INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN
ICED 05 MELBOURNE, AUGUST 15-18, 2005

HOW COMPANIES LEARN FROM DESIGN FLAWS: RESULTS FROM AN EMPRICAL STUDY OF THE GERMAN MANUFACTURING INDUSTRY

B. Gries, K. Gericke and L. Blessing

Keywords: Learning from design failures, design flaws, empirical study

1 Introduction and objectives

Design flaws often become apparent at a time when the product is already in use and its development process, which in many cases includes extensive testing of parts, components and prototypes, is considered complete. Such flaws (see 2.1 for a definition) may reach from poor ergonomics to the total failure of the product. Often, especially when user safety is at risk, design flaws are so severe that companies are forced to announce a product callback. However, the occurrence of a design flaw is always an indication that any methods intended to prevent design flaws – e.g. Failure Mode and Effect Analysis (FMEA) – have failed to some extent.

Petroski suggests that many (if not most) products, which we are familiar with today, have a long history of previously flawed designs [3]. This implies that designers did indeed learn from design flaws in both senses of the word “learn”: discovering the flaw and utilizing the knowledge gained about it to find a solution.

As far as discovering a design flaw is concerned, it can be assumed that the feedback from those who interact with the physical products in practice – the individuals who maintain, repair, recycle but essentially use the products – plays an important role. In their previous work, the authors pointed out that this feedback information could not only be vital for identifying potential product hazards but helps designers to review the effects of their design measures and therefore to improve their products from generation to generation [4].

In order to obtain a better understanding of how designers learn from design flaws, a mail survey was conducted that aimed at investigating company-, process- and product-related factors of this phenomenon and to answer (among others) the following research questions:

- To what extent are design flaws of a company’s (or a competitor’s) product a driving force in the development of new products?
- How do the designers of a company become aware of design flaws of their products?
- How successful are companies in correcting design flaws?
- How do successful and unsuccessful companies differ in terms of size, activity profile of their designers and characteristics of their products?
- What are possible factors that influence the success in correcting a design fault?

2 Study design

2.1 Questionnaire construction

Based on the abovementioned research questions, hypotheses were developed and 28 questions formulated to test them. A questionnaire was designed in which those 28 questions were arranged in three sections.

In section I, the participants were asked to give some general information about their company (number of employees, annual sales, etc.) and to define an activity profile of its designers.

Section II was titled "Questions about the product and its development". Participants were instructed to refer all following questions to a single product which is the result of the development process that they are most familiar with and which is already available on the market (subsequently referred to in the questionnaire as PRODUCT).

In section III, the participants were asked questions about the most severe design flaw of the PRODUCT which was defined as "an unwanted behavior of the PRODUCT which is for the most part caused by its design". The subjects were advised that the study was based on the concept that according to this definition, any product is flawed to a certain extent (be it, that it leaves room for optimization). Again, the participants were instructed to refer all subsequent questions to one and the same flaw (referred to as DESIGN FLAW).

By default, the questions were asked in a "check all that applies" format. Questions that required the participants to choose only one option were clearly marked as such. Quantitative information was to be given by entering the value instead of selecting a category (i.e. 1-50, 51-100, etc.). Open questions were avoided altogether. Following common practice [9], an early draft of the questionnaire was tested on five representatives from industry in order to identify questions that were misinterpreted and/or felt difficult to answer, questions to which the participants were inclined to refuse an answer (a particularly interesting point, given the rather delicate topic of the study) and to pick up suggestions for further questions. In addition, the average time for completing the questionnaire was taken, which was about twelve minutes and felt tolerable by all subjects.

In addition to the paper version of the final questionnaire (which was printed on light green paper in form of a twelve-page A5 booklet), an online version was implemented which can still be viewed at www.ktem.tu-berlin.de/umfrage (enter 'iced2005' as password).

2.2 Survey implementation

The survey population encompassed all German companies whose economic activities can be described with NACE codes [6] from 28.6 (et seqq.) to 37, which include the manufacture of:

- machinery and equipment (29 et seqq.),
- office machinery and computers (30 et seqq.),
- medical, precision and optical instruments (33 et seqq.),
- motor vehicles (34.1 et seqq.) as well as
- aircraft and spacecraft (35.3 et seqq.).

The sample frame was defined by all firms fitting into the above category range listed in the 2003 editions of the “Hoppenstedt” company databases for small and medium [7] as well as large companies [8]. From this list of 18,196 companies, a random sample of 1,000 firms was drawn.

By referring to the information from the database and subsequent research on the internet, 794 recipients of the questionnaire could be identified by name. Since only a small minority could doubtlessly be recognized as designers (the survey’s target group), most of the addressees were owner-managers, CTOs and similar members of upper management. Consequently, the cover letters contained a passage in which the recipients were asked to forward the enclosed questionnaire to “a person in a leading position who is most familiar with a specific product and its development”. The letter also contained the URL and a password for the online version of the questionnaire. In addition to the hand-signed cover letter (and the questionnaire of course), the mailing included a franked return envelope and a ball pen as a small incentive. Two weeks after mailing, reminder letters with an enclosed replacement questionnaire were sent. After 50 working days, the survey was brought to an end.

2.3 Data evaluation

An important aspect of the study was not only to deliver individual results alone but also to analyze their relationship to success (cf. [5]). To make a distinction between successful and unsuccessful cases of correcting a design flaw, participants should decide in question III.8 which of the statements in Table 1 applies most. For the study, only cases in which answer option D is selected are defined as “successful”, whereas options A, B or C represent an “unsuccessful” case.

The confidence level of the study was set to $\alpha = 0.05$, which means that any found difference or correlation is only statistically significant for $p < \alpha$ where p is the probability that the finding is the result of chance. For a lack of space, however, details on the statistical methods used in this study to test the significance cannot be given (refer to e.g. [11]-[13]).

3 Results

In total, 173 responses were received. Ten companies refused to take part in the study and six were ceased, resulting in a response rate of 17.6%. About 84% of respondents preferred the paper questionnaire over its online counterpart. Table 1 reveals that almost 80% of cases are successful according to the definition in 2.3.

Table 1. Frequency and percentage of answers to question III.8 (only one answer allowed)

	Answer option	**Frequency / percentage**
A	The DESIGN FLAW has not been dealt with	13 / 7.6%
B	Attempts were made to develop a solution for the DESIGN FLAW.	13 / 7.6%
C	A solution for the DESIGN FLAW was developed which, however, was not or will not be implemented in the PRODUCT.	10 / 5.85%
D	**The DESIGN FLAW was corrected and the solution has been or will be implemented in the PRODUCT.**	**135 / 78.95%**

3.1 Successful and unsuccessful cases do not differ significantly in terms of company size

Figure 1 shows the distribution of variables that reflect the size of the companies in this study. Typical for the German business landscape, small companies dominate.

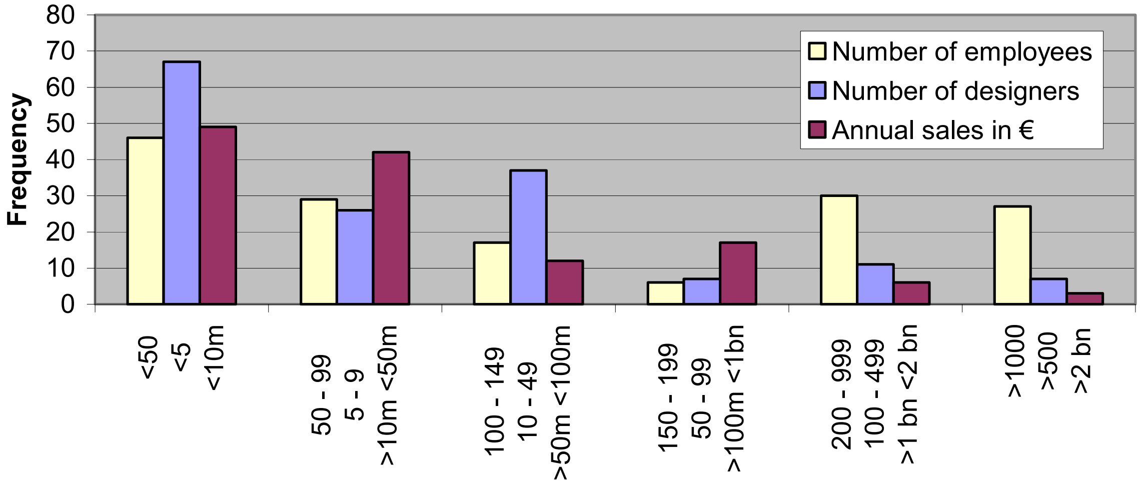

Figure 1. Frequency distribution of answers to questions I.1 through I.3. ($n_{I.1}$ = 155, $n_{I.2}$ = 155, $n_{I.3}$ = 129, open-ended questions)

In Table 2, the mean values of each variable are given for both the successful and unsuccessful group in comparison with the whole sample. Companies belonging to the successful group are apparently larger in terms of all listed variables. Yet, none of the variables differs significantly as the results of a Mann-Whitney-U test reveal.

Table 2. Differences of mean values of company size-related variables in successful und unsuccessful cases and their significance.

Variable	**Mean values**			**p-value (Mann-Whitney-U test)**
	Successful cases	Unsuccessful cases	Sample	
Number of employees	1118	356	981	0.079
Number of designers	126	13	106	0.070
Annual sales	€ 215.0m	€ 18.8m	€ 180.0m	0.069

3.2 Design flaws are a similarly important driver for the development of new products in successful and unsuccessful cases

Figure 2 shows the distribution of answers to question II.7 “Which were the most influential factors in the decision to develop the PRODUCT?”. Individually, the options “Flaws of own products” and “Flaws of competitor products” achieved average ranks, whereas combined, they could be regarded as the third most important factor, surpassed by “Lowering manufacturing costs” just barely.

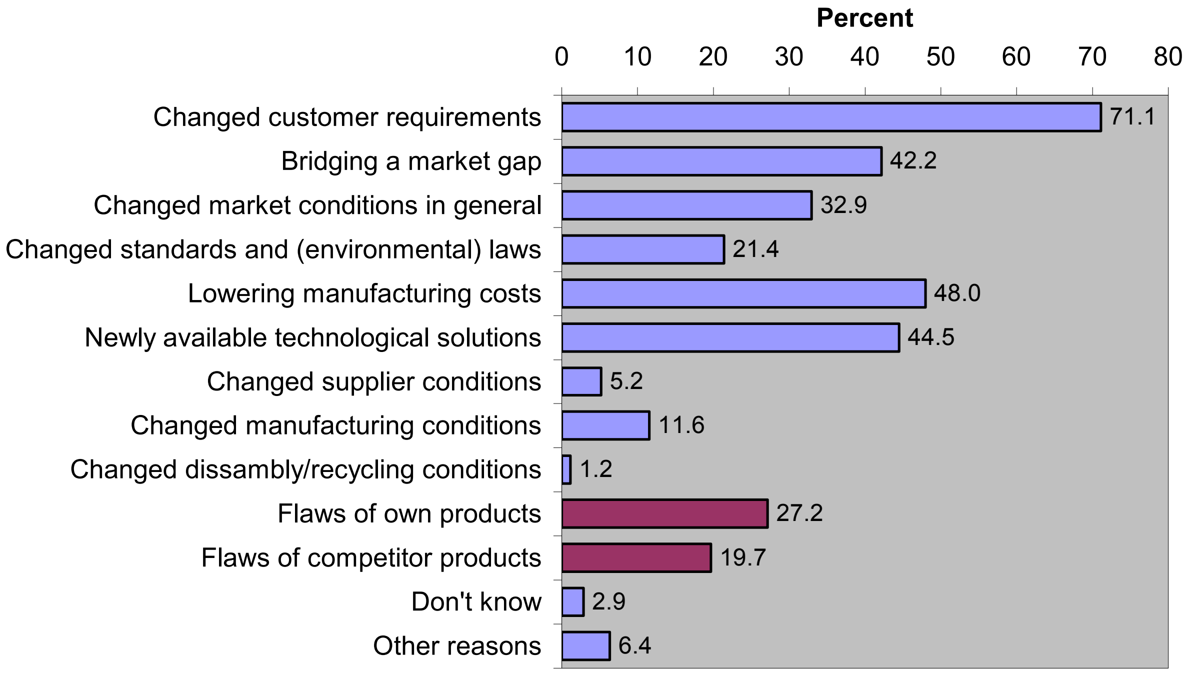

Figure 2. General percentage distribution of answers to question II.7 (n = 173, multiple answers allowed)

Table 3 shows no significant differences in the percentage of answer options in successful and unsuccessful cases whatsoever. The answer "Flaws of own products" was – as many other options – almost equally often selected in both groups. While there is a clearly higher percentage of the option "Flaws of competitor products" in the group of successful cases, this difference is still also insignificant.

Table 3. Percentages of answers in successful and unsuccessful cases and significance of their difference

Variable	**Percentages of answers**		**p-value (Mann-Whitney-U test)**
	Successful cases	Unsuccessful cases	
Changed customer requirements	69.6	80.6	0.196
Bridging a market gap	43.0	41.7	0.889
Changed market conditions	36.3	22.2	0.113
Changed standards and laws	21.5	22.2	0.924
Lowering manufacturing costs	48.1	50.0	0.884
Newly available technologies	48.9	30.6	0.060
Changed supplier conditions	5.9	2.8	0.454
Changed manufacturing conditions	12.6	8.3	0.481
Changed recycling conditions	1.5	0.0	0.464
Flaws of own products	**27.2**	**27.8**	**0.965**
Flaws of competitor products	**28.9**	**16.7**	**0.576**
Don't know	2.2	5.6	0.293
Other reasons	8.1	0.0	0.077

3.3 Successful and unsuccessful cases significantly differ in the degree to which designers are involved in certain activities

In order to identify the task profile of the company's designers, participants were given a list of activities that can roughly be assigned to the life cycle phases of product development, manufacturing and (after) sales. For each task, the subjects should choose from one of the following options:

A. Designers involved or responsible

B. Designers not involved or task outsourced

C. Task generally not undertaken by company

Figure 3 shows the results of question I.4. It becomes obvious that (with the exception of product management) designers are stronger involved in early phases of the product life.

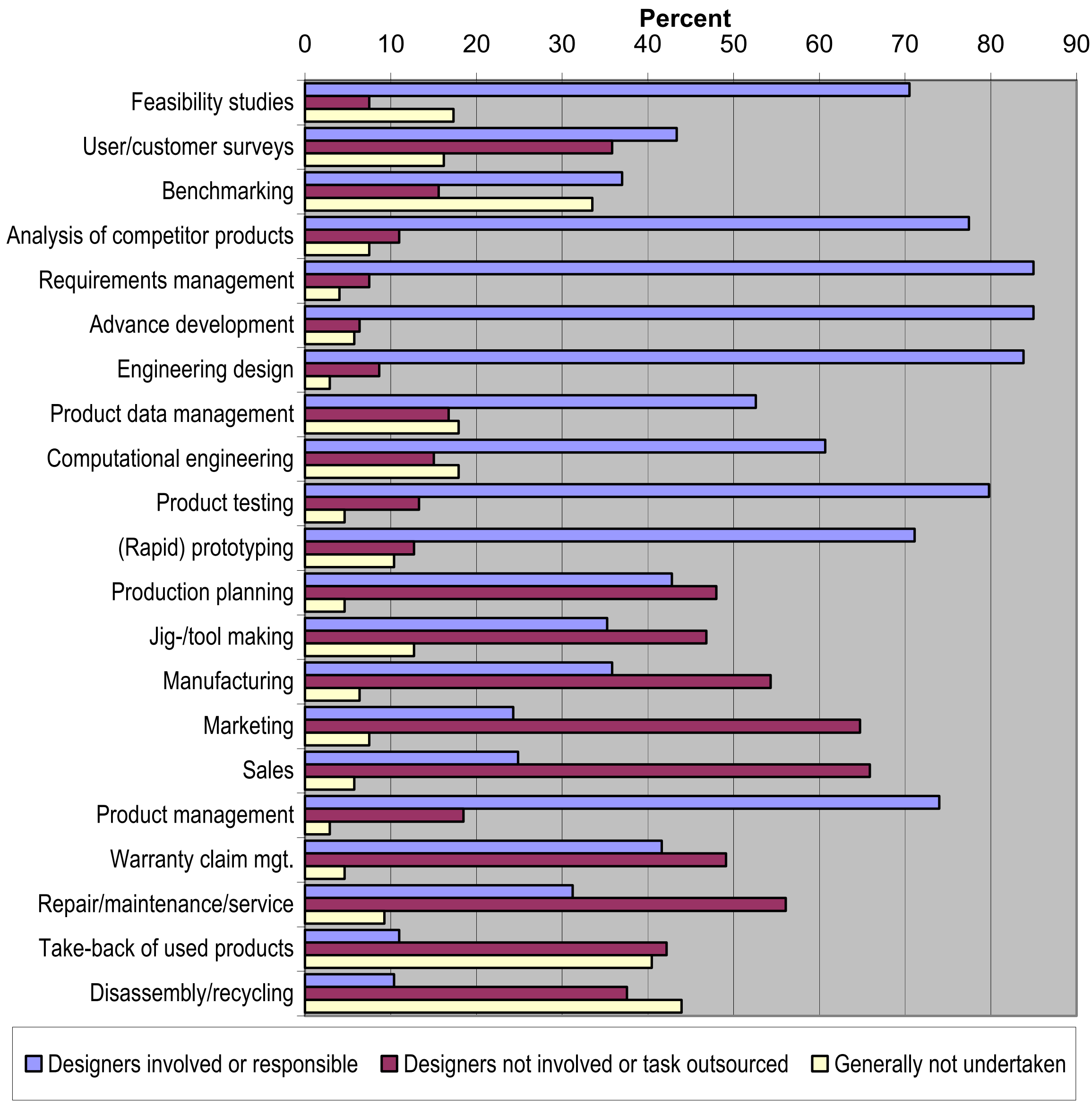

Figure 3. General task profile of designers (question I.4, n = 173, only one answer per option allowed; option "Don't know" omitted)

Table 4. Tasks with answer profiles that differ significantly (except "Take-back of used products"). Refer to text for option key.

Variable	Percentages of answer options						p-value (Mann-Whitney-U test)
	Successful cases			Unsuccessful cases			
	A	B	C	A	B	C	
Customer/user surveys	49.6	32.6	14.8	19.4	50.0	19.4	0.008
Requirements management	89.6	6.7	2.2	69.4	11.1	8.3	0.037
Product data management (PDM)	57.0	17.0	14.8	36.1	16.7	27.8	0.032
Computational engineering	65.9	16.3	14.1	41.7	11.1	30.6	0.022
Product management	79.3	15.6	2.2	58.3	25.0	5.6	0.044
(Take-back of used products)	12.6	45.2	37.8	5.6	30.6	50.0	0.056

Table 4 lists those tasks for which a Mann-Whitney-U test has revealed a significantly different answer profile. The most significant result is that in successful cases, designers are much stronger involved in customer/user surveys ($p < 1\%$). Also, it becomes obvious that lesser involvement in requirements management seems to be characteristic for unsuccessful cases. The only significant finding outside the phase of product development is that successful cases feature a higher percentage of designers engaged in product management.

3.4 In successful cases, testing and warranty claims are significantly more often stated as occasions on which the design flaw became known

To investigate possible information sources of design flaws, participants were asked in question III.6 "On which occasion did the DESIGN FLAW become known?". The figure below shows the distribution of answers.

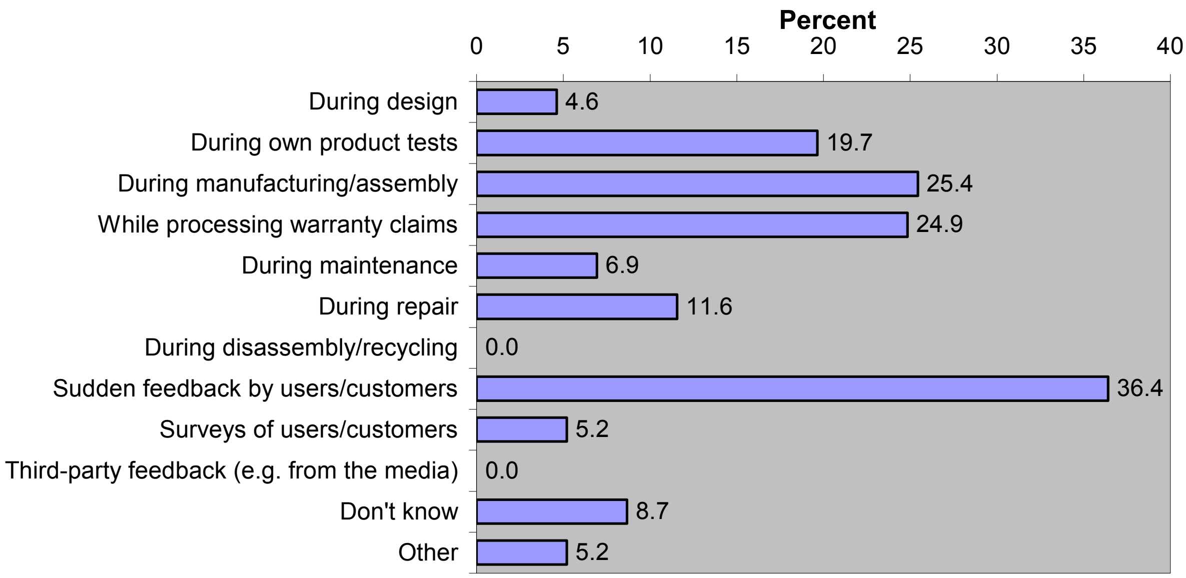

Figure 4. General percentage distribution of answers to question III.6 (n = 173, multiple answers allowed)

Table 5 reveals that only the higher percentages of the answers "Testing" and "Processing of warranty claims" in successful cases are significant. Interestingly, "Surveys of users/customers" was stated almost twice as often in unsuccessful cases; however, due to the overall low frequency of this answer option, this finding is insignificant. Noticeably, not a single DESIGN FLAW was revealed during disassembly/recycling or by third-party feedback.

Table 5. Percentages of answers in successful and unsuccessful cases (significant differences in boldface)

Variable	Percentages of answers		p-value
	Successful cases	Unsuccessful cases	(Mann-Whitney-U test)
During design	5.2	2.8	0.545
During own product tests	23.7	5.6	**0.016**
During manufacturing/assembly	28.1	16.7	0.163
While processing warranty claims	28.9	11.1	**0.029**
During maintenance	7.4	5.6	0.700
During repair	11.9	11.1	0.902
Sudden feedback by users/customers	38.5	30.6	0.380
Surveys of users/customers	4.4	8.3	0.355
Other	5.2	5.6	0.930

3.5 Successful cases feature significantly more products that were tested for their usability

Testing products as part of their development is undoubtedly an established means of recognizing potential design flaws at an early stage. The question remains, whether previous testing has an influence on the success in correcting design flaws of products whose development is finished and which are already in use. In question II.9 participants were asked “How was the PRODUCT tested during its development?” – the diagram in the figure below illustrates the general distribution of answers.

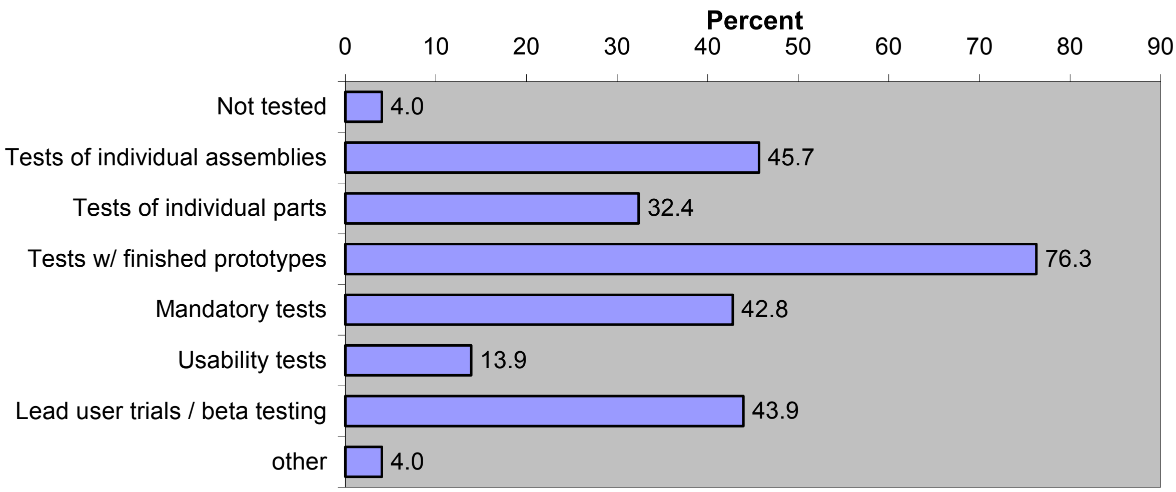

Figure 5. General percentage distribution of answers to question II.9 (n = 173, multiple answers allowed)

The majority – more than three quarters – of participants stated that their PRODUCT underwent some testing at prototype stage. Surprisingly, lead user trials and beta testing were deemed similarly important as tests on individual assemblies and mandatory tests (e.g. for obtaining an operating license for the PRODUCT). However, the least often stated options "No testing" and "Usability tests" [10] are the only variables featuring a significant difference in successful and unsuccessful cases (see Table 6).

Table 6. Percentages of answers in successful and unsuccessful cases (significant differences in boldface)

Variable	**Percentages of answers**		**Significance of difference**
	Successful cases	Unsuccessful cases	
No testing	**2.2**	**11.1**	**0.017**
Tests of individual assemblies	48.9	33.3	0.082
Tests of individual parts	35.0	22.2	0.131
Tests with finished prototypes	78.8	66.7	0.091
Mandatory tests	44.5	36.1	0.330
Usability tests	**16.8**	**2.8**	**0.029**
Lead user trials / beta testing	45.3	38.9	0.452
Other	5.1	0.0	0.164

3.6 Seriousness and likelihood of design flaws do not correlate

In question III.4, the participants were asked to estimate the *seriousness* of the DESIGN FLAW at the time of its occurrence on the scale in Table 7.

The *likelihood* of the DESIGN FLAW was assessed in the subsequent question III.5, again on a scale from "1" (occurrence practically impossible) to "10" (occurrence certain). In Figure 6 the distribution of frequencies for each scale level of likelihood and seriousness is shown.

Table 7. Answer options for question III.4 and corresponding scale level. The term "system" refers to the overall technical object in which the PRODUCT might be embedded.

Seriousness of DESIGN FLAW	**Scale level**
User does not take notice	1
Negligible reduction of usability	2
Noticeable reduction of usability	3
Considerable reduction of usability	4
Failure of the PRODUCT or system with no or insubstantial damage	5
Failure of the PRODUCT or system with damage that requires repair	6
Failure of the PRODUCT or system with substantial but repairable damage	7
Destruction of the PRODUCT or system	8
Direct danger to the user under certain (operating) conditions	9
Direct danger to the user in unforeseeable conditions	10

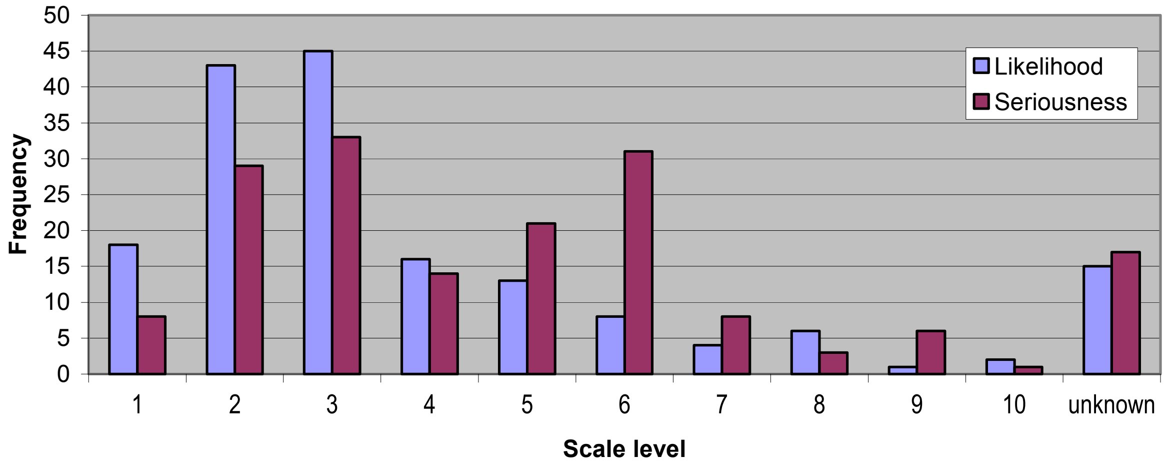

Figure 6. Frequency distribution of answers to questions III.4 and III.5 ($n_{III.4} = 171$, $n_{III.5} = 156$; only one answer allowed)

The possible correlation between the variables seriousness and likelihood was tested using Kendall's τ method. At a confidence level of $\alpha = 0.05$ and for $n = 149$, however, the resulting correlation coefficient of $\tau = 0.037$ is insignificant. It is therefore acceptable to define the overall *severity* of a DESIGN FLAW (which plays a role in the next result) as the product of *seriousness* and *likelihood*.

3.7 Technical characteristics of a design flaw do not determine its severity but "mechatronic" flaws are significantly more often successfully corrected

Question III.2 was put "What was the technological cause of the DESIGN FLAW?". Participants should select all answer options in that applied. Using a hierarchical cluster analysis [13] of all 173 cases, three major answer patterns could be identified.

Figure 7 shows the percentage distribution of causes of the DESIGN FLAW for the identified clusters – which can be characterized as follows:

- **Cluster A:** Cases in which the DESIGN FLAW is predominantly caused by mechanical issues
- **Cluster B:** Cases in which the DESIGN FLAW is mainly caused by software and electronics but also by mechanics ("mechatronic" design faults).
- **Cluster C:** Cases neither belonging to cluster A nor B or where the cause of the DESIGN FLAW is unknown.

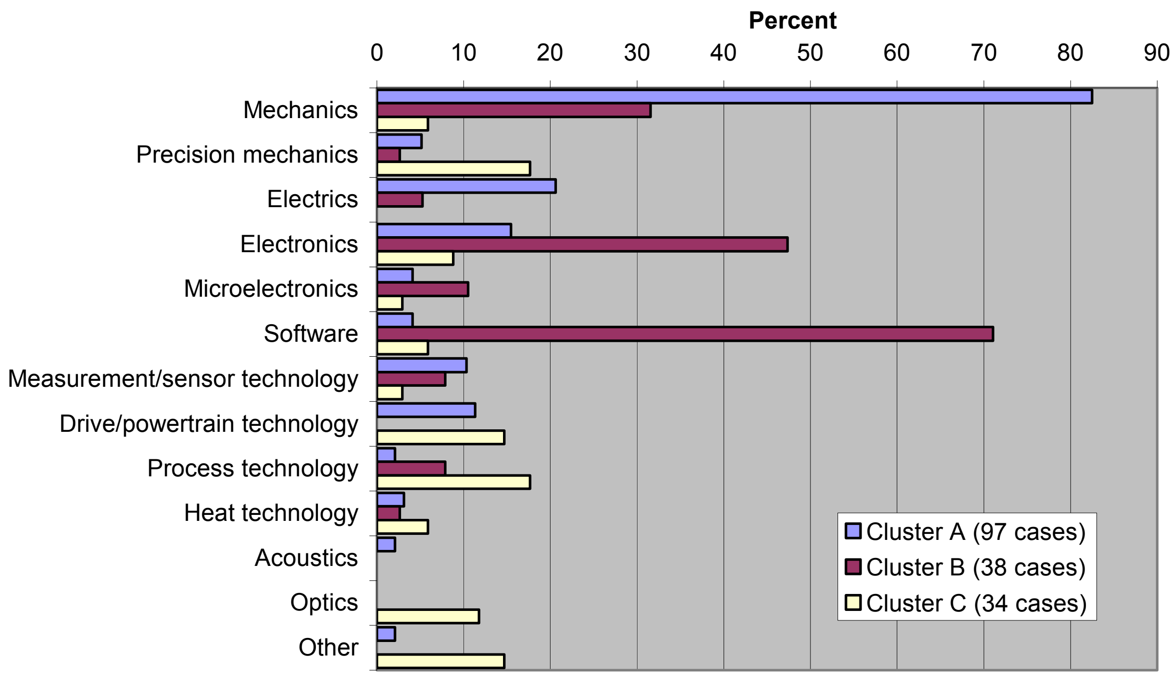

Figure 7. Percentage distribution of answer options to question III.2 for each cluster (multiple answers allowed)

Table 8 shows for each cluster the mean severity of the DESIGN FLAW (see 3.6) and the percentage of cases belonging to the successful group. A p-value of 0.679 confirms that the obviously minor individual differences in mean severity are indeed insignificant. However, the finding that cluster B features the highest percentage of cases belonging to the group where the DESIGN FAULT has been successfully corrected meets the significance criterion of $p < \alpha$, $\alpha = 0.05$.

Table 8. Mean severity for each cluster of answer options.

Variable	**Value**				**p-value (Kruskal-Wallis-H test)**
	Cluster A	Cluster B	Cluster C	Sample	
Mean severity	14.7	14.3	15.6	14.9	0.679
Percentage of successful cases	**76.3**	**92.1**	**68.8**	**79.0**	**0.045**

3.8 Products in successful cases are significantly more often original designs

In question II.8 participants were asked to chose from the following statements the one that best describes the development process of the PRODUCT:

A. Development of new solution principles for a substantially new problem or task and/or entering a new technological territory

B. Adaptation of an existing design to new boundary conditions using well-tried solution principles; task not basically new

C. Adaptation to new boundary conditions by variation of existing parts/assemblies as part of processing an order

These options directly correspond to original design (A), adaptive design (B) and variant design (C) as defined by Pahl and Beitz [1].

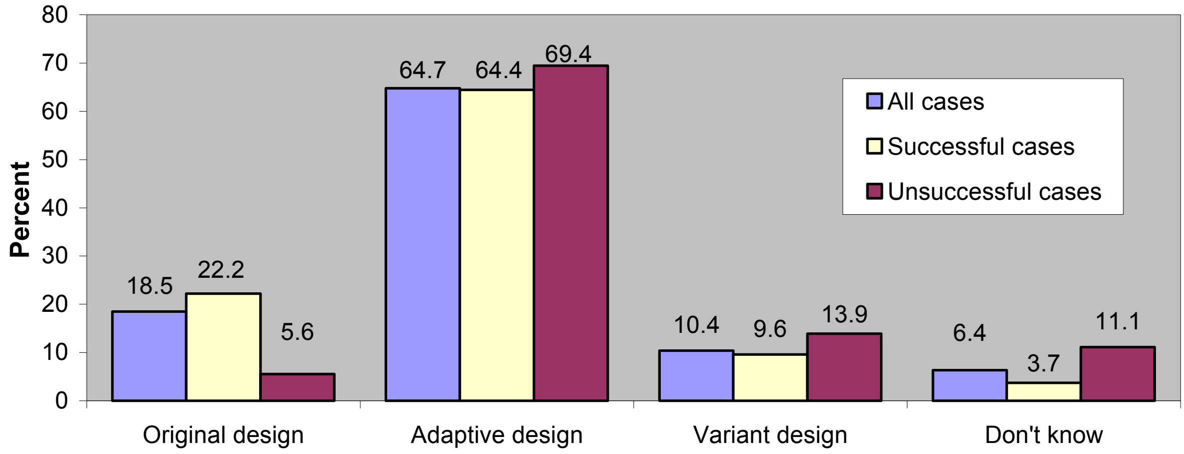

Figure 8. Percentage distribution of answers to question II.8 (n = 173, only one answer allowed)

For all cases, the percentage distribution of design types shown in Figure 8 is according to expectations with a majority of adaptive designs, followed by original and variant designs [1], [2]. Noticeably, however, only 5.6% of products in unsuccessful cases meet the definition of an original design compared to 22.2% of products whose design flaws were successfully corrected. The significance of the overall discrepancy was examined in a similar fashion as in 3.3 by carrying out a Mann-Whitney-U test which revealed a p-value of 0.036.

4 Conclusion

In this study, some company-, process and product-related factors of an everyday design phenomenon have been investigated. The chosen approach allowed for a differentiated view on the results by not only revealing what perceptions exist in general, but also how they differ in successful and unsuccessful cases and where these differences are significant.

As far as company-related factors are concerned, result 3.1 shows that a company's economic figures (i.e. employees and annual sales) are indistinctive for its success in dealing with design flaws. It is, however, incorrect to conclude that smaller companies are equally successful as larger ones.

Result 3.4 shows that companies most often learn from users or customers that their products are flawed. It confirms the assumption that the feedback from this group of individuals plays an important role in discovering design flaws – being more important for successful companies than for unsuccessful ones (however insignificantly). A feedback-related determinant that has been identified as significant for success is the processing of warranty claims which, additionally to providing a channel for some kind of user feedback, also makes the flawed products accessible to its designers.

As various results imply, feedback seems to be a recurring aspect of many significant process-related factors. Result 3.3 reveals that successful companies are characterized by their designers being more committed to activities such as product management and, most significantly, customer/user surveys. Yet, a more active involvement in the take-back of used products – that could give designers the opportunity to review their own design measures based on the condition of products that have reached the end of their service life (as discussed in [4]) – slightly fails the test of being a significant characteristic of successful companies. In result 3.5, the importance of feedback is reflected by the relatively high percentage of products that underwent lead user trials or beta testing, awarding these methods similar weight as mandatory tests and tests of individual assemblies.

Result 3.7 delivers a reasonably strong product-related factor by proving that different archetypical technical characteristics of a design flaw, which inevitably correspond to the technical characteristics of the product, are distinctive for success. It turns out that more "highly-developed" products with design flaws mainly caused by software, electronics and mechanics are handled significantly more successfully than products whose design flaws are of predominantly mechanical nature.

Repercussions of product-related factors may be evident in several results, most obviously in 3.5 where it has been found that successful cases feature significantly more products that were tested for usability. In general, there seems to be a slight but noticeable tendency that indicators for a more "sophisticated" development processes (e.g. PDM and computational engineering in result 3.3) are distinctive for success. However, the question whether this observation can be explained by the finding that more "highly-developed" products are more successful – implying that more "highly-developed" products require a more "sophisticated" development process – cannot be answered in this study.

Result 3.8 is interesting inasmuch as being unsupportive of the idea that adaptive designs benefit from a possibly greater amount of knowledge (which could include knowledge about design flaws). Following the above reasoning, a possible explanation would be that less "highly-developed" (and therefore less successful) products are per se less original in design.

The high percentage of cases in which "The DESIGN FLAW was corrected and the solution has been or will be implemented in the PRODUCT" (almost 80%) is a strong indicator that the survey suffered – not quite unexpectedly – from considerable nonresponse error: designers who felt they deal successfully with the issue probably had a higher motivation for participating in the study. As a result of the imbalance of successful and unsuccessful cases (135 vs. 36), significant differences were more difficult to identify, but are, however, not less valid.

Further work involves the evaluation of the remaining data, focusing on the relationship between product- and process-related factors and investigating the aspect of "learning" from design flaws in the sense of avoiding their repetition. In order to obtain a more in-depth look into these issues, complementary methods, e.g. case-study interviews, might be considered.

References

[1] Pahl G. and Beitz W., "Engineering Design – A systematic approach", 4th ed., Springer, Berlin, 1996

[2] Hundal, M., "Systematic mechanical designing: A cost and management perspective", ASME Press, New York, 1997

[3] Petroski, H., "The Evolution Of Useful Things", Vintage Books, New York, 1992

[4] Gries, B., Blessing, L., "Towards a Disassembly Process-Oriented Design of Sustainable Products", Proceedings of the 14th International Conference on Engineering Design, Folkeson A. et. al (editors), The Design Society, Stockholm, 2003

[5] Blessing, L., "What is this thing called Design research?", Annals of the 2002 CIRP Design Seminar, Hong Kong, 2002.

[6] Eurostat, "NACE Rev. 1 - Statistical Classification of Economic Activities in the European Community – Introduction", Office for Official Publications of the European Communities, Luxembourg, 1996

[7] CD-ROM database, "Hoppenstedt Mittelständische Unternehmen", Bonnier Deutschland GmbH & Co. KG, Darmstadt, 2003

[8] CD-ROM database, "Hoppenstedt Großunternehmen", Bonnier Deutschland GmbH & Co. KG, Darmstadt, 2003

[9] Dillman, D., "Mail and Internet Surveys – The Tailored Design Method", 2nd ed., John Wiley & Sons, New York, 2000

[10] Rubin, J., "Handbook of Usability Testing", John Wiley & Sons, New York, 1994

[11] Bortz, J. and Lienert, G., "Kurzgefasste Statistik für die klinische Forschung", 2nd ed., Springer, Berlin, 2003

[12] Bortz, J., "Statistik für Sozialwissenschaftler", 5th ed., Springer, Berlin, 1999

[13] SPSS Inc., "SPSS® Base 12.0 User's Guide", SPSS Inc., Chicago, 2003

Corresponding author:
Bruno Gries
Engineering Design and Methodology Group
Technical University Berlin
Strasse des 17. Juni 135
D-10623 Berlin, Germany
Tel: +49-(0)30-314 23161
Fax: +49-(0)30-314 26841
E-mail: gries@ktem.tu-berlin.de
URL: http://www.ktem.tu-berlin.de/