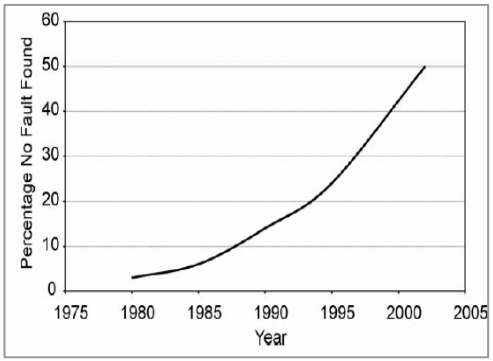

Figure 2. Rapid growth of "no fault found" cases in modern high-volume consumer electronics in time [3]

User-Centered Failure Classification Model

In order to systematically utilize valuable field feedback information to better meet user expectations and to tackle the growing problem of currently unknown "no fault found" cases, we developed the user-centered failure classification model in Figure 3 [7]. Taking each individual user-feedback (e.g. question, complaint, remark) as the unit of analysis, this model intends to provide an efficient and effective means to link problems to their respective originator activities in a product development process, in the form of actionable items to work with, for improvement. Furthermore, the model is built to be generic and hence extendible for different (i) products, (ii) sources of usage data, (iii) audiences involved in product development, (iv) model implementing tools and coding mechanisms (as in [4]).

Due to the scope of this paper, we only focus on a part of our depicted model here, which offers the typically missing account for soft (versus hard) *product* failures, besides all other types of *related* failures[1]. The model recognizes hard failures as product failures where the product is incapable of performing its functions as listed in its technical specifications without the intervention of authorized technical support for recovery by means of repair or replacement of parts. On the other hand, soft failures are recognized as product failures where the product, despite being capable of performing its functions as listed in its technical specifications, still necessitates professional intervention for recovery (but not repair) through instructions or information from an unexpected user-product interaction state.

1 N.B. Every customer report to a company-official in the form of questions, complaints, etc. is each considered a "failure", to be used as feedback for improving related products and services.

product-level (i.e. product capability based) distinction of soft failures into two sub-classes

user-level (i.e. use experience based) distinction of soft failures into six sub-classes

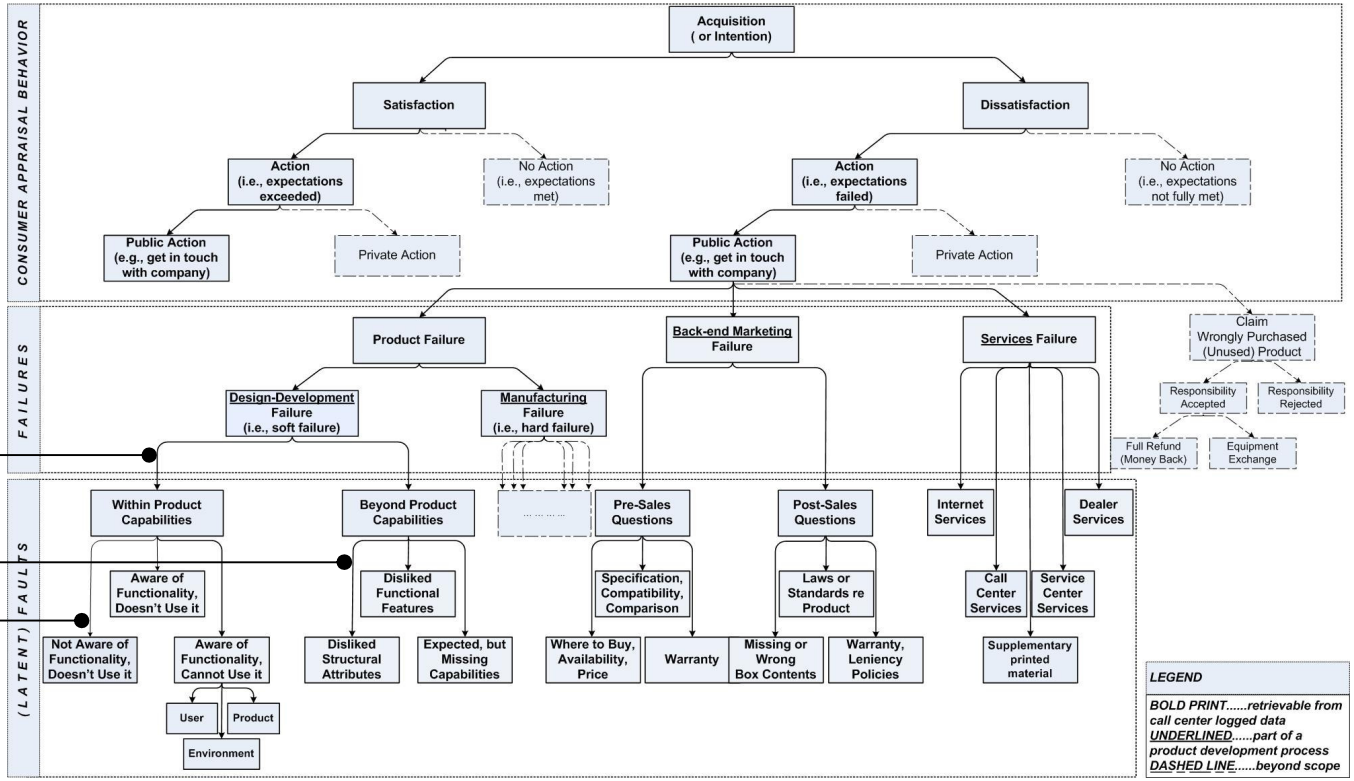

Figure 3. User-centered failure classification model

Soft failures are sub-classified first at the product-level, before they are classified at the user-level (Figure 3). The product-level classification makes a distinction between (i) problems that can be resolved *within* the capabilities of the product by the user upon getting supporting guidance and instructions; versus (ii) those due to users' higher or other expectations that are *beyond* what the product is capable of, and hence can only be allegedly addressed -for the moment- by providing (the missing) information to the user in order to compensate for the discrepancy. This distinction corresponds to the widely acknowledged "errors of omission" and "errors of commission" as referred to by Krippendorff [9], in that order. Consequently, failures captured at this level indicate if the product can be improved by (i) *adjusting* current capabilities so that they are easily e.g. noticeable, inviting, and accessible in general; (ii) *enhancing* current capabilities or *adding* new capabilities to eliminate particular explicit user disillusionment.

The user-level classification of soft failures is based on different "phases of use" [2] that need to be consecutively passed through to ensure satisfied users who take full advantage of all capabilities of the product. These phases include *awareness* of capabilities, *motivation* to use, *orientation* for figuring out how to use, *adoption* in daily life, and *incorporation* for extended use. Failures captured at this level indicate weaker usage aspects of a product, likely to prevent its successful communication to its users, ultimately leading to poor acceptance and adoption.

The part of our model that we discussed here, offers relevant classes to work with, as an eloquent replacement to the currently named "no fault found" category of failures. Since the unit of analysis of this model is an individual user-feedback each of which is registered as freetext summary, further product and feature specific information can be retrieved from each classified feedback, either manually, or automatically by text mining. This can be done most effectively by first observing the failure distribution proportions over all classes of our model, and hence identifying the largest proportions which reveal the weakest aspects of a product that require the most attention. As a result, prioritization among necessary improvement points can be made, in connection with corporate priorities.

Model Test & Results

A specific test of our user-centered failure classification model on user-feedback data collected at a call center of a multinational consumer electronics company revealed some intriguing findings in terms of failure distribution proportions.

The data comprises user reports from the UK and Germany, collected during the first half of 2007 (i.e. 3 January-4 July), about a newly released DVD recorder. The reports were initiated by emails from users and handled over time with both email and telephone exchanges between call center agents and users. The complete dataset has 244 "service requests" initiated by users, as registered in the database, an equivalent to 336 individual user-feedback relating to failures.

The result of classifying 336 failures showed that 251 were product related failures. From 251 product failures, *74% could be identified as soft failures*, whereas 8% could be identified as hard failures. In other words, product design-development failures occur 9 times more than product manufacturing failures.

Furthermore, the result of sub-classifying all 186 soft failures showed that 55% were setbacks *within* current product capabilities, while 45% were due to users' higher or other expectations that are *beyond* what the product is capable of.

Discussion

To keep developing desired products on the competitive edge, voice-of-the-customer and hence user-centeredness is a key factor: According to Kano model of customer satisfaction, the trend is that "what (capability) was considered exciting yesterday, becomes asked for today and expected tomorrow".

We presented call center data as a beneficial and natural resource for getting *recent* feedback data from *real* users in *large* quantities. Due to direct contact with customers, it is also possible to collect *relevant* information about their experiences and expectations. Furthermore, since all contacts with customers are registered over time, it is possible to track how their product-experience and sources of dissatisfaction change over time, by analyzing their respective data.

Currently call centers function mainly with the purpose to serve many customers, but not to help improving the quality related information flows in-house. Optimal ways to ensure both seems possible with structured and generic tools such as our classification model.

Acknowledgements

This work is being carried out as part of the "Managing Soft Reliability in Strongly Innovative Product Creation Processes" project (http://softreliability.org/) since October 2005, sponsored by the Dutch Ministry of Economic Affairs under the IOP-IPCR program.

References

[1] Bly, S., Schilit, B., McDonald, D.W., Rosario, B. and Saint-Hilarie, Y. Broken Expectations in the Digital Home. *Ext. Abstracts CHI 2006*, ACM Press (2006), 568-573.

[2] Bouwmeester, K. den and Bosma, E. Phases of Use: A means to Identify Factors that Influence Product Utilization. *Ext. Abstracts CHI 2006,* ACM Press (2006), 117-122.

[3] Brombacher, A.C., Sander, P.C., Sonnemans, P.J.M. and Rouvroye, J.L. Managing Product Reliability in Business Processes 'Under Pressure'. *Reliability Engineering & System Safety 88*, 2 (2005), 137-146.

[4] Gorlenko, L. and Englefied, P. Usability Error Classification: Qualitative Data Analysis for UX Practitioners. *Ext. Abstracts CHI 2006*, ACM Press (2006), 803-808.

[5] Kano, N., Seraku, N., Takashi, F. and Tsuji, S. Attractive Quality and Must-Be Quality. *The Journal of Japanese Society for Quality Control 14*, 2 (1984), 39-48.

[6] Koca, A., Lu, Y., Brombacher, A.C. and Hartmann, J.H. Towards Establishing Foundations for New Classes of Reliability Problems Concerning Strongly Innovative Products. *Proc. ICMIT 2006*, IEEE (2006), 581-584.

[7] Koca, A., Schouwenaar, A.J.M. and Brombacher, A.C. Analysis of User-Centered Failure Mechanisms in New Product Development for Quality Improvement. *Proc. EurOMA 2007,* (2007).

[8] Koca, A., Schouwenaar, A.J.M. and Brombacher, A.C. Field-Feedback in Innovative Product Development: A Comparison of Two Industrial Approaches. *Proc. IPDM 2007*, (2007), 657-668.

[9] Krippendorff, K. On the Essential Contexts of Artifacts or on the Proposition That 'Design is Making Sense (Of Things)'. *Design Issues 5*, 2 (1989), 9-39.