

Extended Web Assessment Method (EWAM) - Evaluation Of E-Commerce Applications From The Customer's Viewpoint

(Copyright 2002 IEEE. Published in the Proceedings of the Hawai'i International Conference on System Sciences, January 7-10, 2002, Big Island, Hawaii.)

Petra Schubert

*IAB – Institute for Business Economics
FHBB – University of Applied Sciences, Basel
Peter Merian-Str. 86, 4002 Basel, Switzerland
tel +41 61 279 1774
p.schubert@fhbb.ch*

Walter Dettling

*IAB – Institute for Business Economics
FHBB – University of Applied Sciences, Basel
Peter Merian-Str. 86, 4002 Basel, Switzerland
tel +41 61 279 1790
w.dettling@fhbb.ch*

Abstract

The paper presents and applies the Extended Web Assessment Method (EWAM) an evaluation tool which has been specifically created for the assessment of e-commerce applications. EWAM builds on the Web Assessment Method developed at the University of St. Gallen, Switzerland [12] and integrates findings from the Technology Acceptance Model [1] and several alternative approaches. It defines an evaluation grid including a set of criteria to appraise the quality and success of existing e-commerce applications. The focus lies on consumer perspectives and the specific features of the Internet as medium. Using the EWAM tool, an analysis of eight commercial Web sites in two different business sectors – consumer goods and e-banking – was performed. The findings show that according to the assessors most of the Web sites assessed do still not fully meet user expectations. Consumer goods are a local business. Despite of this possible disadvantage they showed a good performance in satisfying their customers. E-banking applications do not present differentiated features, one can seemingly be replaced by the other with no problem.

1 Introduction

Digital marketing calls for new marketing paradigms and places some of the existing rules in a new light. It is of vital necessity that Internet companies should also take these

paradigms into consideration for their Internet business. In the online medium, where customers are only a mouseclick away from comparable offers, distinguishing characteristics are the only means to bind customers long term to one's own offer. The most important thing here is to concentrate above all on features which are only possible online. One possibility, for example, is the setting up of customer communities [11], which pool their knowledge and their profiles for communal use and render economies of scale effective in a new way.

The paper presents the Extended Web Assessment Method (EWAM), an evaluation tool which has been specifically created for the assessment of e-commerce applications. A major problem posed when searching for possible improvements of e-commerce applications is the question of the criteria used to determine the success of such a system. This is where EWAM comes in, by offering a set of relevant criteria which provide a basis for measuring the success of Web sites of individual Internet companies. The prescribed criteria enable existing Web sites to be evaluated from the customer's point of view. To do this assessors have an online tool at their disposal with whose help data can be collected and evaluated. By defining sector profiles, comparisons of a Web site with the sector average or with the Best Practice Profile can be made. This benchmarking can ultimately be used to suggest steps for improvement.

The paper presents the findings of an empirical study carried out in April 2001. Twenty students from an execu-

tive training class in E-Business completed the web assessment forms, evaluating web sites of two different business sectors: consumer goods and e-banking. Before starting this project they were thoroughly instructed in the use of the tool. Additionally, they had to provide a qualitative analysis of the eight web sites in the form of a written report. The fulfillment of this task was needed for a successful attendance of the class.

Since the study was conducted in Switzerland, the results represent a Swiss perspective. E-business is still a very regional affair, making it hard for consumers to assess Web sites in countries other than their home country. Nevertheless, we included a German and a US company in our sample to make the study more interesting and to add an international perspective in a study about the inherently global Internet commerce environment.

Only Web sites which were intensively used or tested by the assessors were investigated. The assessors being well-instructed Web experts, the sample was intended to represent a small quantity of highly-qualified user opinions. The evaluations resulted into Web assessment reports which compare the individual performance profile of one company with the "best of its class" (best practice profile) and the general sector profile (aggregation of all companies in the same sector).

The paper starts with a description of the methodology of the Web Assessment explaining its origins and the latest adjustment to advances in technology and Web site design. We then describe our empirical study and the way in which we prepared the results. Four companies from both sectors (consumer goods and e-banking) are briefly introduced and the findings for the sector are discussed in detail. The paper concludes with some remarks about the current state-of-the-art in Web sites design and functionality and their suitability for e-commerce transactions.

2 Theoretical background

2.1 Original Web Assessment Method

The Web Assessment Method was developed in 1997 at the Competence Center for Electronic Markets (CCEM) at the University of St. Gallen in conjunction with the company partners [12]. The original impulse arose from the discontent felt by industrial partners of the CCEM about unsatisfactory results of e-commerce applications already carried out.

The method defines an evaluation grid with a set of criteria to appraise the quality and success of existing e-commerce applications. Besides the rigorous focusing on consumer perspectives, the success in implementing the offer of products and services is considered with reference

to the specific features of the electronic medium. Successful Internet business activity requires the following new paradigms to be taken into account:

Electronic markets and transaction phases: The Web Assessment Model examines the three classic transaction phases of electronic markets (information, agreement and settlement). A fourth element, the 'community component', is integrated as a link between the actual purchase transaction and the necessary trust relationship in the virtual realm.

Information technology / Media-inherent Characteristics: Where marketing aspects are concerned, the Web Assessment model focuses on the special features that are inherent in the Internet. The assessment criteria are, besides the transaction phases and community component, derived from the characteristics of the electronic medium: hypermedia presentation, database interface (expert system), 24-hour access, anonymity, ubiquity, configuration possibility of the user surface, integration with the customer and asynchronous communication.

Performance marketing: The underlying idea of performance marketing is that the client should not only be sold the core product, but should additionally be offered a range of complementary products to maximize customer benefit. These additional services customize the product and make it attractive for the client. In the difficult arena of international competition product differentiation is thus made possible.

In an empirical study carried out in 1997/1998 more than 70 questionnaires were collected and evaluated from over 55 different participants (researchers and practitioners) [12]. The result of the study was a comparison of the individual performance profile of two companies (SwissAir and Amazon.com) with comparable offers from competitors in the same sector. At the same time in this initial application of the Web Assessment Tool the special strengths and weaknesses of the companies could be identified.

As a result of technological progress and the accompanying value change among users since the development of this method, it was thought necessary to adapt the questionnaire for the purpose of data collection and identification of success criteria.

In summer 2000 the method was fundamentally revised; besides taking account of new research findings – especially in the Internet marketing field – it also incorporated the Technology Acceptance Model (TAM) [1], established for the acceptance of information systems, which will be briefly described in the following paragraph.

2.2 Technology Acceptance Model (TAM)

With the Technology Acceptance Model Davis describes the effect of system features on the acceptance of users with regard to new computer-based information systems. Davis chose the 'Fishbein Model', a psychological behavior model, as the basis for the development of the TAM. This model, specified by Fishbein in 1967, was extensively analysed by Fishbein & Ajzen [2] and further elaborated into the 'Theory of Reasoned Action'. The TAM follows the causal chain of the TRA: "Attitude" → "Intention" → "Behavior".

The hypothetical relationships defined in the basic model were tested empirically using a questionnaire (with questions on e-mail and XEDIT) followed by regression analysis.

Out of 120 questionnaires distributed to employees of IBM Canada's Toronto Development Laboratory, 112 were returned completed (n = 185, of which 109 regarding e-mail and 76 about XEDIT). The resulting evaluation confirms several hypotheses.

(H1) 'Attitude' has direct influence on 'Actual System Use'.

(H2) 'Usefulness' has significant effect on 'Attitude toward Using'.

(H3) + (H4) 'Ease of Use' has an effect on 'Attitude' and 'Usefulness'.

... and provides further findings which were not expected in this form:

(H5) 'System' has a significant effect on 'Ease of Use' but not on 'Usefulness'.

(H6) 'Usefulness' has a direct effect on 'Actual System Use'.

(H7) 'System' has a direct effect on 'Attitude toward Using'.

In the original form of the TAM the "Subjective Norm" of the Fishbein Model is not integrated, as Davis did not consider it to be relevant. Malhotra/Galletta [9] conclude in 'Extending the Technology Acceptance Model to Account for Social Influence, Theoretical Bases and Empirical Validation' that Davis had underestimated the importance of the 'Subjective Norm' in the TAM. Their empirically tested study points to the fact that for the user social influences also play an important role in the acceptance and use of new information technologies.

The TAM is an important contribution to understanding the use, behavior and acceptance of new information systems by the user. Thus it also appears suitable for further application as a theoretical base in the Extended Web Assessment Method.

The motivational variables "Attitude toward using", "Perceived Usefulness" and "Perceived Ease of Use" serve

as a link between system features and the behavior of an individual in relation to the use of new information systems (Actual System Use). From this point of view the model is successful [1].

The fact that social components (Subjective Norm) were not taken into account in TAM was criticized in ongoing research [9], which underlines the importance of these components. Hence the social influences were also incorporated into EWAM in the form of criteria regarding "Trust".

2.3 Alternative Approaches

The following alternative approaches to Web evaluation listed below were also taken into account during the revision of the Web Assessment Method:

- Design Quality of Web sites for Electronic Commerce: Fortune 1000 Webmaster's Evaluations [8]
- Perfekte Webseiten – wie sieht die Realität aus? [7]
- GomezPro.com [4]
- JurisNET [6]

3 Extended Web Assessment Methode (EWAM)

3.1 Description of the method

The Extended Web Assessment Method (EWAM) builds on the original Web Assessment Method and integrates findings from the Technology Acceptance Model and several alternative approaches [3]. It defines an evaluation grid including a set of criteria to appraise the quality and success of existing e-commerce applications, which will be presented below. The focus lies on consumer perspectives and the specific features of the Internet as medium.

A successful e-commerce application must meet the needs of the user according to "Perceived Usefulness" (Criteria USEF1-USEF15) and "Ease of Use" (Criteria EOU1-EOU8). Under the headword "Trust" (Criteria TRUST1 - TRUST2) questions about the "Subjective Norm" are additionally taken into account. A success or quality feature must be assigned to one of these categories. An excerpt from the lists is depicted in figure 1.

When evaluating an e-commerce application according to the Extended Web Assessment Method a Web site is first of all allocated to a sector. This serves later during evaluation to identify the reference sector for benchmarking. The success and quality criteria are formulated in general terms and are valid in every sector, but are differentiated by their importance rating. In order to take due ac-

count of the differences in the individual sectors, criteria are weighted corresponding to different sector profiles and their relevance in the sector. Hence being up to date with information is of greater importance for the supplier of financial information (e.g. Stock Brokerage, real-time share prices) than for a supplier of consumer goods. With the distribution of digital goods (e.g. software) the choice of generic services (EOU5), for example, tracing and tracking of a parcel, is of lesser importance than with the delivery of a book. In order to be able to undertake specific and high-quality analyses, it is essential to record the level of importance per criterion and per sector exactly. The importance per criterion is recorded on a scale from “unimportant” (-2) through “less important” (-1), “important” (+1) to “very important” (+2).

Trust is the *sine qua non* of e-commerce; without trust no business is done. The creation of a trustworthy environment on the Internet is a major challenge for the success of e-commerce. According to the “Theory of Reasoned Action” [2] “Subjective Norm” refers to how an individual is influenced by a person particularly close to them, and by this person's opinion of a certain way of behaving.

3.2 Summary of the Methodology

An EWAM criterion is first of all assigned to a criteria category (“Ease of Use”, “Usefulness” or “Trust”) and within these three categories allotted to one of the four transaction phases of electronic markets (the information, agreement, settlement and after-sales phase), to the community component or to the category “criteria which concern all phases”. Compared with the original Web Assessment Method, the EWAM was extended to include the after-sales phase and a category group “criteria which concern all phases”.

3.3 Data collection: the EWAM tool

Data is collected over the Internet with an online questionnaire (“EWAM tool”) in which the usual structure of the original tool [13] was almost retained. The individual criteria were assigned to a transaction phase or a community component, and two new sections were added (“After-sales phase” and “criteria which concern all phases”).

When the assessor starts the evaluation with the EWAM tool, as a first step he must record the URL of the Web site under examination and assign it to a sector. The scale of the possible choices is so arranged that the assessor must decide on a positive or negative statement with each value. The scale has four values (+,+,+,-,-). The alternative value “n. a.” (not applicable) can be used if a criterion is not relevant or not available in a particular context. The criteria are so formulated that a positive (negative) evaluation will also lead to a positive (negative) result. “I strongly agree” always scores (+2), “I slightly agree” (+1), “I slightly disagree” (-1) and “I strongly disagree” (-2). “N. a.” scores nil, which is disregarded with further calculations (e.g., of averages).

3.4 Data preparation and analysis

For the drawing up of meaningful evaluations of any Web site under examination (the Web site compared to the sector average, to best practice profile or to one of its competitors) three profiles were defined in the EWAM.

- *Sector Profile*: the profile of the relevant sector
- *Company Profile*: the profile of the Web site
- *Best Practice Profile*: the profile of the best-of-breed in the relevant sector

The EWAM judges a Web site purely from the customer's point of view. The best EWAM result does not necessarily mean that this Web site is the most successful in reality, since success is influenced by further factors

1. Information Phase							
		ID	‡	+	,	i	n.a.
Web Page and Specific Offer are Easily Found	EQU1		<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Good Structure of Content	EQU2		<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reasonable Information Quantity	EQU3		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Quality of Content Meets User Expectations	USEF1		<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cost Benefits Passed on to the Client	USEF2		<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bundling: Good Combination Possibilities for Products/Services	USEF3		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Good Recommendation Systems	USEF4		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
Adequate Application of Hypermedia	USEF5		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 1: EWAM tool: List of criterias

(e.g. e-business relevance of the offer, profitability, back-end integration, financing aspects, etc). A Best Practice Profile can only be established when (a) a sufficient number of different Internet businesses per sector have been evaluated and (b) these have been compared with their performance in the real world, in which case a best practice must also be effectively successful there.

Accordingly, an adequate best practice data reference base can only ensue from the combination of points (a) and (b).

3.5 Personal Web Assessment Report

One of the results of a Web site assessment is a Personal Web Assessment Report containing the following analyses and graphical representations:

- a) Summary of the individual criteria and results in the categories “1. Information Phase”, “2 Agreement Phase”, “3 Settlement Phase”, “4 After-sales Phase”, “5 Community Components”, “6 Final Section” and a calculation of the total score.
- b) Comparison of the Web sites examined with the sector average and with the sector Best Practice, in the form of a quantitative and graphical analysis, taking no account of the importance rating of the criteria.
- c) Graphic comparison of the results of the first six categories (a, b) with their importance rating for company and sector profile.
- d) Comparison as b) above, but taking full account of the importance rating of criteria.

We identified three different target groups which gain the following benefit from the personal Web assessment report:

- *Internet merchant*: Comparison of the quality of the customer orientation of his Web site with the sector profile or with a direct competitor. Possible suggestions for improvement can be drawn from the result.
- *Potential Internet merchant*: Awareness of the success criteria of commercial Web sites in the target sector.
- *Internet buyers*: Examination of the quality of the customer orientation of the assessed Web site.

4 EWAM findings: Consumer Goods and E-Banking

The following paragraphs present the findings of our study for two different sectors: consumer goods and e-banking. The Web sites chosen stem from Switzerland, Germany, and the USA. Our aim was to also study differences between the countries.

4.1 Consumer Goods

Web sites assessed:

- <http://www.webvan.com> (***)
- <http://www.le-shop.ch>
- <http://www.migros.ch>
- <http://www.mcfood.ch>

The companies which were chosen for the consumer goods sample are from two different countries: Switzerland and USA. **WebVan** was an ambitious American online venture with a vision “of 26 massive automated warehouses and a national fleet of vans that would deliver groceries – and just about anything else – to everyone in America” [5]. Based on big money investments it is not surprising that WebVan came out as “Best of Class” although the company had been experiencing some problems [10] and had to shut down its operations by mid 2001. **Le-Shop** is a Swiss independent online business which acts as a reseller for food and drugs. Acting solely on the Internet it is not surprising that it was voted “best in Switzerland”. Migros is a Swiss store chain selling all kinds of consumer goods with a focus on the food sector. McFood, finally, was a small Swiss “newcomer” which had never been able to setup its logistics and was picked as the “bad example” by the course instructor.

4.2 Interpretation of the assessment forms

In order to be able to assess the Web sites we had to interview the assessors about their perceived importance ratings for the consumer goods sector. As mentioned before, the “raw data” for a data set does not allow interpretation of the performance of a Web site. We distributed two general Web assessment forms to a selected set of people who are familiar with using the Web as a source of information and shopping. These people were asked to state their expectations for Web sites in the two different sectors. Their forms were collected and the data was entered into the tool as the so-called “importance rating”. The importance rating shows the expectation of a user for a certain criterion. If expectations are high, meaning that a certain feature is important for a Web site from the user perspective, and cannot be found during the assessment the result will be a bad evaluation for the respective criterion. For the evaluation, the criteria are then summed up and aggregated to show only one value for a certain phase.

For an explanatory presentation of our results we picked www.webvan.com (the best practice profile). Figure 2 shows the aggregated values for the consumer goods sector. The column labeled “importance” shows the

Phase	Results (Range -2/+2)	Importance	Company Profile		Best Practice Profile (BPP)	Sector Profile (SCP)
			Difference to			
			BPP	SCP		
1. Information Phase	3. 0.71	0.67	0.00	0.30	0.67	0.38
2. Agreement Phase	0.50	0.61	0.00	0.14	0.61	0.47
3. Settlement Phase	0.56	0.45	0.00	0.41	0.45	0.05
4. After-Sales Phase	2. 0.83	0.83	0.00	0.55	0.83	0.27
5. Community Component	0.17	-0.68	0.00	0.11	-0.68	-0.80
6. Final Section	1. 1.43	1.04	0.00	0.46	1.04	0.58
7 Overall Score	n.a.	0.49	0.00	0.33	0.49	0.16

Figure 2: Aggregated data set for the consumer goods sector, company profile: www.webvan.com

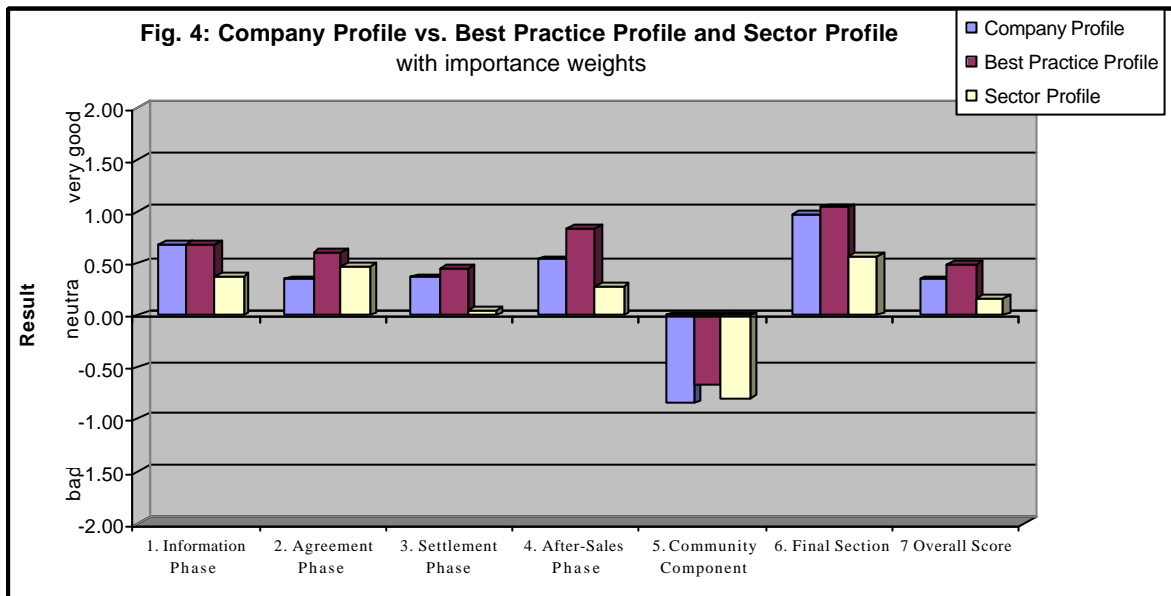


Figure 3: Company Profile of www.le-shop.ch

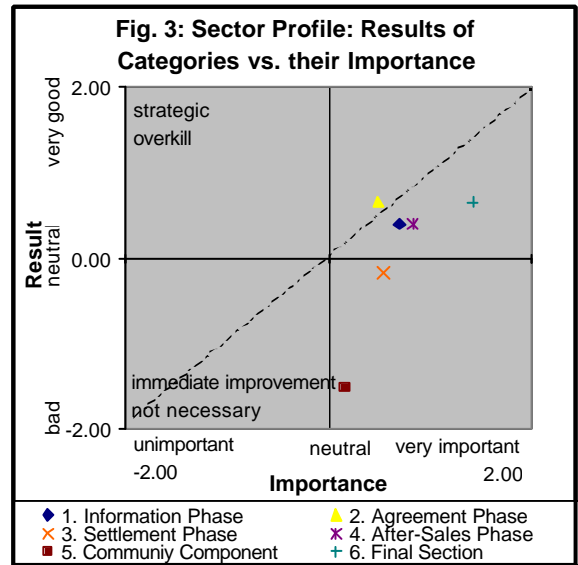
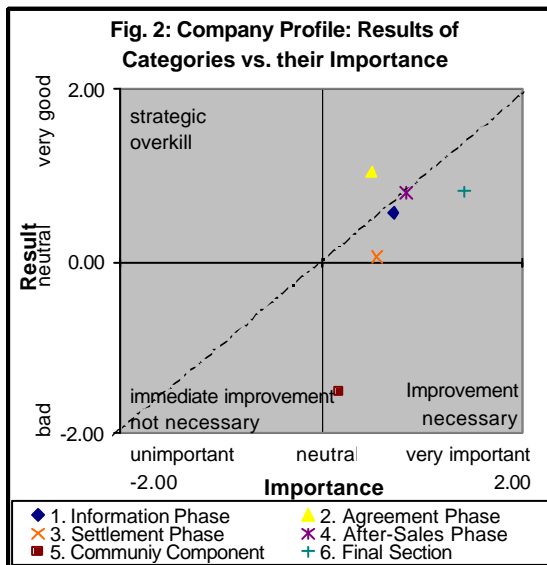


Figure 4: Company Profile of www.migros.ch

perceived user expectations. The possible range of values is +2 to -2. For the calculation, the importance values are later transferred to a scale of 0 to 1. The final section is the most important followed by the after-sales and the information phase. The final section contains general criteria which apply to all phases (availability, user interface, interactivity, trust) and seems to be very important for consumer goods.

The community component received a very low importance rating. This could be due to the fact that the experts which we polled for the importance rating did not think about the power of community building on the Internet as e.g. pointed out by Schubert [11]. On the other hand, purchasing food and drugs might indeed not be an activity during which community-support is appreciated although examples could be recommendations, product ratings, aggregation of demand to lower prices, etc.

Figure 3 shows the company profile for www.le-shop.ch, the best Swiss Web site for consumer goods (second after the American company WebVan). As we can see, Le-Shop scored almost as high as WebVan in two of the three important sections, the final section and the information phase. There seems to be room for improvement in the after-sales phase and especially in the agreement phase which received a lower value than the sector average.

Figure 4 shows a comparison of perceived user expectations and assessment values. It contains recommendations for generic strategies which should be applied dependent on the results in the various sections.

- *Strategic Overkill*: Entries in the upper left field indicate (very) good results in a (rather) unimportant category. Available resources are possibly not being applied effectively.
- *Maintain Strategy*: Entries in the upper right field indicate (very) good results in (very) important categories.
- *No immediate improvement necessary*: Entries in the lower left field indicate (very) poor results in (rather) unimportant categories.
- *Improvement necessary*: Entries in the lower right field indicate (very) poor results in (very) important categories.

As we can see from figure 4 www.migros.ch scored higher than the expectations in the category Agreement Phase. Interestingly, this is a category which were deemed unimportant for the consumer goods sector. The diagonal shows the points where expectations meet actual assessments. Most assessments for www.migros.ch are grouped around the diagonal. The general advice for www.migros.ch is to focus its future improvements on the settlement phase and the community component.

The aggregated results for all four companies (sector profile) show values which are lower than the individual ratings achieved by www.migros.ch. This is due to the fact that one of the companies, McFood, scored unusually low in all areas because it has not really started operations yet (without telling the customer on the Web site).

4.3 Electronic Banking

Web sites assessed:

- <http://www.credit-suisse.ch> (***)
- <http://www.deutsche-bank-24.de/>
- <http://www.ubs.ch>
- <http://www.bekb.ch>

Credit Suisse and UBS are the two biggest banks in Switzerland. The Berner Kantonalbank (www.bekb.ch) is a small Swiss regional savings bank which, nevertheless, offers a comprehensive e-banking application. Bank-24 is the online venture of Deutsche Bank, the biggest bank in "Euroland". It was the first purely Internet-based bank in Germany. Deutsche Bank has a proactive Internet strategy – they were the only bank to launch a project with Ecash which so far has not really taken off (and probably never will since new and better payment systems are already on the market).

The e-banking sector resulted to be difficult to assess with the EWAM. Since most of our assessors did at most already use one of the selected e-banking services (who has more than one bank account?) they had to rely on the demo versions which are offered on all four Web sites. Credit Suisse was chosen to be best of its class in this sample. The results clearly state that the four Web applications are very similar to each other in terms of functionality, look and feel, contact possibilities, and so on.

The results suggest that differences are almost negligible. What catches the eye are the almost zero values for the community category. Neither did the assessors attribute great importance to community (-1.75) nor did they report any kind of community-supporting components on the four Web sites. Interestingly, the after-sales phase seems to be of high interest. This is where the interaction between company and customer often relates back to the "real world", namely to the telephone. Most after-sales services, such as the handling of complaints, are not processed via the Internet.

5 Mathematical Derivation

The following paragraphs describe the calculations for the comparison of any Web site examined with the average and with the best practice of this sector, including the importance rating of criteria, for the "Sector Profile". The cal-

calculations for the “Company Profile” and “Best Practice Profile” are similar.

Criteria, Importance and Evaluation				
Criteria: X_i	Importance: W_i (Range -2 to +2)	Importance: Wg_i (Range 0-1)	Evaluation: \bar{r}_i (Range -2 to +2)	Weighted Value: $R_i = Wg_i * \bar{r}_i$
X_1	W_1	Wg_1	\bar{r}_1	R_1
X_2	W_2	Wg_2	\bar{r}_2	R_2
...
X_{26}	W_{26}	Wg_{26}	\bar{r}_{26}	R_{26}

Table 2: Criteria X_i , Importance W_i & Wg_i , Evaluation r_i , Weighted Value R_i

The criteria ($X_{i; i=1..26}$) are grouped in six categories ($K_{k=1..6}$) according to the three transaction phases (information, agreement, settlement), the “After-Sales Phase”, the “Community Component”, and the “Final Section”

Category K_k	Description
K_1	1. Information Phase
K_2	2. Agreement Phase
K_3	3. Settlement Phase
K_4	4. After-Sales Phase
K_5	5. Community component
K_6	6. Final Section

Table 3: Categories K_1 - K_6

Transformation of the importance of a criterion:

The evaluation of the individual criteria and their importance uses a scale from (-2) to (+2). To avoid the problem of the multiplication of two negative values, the importance of each individual criterion (W_i) is transformed into a range from (0) to (1) (Wg_i).

$$Wg_i = \frac{1}{4}(W_i + 2)$$

Average evaluation of each criterion

$$\bar{r}_i = \frac{1}{m} \sum_{j=1}^m r_{ij} \quad \text{where:}$$

m = number of assessors evaluating a criterion X_i
 r_{ij} = individual result of a criterion in range (-2/+2)

Multiplication of average result with importance rating:

The definitive result for every criterion is arrived at by multiplying the average result of any criterion (\bar{r}_i) with its importance rating (Wg_i). Thus a criterion only reaches the top score when the importance rating ($W_i = 1$) is top. With lower importance ratings ($1 = W_i = 0$) the final result of a criterion decreases, but does not fall below zero.

$$R_i = \bar{r}_i * Wg_i$$

Addition of the individual evaluations per category:

The result for the category ($K_{k=1..6}$) is arrived at by adding the R_i per category. The example of the Information Phase (K_1) and the Agreement Phase (K_2) is given below.

$$K_1 = \sum_{i=1}^8 R_i \quad K_2 = \sum_{i=9}^{10} R_i$$

Calculation of the percentage attainment of top score for K_k :

a) Calculation of the minimum (R_{MINi}) and maximum (R_{MAXi}) evaluation for a criterion with given importance in range (0-1).

$$R_{MINi} = Wg_i * -2 \quad -2 = R_{MINi} = 0$$

$$R_{MAXi} = Wg_i * 2 \quad 0 = R_{MAXi} = 2$$

b) Calculation of the minimum (K_{MINi}) and maximum (K_{MAXi}) evaluation for a category. The example given is for the Information Phase ($K_{k=1}$).

$$K_{MIN1} = \sum_{i=1}^8 R_{MINi} \quad K_{MAX1} = \sum_{i=1}^8 R_{MAXi}$$

c) Calculation of the percentage attainment of top score

$$K\%_k = \frac{(K_k + K_{MAXi})}{K_{MAXi} + (K_{MINi} * -1)} * 100\% \quad \text{simplified:}$$

$$K\%_k = 0.5 \left(\frac{K_k}{K_{MAXi}} + 1 \right) * 100\%$$

Transformation of $K\%_k$ in a range (-2/+2):

In further analogies the result of a category (K_k) should be compared with its importance rating. Additionally the result $K\%_k$ is transformed in a range of (-2/+2) as the definitive result of this category (KR_k).

$$KR_k = \left(\frac{K\%_k}{100} * 4 \right) - 2$$

A value (e.g. 0.96, Information Phase- KR_1) in a range from (-2/+2) indicates that in the sector average the Information Phase has been evaluated by users as relatively good (evaluation scale: $KR_k = 1.0$ corresponds to ‘good’). If the

Web site examined shows a lower value, then it comes off comparatively worse in the Information Phase.

Comparison of the results of the individual categories (KR) with their importance ratings:

Calculation of the average value of the importance ratings of individual criteria (W_i) for a particular category (K_k). The Information Phase is given as an example:

$$KW_k = \frac{1}{8} \sum_{i=1}^8 W_i$$

Overall Score:

The Overall Score is the final result of a profile (PR). It is calculated from the sum of the six categories (KS) in relationship to the theoretical maximum result of the respective profile.

a) Calculation of the sums of all categories (KS)

$$KS_o = \sum_{k=1}^6 K_k \quad \text{where } o \text{ (} o= 1..3 \text{) indexes the three profiles. (Sector Profile: } o=1 \text{)}$$

b) Calculation of the percentage attainment of the top score for KS

Calculation similar to 5c), but for all categories ($k = 1..6$). The percentage attainment of the top score of $KS\%_o$ is calculated with the following formula:

$$KS\%_o = \frac{(KS_o + RS_{MAXo})}{RS_{MAXo} + RS_{MINo} - 1} * 100\% \quad \text{simplified:}$$

$$KS\%_o = 0.5 \left(\frac{KS_o}{RS_{MAXo}} + 1 \right) * 100\%$$

where

RS_{MAXo} = The theoretical top score for all criteria in profile o

RS_{MINo} = The theoretical minimal score for all criteria in profile o

c) Transformation of $KS\%_o$ in a range (-2/+2)

$$PR_o = \left(\frac{KS\%_o}{100} * 4 \right) - 2$$

For the sector profile this results in a value of 1.02, which means “good” on the scale (-2/+2). The relevant interpretation is: “In this sector Web appearances are considered good within the sector average”.

6 Discussion of findings

Web assessment is a very ambitious and labor-intensive work. The assessors have to meet certain criteria:

- They need to understand the criteria of the Web assessment form very well, hence they must undergo a thorough instruction
- They must be experienced Web users

- They must take the time to go through all four transaction phases for each Web sites assessed (including delivery and payment!)

The two sectors which we put to the test were not equally well-suited to be assessed with the tool. The Web assessment method is based on the four transaction phases of a typical purchase. Thus, e-banking is difficult to test with the existing categories. Nevertheless, we still think that it is possible to evaluate almost any kind of Web site since most of the criteria refer to Web site functionality and not purely to the process.

The e-banking sector shows homogeneous Web applications which do not differentiate themselves from the competition. This is not surprising since banking applications are quite generic and involve little personal involvement on the part of the user.

The consumer goods business is a local business – even on the Internet. This is quite evident because of the nature of the products (have to be physically delivered, are prone to spoil or break). It is even more surprising, that the Web sites selling consumer goods seemed to be very mature. The services were found to be quite reliable, the problem of cooperation with logistics partners had been solved satisfactorily. Some complaints were the limited range of products (e.g. frozen food is not available) or the missing link to the inventory (availability check). Some orders arrived with missing items because they had been out of stock at the time when the package was prepared.

Generally, it seems that few Internet merchants have ever tried to assess their Web sites from a consumer perspective. Most are still driven by technical possibilities. It is noticeable that there is a missing link to “traditional” information systems, e.g. an existing ERP, which could help to perform an availability check which is crucial to an on-line transaction. Most systems seem to be stand-alone Web applications. This can be either due to the consideration of security (a bad reason from the customer perspective) or due to the high costs of an integration of the different systems.

7 Possible restrictions to interpretation

The authors are aware that the empirical study with a sample set of 20 expert opinions can only reflect a limited and somehow biased picture of current practice in the two sectors. The bias exists because the students share similar opinions of e-commerce and are a homogeneous group (because they attended the same E-Business class, their opinions tend to be less universal than if they were hand picked at random). For e-commerce as a whole, 20 people from Switzerland are not representative of the thousands of Web users in Switzerland’s B2C field. Nevertheless, since

the EWAM is a very knowledge-requiring process we cannot ask a random sample of people to do the assessments.

The Web sites chosen for evaluation were not very diverse. Only four companies per sector were chosen. In the physical world these three sectors have hundreds of companies that offer services – on the Web, nevertheless, it took us a week to find four Web companies for each sector that were suited for testing.

8 Conclusions

EWAM lays down a conceptual framework for the evaluation of commercial Web sites whose basic form – the Web Assessment Method – has already proved itself in operation for several years. With the EWAM this work is placed in the context of current developments in e-commerce and united with an established scientific model (the TAM). Web sites can thereby be thoroughly appraised and implicitly the degree of customer orientation can be judged. An important improvement was considering the importance of a criterion for a specific business sector which led to a more differentiated assessment of Web sites in different industries.

However, further research and development work is needed for the successful launch of EWAM. In a first phase the important thing is to gather for all sectors a quantitatively and qualitatively sufficient database, whereby the importance of each individual criterion per sector must be recorded and existing Web sites must be evaluated. Together with a sufficient database and knowledge about commercial Web sites which have proved successful in practice, the best practice profiles for each sector must be identified.

With EWAM, the focus is fixed exclusively on the customer's perspective in the B2C field. Further success factors such as the integration of the supply chain, the linkage of in-house information systems or the consideration of financing and yield aspects are excluded from the present version.

9 Acknowledgements

The authors are indebted to a student class who – within the scope of their executive training on E-Business at the University of Applied Sciences Basel – contributed the data to this research paper.

10 References

- [1] Davis, Fred D. Jr. (1985): A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results. Doctoral Thesis, Sloan School of Management, Massachusetts Institute of Technology, 1985.
- [2] Fishbein, M.; Ajzen, I (1975): *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*, Reading, MA: Addison-Wesley, 1975.
- [3] Giger, Adrian (2000): *Erweiterte Web Assessment Methode – Beurteilung von E-Commerce Applikationen aus Kundensicht*, Master Thesis, University of Applied Sciences (FHBB), Basel.
- [4] Gomez Advisors, Inc. (2000): [<http://www.gomezpro.com/index.asp>], [accessed 16.08.2000].
- [5] Helft, Miguel (2001): Reality Check, Aisle 5, in: *The Standard*, [<http://www.thestandard.com/article/0,1902,24035,00.html?nl=dnt>], 30.04.2001. [Zugriff: 10.06.2001].
- [6] JurisNET GmbH (2000): *JurisNET-Gütesiegel*, [<http://www.jurisnet.ch>], [accessed 10.08.2000].
- [7] Kamenz, Uwe (2000): *Perfekte Webseiten - wie sieht die Realität aus*, in: Pförsch, Waldemar A. (Ed.), *Living Web: Erprobte Anwendungen, Strategien und zukünftige Entwicklungen im Internet*, pp. 21-33, Landsberg: verlag moderne industrie, 2000.
- [8] Liu, Chang; Arnett, Kirk, P.; Litecky, Chuck (2000): *Design Quality of Websites for Electronic Commerce: Fortune 1000 Webmaster's Evaluations*, in: *The International Journal of Electronic Commerce & Business Media*, Vol. 10. No. 2, pp. 120-129.
- [9] Malhotra, Yogesh; Galletta, Dennis, F. (1999): *Extending the Technology Acceptance Model to Account for Social Influence: Theoretical Bases and Empirical Validation*, in: *Proceedings of the 32nd Annual Hawaii International Conference on System Sciences*, [<http://www.computer.org/proceedings/hicss/0001/00011/00011006abs.htm>], [accessed 08.08.2000].
- [10] Platoni, Kara (2001): *The Last Mile*, in: *Eastbay Express Online*, [<http://www.eastbayexpress.com/issues/2001-05-25/feature.html/page1.html>], 25.05.2001. [Zugriff: 10.06.2001].
- [11] Schubert, Petra (2000): *Virtuelle Transaktionsgemeinschaften im Electronic Commerce – Management, Marketing und Soziale Umwelt*, Lohmar, Köln: Josef Eul Verlag, 2000.
- [12] Schubert, Petra; Selz, Dorian (2000): *Measuring the Effectiveness of Electronic Commerce Websites with the Web Assessment Method*, in: Hunt, Brian (Ed.), *E-Commerce and V-Business*, London: Imperial College, 2001.
- [13] Selz, Dorian; Schubert, Petra (1998): *Web Assessment - A Model for the Evaluation and Assessment of successful Electronic Commerce Applications*, in: *Proceedings of the 31st HICSS Conference, Hawaii*, Vol. IV, 1998, pp. 222-231.