CHI-SA 2000

Human-Computer Interaction Conference

May 8-10 Pretoria

A Usability Metric for Internet-based Banking: Lessons learned

Tobie van Dyk
Department of Computer Science and Information Systems
University of South Africa
Pretoria

Telephone: +27124296676 (w)
E-mail: vdyktj@unisa.ac.za

Content

(1) Introduction
(2) Internet Banking
(3) Task analysis
(4) Web usability and its evaluation
(5) Conclusion and lessons learned
(6) List of references

Abstract:
The application of general usability principles and properties to Internet-based banking is reviewed. Existing usability guidelines for hypermedia and Web design are examined, and applied during a structured evaluation of local and international online banking facilities. A low-cost usability metric is proposed that is a combination of a checklist-based and heuristic evaluation. Such measurements should be preceded by a structured task analysis approach.

Key terms:
Usability, Human computer interaction, HCI, Hypermedia, Hypertext, Heuristic evaluation, Checklist evaluation, Task analysis, Task analysis for knowledge description, TAKD, E-commerce, Internet banking, PC banking, Home banking, Online banking, Web page design.

1 Introduction

Internet banking is but one example of an application in the rapidly expanding Internet E-commerce domain. It has many of the typical characteristics of GUI software products in addition to its expected hypertext/hypermedia inheritance.

Accurate, repeatable, and cost-and-time efficient usability measurement approaches and techniques are a precondition for an interaction quality improvement – else any usability changes may go unnoticed.

The research study (van Dyk, 1999), covered a two-year period from April 1997 to May 1999. Four of the local banks that offered Internet banking facilities (ABSA, Nedbank, Standard Bank, and First National Bank), and a selection – based on a spread in observed usability – of six major international banks (such as the Bank of America), constituted the research sample.

For one of the local banking sites (Standard Bank), an in-depth longitudinal study was conducted. The observation of Standard Bank's facility is continuing.

The main evaluation instruments were a set of adapted heuristic evaluation guidelines and an evaluation checklist. These are both low-cost usability evaluation methods – i.e. discount methods – and are exemplary examples of this high-yield low-resource-cost approach.

These evaluation approaches all suffer from one important deficiency – ***it is difficult to obtain reliable quantitative and baseline data through their use.*** As normally implemented they are more suitable for yielding comparative and qualitative information. Approaches where severity and frequency ratings are assigned to heuristic usability problems and classes of problems can ameliorate this. Although a checklist may seem to generate statistically derived absolute usability values through scored rating scales these values remain largely subjective.

It can be shown that the general usability principles and properties also apply to the design of Web-based software products. Web-based applications have an increased focus on certain usability dimensions, and are characterised by a very wide range of client characteristics. In addition it will be advantageous to integrate the large pool of hypermedia and hypertext research into Web usability studies.

Because ***Internet-based banking is a task-oriented, goal-directed, Web-based application*** (where users access and manipulate bank accounts through a Web browser interface), the advantages associated with the ***use of a structured approach such as task analysis (and specifically TAKD) of the banking task in uncovering usability problems*** are examined, and suggestions for improving the interaction quality through this approach are offered.

2 Internet Banking

Lagoutte (1996:58) notes that *'We are witnessing the beginnings of a fundamental shift in retail banking. Tomorrow, the bank will not be concentrating on achieving competitiveness across the board. Rather, fostered by deregulation and technological change, it will revolve around achieving excellence in only a small number of areas in which it genuinely believes it can achieve advantage. Retail banks today depend heavily on their branch networks and see it as a core of their business. In the future the branch network is but one of the several channels with which the bank accesses its stand-alone client base. Indeed, the branch network is in the future a privileged channel, used only for accessing that small percentage of the client base that offers the potential for a genuinely profitable relationship. The rest of humanity will deal with technology to transact its banking. Alternatively, it will use the telephone or a home TV based system as its access to retail banks.'*

Home Banking is defined as the use of proprietary financial software running on a computer in the home, to perform transactions such as fund transfer and bill payment. Banks and building societies usually allow their proprietary software to be downloaded from their Web pages. ***PC Home Banking*** refers to banks (or building societies) sending out a financial software package on disks, allowing users to fill in details off-line and then to send them into the bank over the bank's private network. ***Internet Banking*** is similar in functionality to Home Banking, with the important difference that it allows access to accounts from a standard Web browser, eliminating the need for proprietary software. (Hennigan and Gourvennec, 1998.)

E-commerce. This is conducting business communication and transactions through networked computers and digital communications. It includes the buying and selling of goods and services, and the transfer of funds over the World Wide Web. Also included are all forms of electronic intra-company and inter-company network-based functions that enable commerce such as EDI, video conferencing, e-mail, workflow, and remote interaction (Schneider and Perry, 2000).

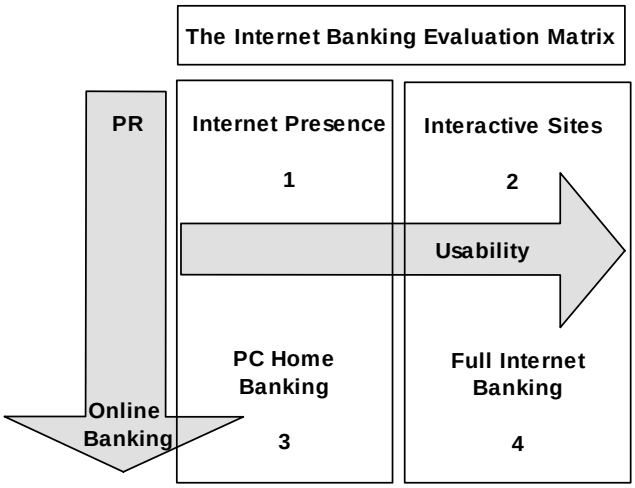
Hennigan et al. (1998) offer the following comparison between four classes of banks with a Web presence (figure 1). They compare the site purpose (public relations to full online banking) on one dimension to 'user-friendliness' (usability) on the other dimension – correctly Internet browser-based banking is indicated by them to be the most usable of the four classes.

Their four classes are: (1) **Net presence**. This is the most basic category, describing sites that concentrate on providing information to their clients. (2) **Interactivity**. This is the next step up, describing sites that attempt to be interactive with the user, through for example feedback forms and e-mail contact details. (3) **Home banking** and (4) **Internet banking** are as defined previously.

The 'Big Four' group of South African banks started offering Internet-based online banking services from the end of 1996. ABSA (Amalgamated Banks of South Africa) – which previously included United, Trust, Allied and Volkskas Banks – was the first to offer an initially limited in range online facility in December 1996. In the period January to May 1997

Nedbank (including Peoples Bank and Permanent Bank), First National Bank and lastly Standard Bank established their banking services.

Figure 1 Internet banking evaluation matrix



Source: Adapted from Hennigan et al. (1998).

Standard Bank (URL: Sbic), introduced their internet-based banking service during the May 1997 Computer Faire at Gallagher Estates in Midrand (South Africa), about two years after their first (non-interactive) presence on the Internet. Standard Bank claims in their 1997 information brochure that within only eight days of going live, the 1000th client registered for their Internet banking service. Currently (April 2000), Standard Bank has about 120 000 Internet-banking clients, and this is growing at about 6 000 clients per month.

These Internet-based services allow banking clients to obtain account information, balance enquiries, execute third-party account payments, and inter-account fund transfers. Additionally, by linking their accounts to personal finance software (such as Intuit Quicken and Microsoft Money), they will be able to track their spending offline, and later reconcile that with their bank statements online. The South African Banks Act restricts a bank's ability to provide certain online services such as applications for a credit card, as original paper documentation is required.

The Financial Mail (October 10, 1997) reports that in the USA increased competition, deregulation, and the spread of electronic banking reduced the number of banks by more than a quarter between 1985 and 1994. Such a banking consolidation is now evident in Europe and has also been observed here – as witnessed by the local mergers that led to the ABSA, Nedcor, and NBS-Boland (and possibly Nedcor-Stanbic), groups of banks, as well as the declining fortunes of niche banks such as the African Bank and Islamic Bank of South Africa, and the emergence of the availability of 'in-house' banking services from certain of the larger retail groups. Electronic banking offers a route to salvation for this group of 'traditional' banks, which believed in the past that their (many) branches offered them distinct advantages, but these have now proved themselves nothing more than costly millstones.

For the banks, one of the most obvious advantages in offering these online facilities is cost. Table 1 shows the relative costs to the bank per transaction for the various channels. Although the table was compiled for US Banks the *'figures should be enough to make any bank, in any country, sit up and take notice'* (Hennigan et al., 1998).

Table 1 Channels and transaction costs

Channel - Cost/Transaction
1. Branch full service: $ 1.07
2. Telephone average: $ 0.54
3. ATM-full service: $ 0.27
4. PC Banking: $ 0.015
5. Internet Banking: $ 0.010

3 Task Analysis

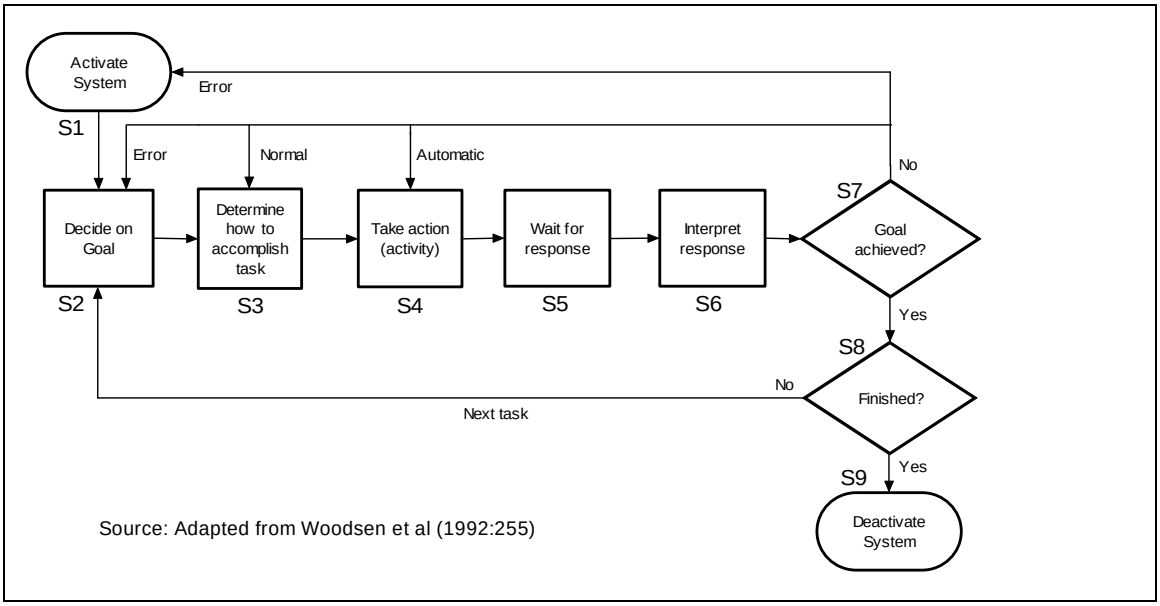
Task analysis may be defined as the study of what a user is required to do in terms of actions and/or cognitive processes to achieve a task. A detailed task analysis can be conducted to understand the current system and the information flows within it. These flows are important to the maintenance of the existing system and should be incorporated in any new or replacement, system (Maguire, 1997). Kirwan and Ainsworth (1992:1) present a similar definition but they substitute 'to achieve a task' with 'to achieve a system goal'. Newman and Lamming (1995:20) also emphasize the goal-directed nature of tasks when they state succinctly that *'a task is a unit of human goal-directed activity'*.

Johnson, Diaper and Long (1984:500) gave a motivation for the use of task analysis: *'The evolution of computer applications and the need for good human-computer interaction has created a demand for knowledge-based descriptions of human task performance'*

One specific task analysis approach – TAKD (**Task analysis for knowledge description**) – had been applied to an Internet banking task where the user logs into the banking facility and displays a balance for a selected account. The advantages associated with the use of such a structured approach to the analysis of the banking task in uncovering usability problems had also been examined (Van Dyk, 1999).

An often-encountered, and simple computer operation model is presented in figure 2 – this specific example being adapted from Woodson, Tillman, and Tillman (1992:255), which is in turn based on the set of USA Military User Interface Design Guidelines. This type of model is eminently suitable to be used as the basis for formulating a strategy for task analysis. It is noted that Norman's execution-evaluation model (as discussed by Dix, Finlay, Abowd, and Beale, 1998:105-106), has stages similar to this model. Of particular interest is the reduction of Norman's seven stages of user activity into the three as presented by Diaper and Addison (1991:389), namely **establishing goals**, **behaving**, and **perceiving. As** a result this type of model has an immediate and direct application to task analysis, of which task sequence construction, and Diaper's object/action pairs task 'generification', are but two examples.

Figure 2 Task analysis based on a simple iterative computer operating model



Notes: Figure 2

(1) The S7-S4 feedback loop labelled 'automatic' corresponds to Johnson's (1989:121) notion of 'procedures' or well-practiced behaviours.

(2) The error feedback loops S7-S2 and S7-S1 are included, as they are especially appropriate viewed within the unreliable communication context of Internet-based interaction. The 'normal' assumption of system deactivation only on completion of the task(s) is then often untrue.

(3) The original (Woodsen's) term 'entry' was replaced by 'action' in S4 to correspond to the five-level hierarchy from job to action – a set of these (i.e. multiple loop backs), would be an activity.

(4) Johnson's (1989:121) definition of a goal: ' a purpose or objective to be achieved, it provides the purpose for which a task is undertaken', is appropriate within the context of this model.

(5) The S9-S2 loop represents the execution of different tasks within single system activation – it corresponds to the previously discussed notion of a task-procedure or project.

(6) S2 could be relabelled as 'Goal adjustment' if it is reached via an error feedback loop.

3.1 Task analysis for knowledge description (TAKD)

TAKD is a method for task analysis that generates a hierarchical description of tasks – the task descriptive hierarchy (TDH) – and translates these into a set of knowledge representation grammar (KRG) sentences. Its original purpose was to analyse data from the observation of relevant tasks and then re-describe them using a single consistent representation of specified information technology knowledge that the user should possess (Diaper, 1989:108). Importantly, Johnson, et al. (1984:499), originally described TAKD as a *'method for producing descriptions of knowledge'*. It is therefore a descriptive rather than normative approach to activity modelling for task analysis (Newman et al., 1995:114). TAKD is also suitable for the evaluation of existing systems – provided the focus is on behaviours that can be measured and described rather than being broadly psychology-focused, and it can also be used to include a task analysis stage during the production of requirements specifications.

Importantly Diaper and Addison (1992:127) emphasize that: *'TAKD places few constraints on analysts, but it does force a systematic, explicit way of thinking about tasks and their more*

general context'. They continue: *'TAKD (expressly intends to be limiting in that it) forces an inherently user-centred perspective, because TAKD actions are always those of people, (and) not of machines'*. The focus of this analysis is then on the observable behaviours of the user rather than the computer, and the value of the analysis was proved to be realized as much from simply completing the exercise as it was from TAKD's formal outputs (such as the TDH diagrams and KRG sentences). This does not mean that TAKD ignores agents other than the user – indeed Diaper (1989:111) states that TAKD concentrates on the complete system, but it only includes these other aspects where they have an impact on the user behaviour.

3.2 Applying TAKD to an Internet-based banking task

The five stages of TAKD – **construction of a activity list (AL), selection of specific objects (SO) and specific actions (SA), construction of the task descriptive hierarchy (TDH),** the **specification of specific and generic knowledge representation grammar (KRG),** and the **specification of task sequence representation –** described in Van Dyk (1999), formed the basis of an application of TAKD to an actual Internet-based banking application. (Actual refers to 'real' or 'live' banking accounts and clients as opposed to using an Internet-based banking demonstration Web site.) The simple computer operating model (figure 2) was used to formulate the steps of this Internet banking task.

The problem area is the execution of a simple banking task – obtaining a detailed statement for a current account – via the Internet. This task is preceded by other tasks: obtaining Internet access, navigating to the Bank's (Standard Bank South Africa) Internet banking Web site, and logging into the site. The task is concluded (the goal achieved), when the statement has been viewed, and the user logged off from the Bank's Web site. During this analysis the three tasks of logging in, obtaining the statement, and logging off had been analysed as a single task with component activities (sub-tasks). This was followed by the construction of the AL, selection of SO and SA, construction of the TDH and the specification of the specific and generic knowledge representation grammars. These stages are not reproduced here as they are rather lengthy – they are described in full in Van Dyk (1999). One final result of the analysis is a set of KRG's as presented in table 2:

Table 2 Final set of user KRG's

Internet-banking-task (*) *press button/function(proceed)/* *Internet-banking-task (do-transaction)* *press button/function(select)/* *Internet-banking-task (*)* *select menubar (option)* *Internet-banking-task (login)* *enter (account/authorization/)*

What are considered to be some of the more important analysis results are presented below:

The small number of generic KRG's. Apart from showing this to be a simple task, it more importantly also emphasizes that:

(a) Internet-based HCI is at present operating within a very (technology) restrictive task environment.

(b) It is a relatively new technology that has yet to evolve towards complexity. It has few interaction variations. These mainly consist of a few simple actions executed repetitiously. Within this new application domain Internet-based banking is even newer.

Table 4 Frequency analysis of the final KRG's

The observations above are supported by the frequency analysis of the final KRG's – for the user there are only four final KRG's with the following frequencies of occurrence:

Partial KRG	**Frequency**
press button/function(proceed)/	**5**
select menubar (option)	**2**
enter (account/authorization/)	**2**
press button/function(select)/	**1**

Similarly for the computer the actions are of two types only (excluding the invisible computer processing actions):

prompt user (do one of four generic actions)	**13**
display page	**12**

Human-computer interaction within this specific task environment is therefore characterized by a few simple actions that may be repeated many times. This has obvious favourable implications for task knowledge requirements and for usability learnability, but it does post some usability warnings: the large number of similar actions will tend to make the interaction boring, and the inherent delayed response resulting from unnecessary communication through these extra actions will further decrease the usability and reliability of the interaction.

Interestingly, Diaper et al. (1991:397) suggest that frequently occurring task elements as described in the TDH are more likely to be treated in a cognitively automatic manner than infrequent task elements. These actions are shown in the simple computer-operating model (figure 2), through the feedback loop S7-S4. These automatic actions would include the final KRG's 'press' and possibly also 'select'. An action that would not be automatic (in part because of its infrequency, but more likely because of the importance derived from the consequence of its activation), would be the transfer of funds, and even the act of logging in – i.e. all tasks that require a form of data entry (the 'enter' KRG).

Coincidentally, the small set of final generic KRG's is a vindication of TAKD's appropriateness for this type of analysis – any other result would have been counter-intuitive and confusing.

An examination of the four user-agent KRG's show two of these to be related to a navigation action (*press button/function(proceed)/ and select menubar (option)*), with a frequency of seven, and only three other action/object pair instances being part of the task. This result, given the nature of the task is hardly surprising: Interacting with an Internet-based application (and hypertext structures in general), is mainly about navigation, and this aspect of usability deserves the attention it usually gets. The primary functionality of the system is therefore related to these navigation actions and the focus should be on making it possible for the user to access these functions safely and easily.

The frequency results also indicate where an unreasonably large user effort is required. This is also related to reducing the need for navigation.

It should be noted that Standard Bank did redesign their banking Web site during early 1999, which then reduced the navigation effort. It was then possible to reach the Internet banking sign-on page directly from the home page.

4 Web Usability

'The smallest of usability problems, when multiplied across thousands or millions of users, becomes a source of massive inefficiency and untold frustration' – Cockburn and Jones (1996:105), quoting Sun's Jakob Nielsen (1993).

Thimbleby (1997) cautions that the enormous size, number and variety of users, and dynamic nature of the Web do mean that hypertext usability problems are likely to be magnified by a 'factor of millions'.

4.1 Heuristic evaluation for the Web

All of the general usability principles (learnability, flexibility, and robustness), and their properties (such as consistency, customizability, and responsiveness), can be applied with good effect to hypermedia-hypertext designs. Many of the Web-specific usability guidelines are either relabelled general usability properties, or combinations of these, and could also qualify for inclusion as general evaluation criteria. Certain of these Web-specific criteria combine usability properties through an overlap between the three classes of general usability principles – but the boundaries for these three are in any case not always distinct – as shown in figure 3. Mayhew (1998:4) correspondingly observes that the difference between Web (interface) design and traditional software design is a matter of degree – the knowledge acquired in designing these latter systems can be effectively applied to the World Wide Web but seldom are.

It can therefore be concluded that what is significant is that the importance of some of the general criteria or usability properties is amplified (or prioritised) for Web-based or hypermedia applications. An aspect such as navigation, which is related to general properties such as reachability, user guidance, and browsability, is one example of this. For many entertainment focused Web sites properties such as enjoyability and responsiveness (for streaming audio and video applications) are disproportionately important. Other Web sites that focus on the presentation of their information content should be evaluated more strongly in terms of properties such as visual clarity, user guidance, and task conformance.

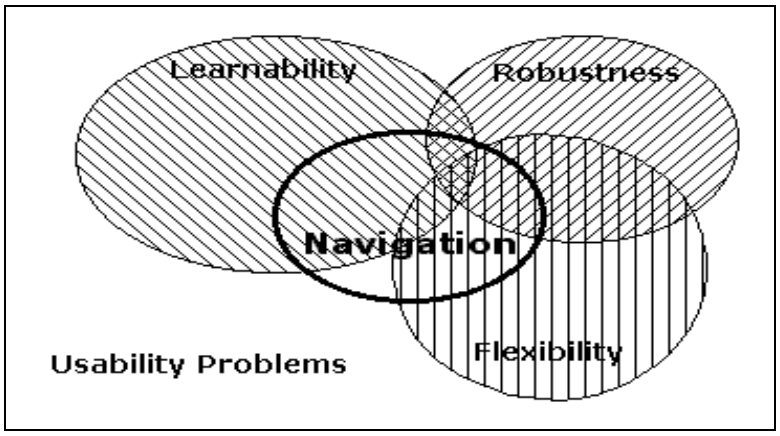
Bevan (1997, URL: NPL1) also examines some of the more important usability issues in Web site design. He regards the main reasons for the slow and confusing use of many Web sites to be due to the following: (1) The Web site has a content and structure that mirrors the internal concerns of the organization rather than that of the user. (2) The Web site contains material that is appropriate in printed form, but needs to be adapted for Web presentation. (The

opposite case also frequently applies to Internet banking Web sites.) (3) Web site publishing has now become very easy, and it is usually not subject to the same quality criteria used for other forms of publishing. A fourth major cause should be added: Usability-naive Web site and page designers.

Mayhew (1998:7), continues and recommends that: *'Usability engineers must take traditional style guides, which have been rigorously developed using usability engineering techniques, and tailor them to be Web-specific – not by shortening them or even changing the basic principles, but by presenting new examples of the realization of these general principles on actual Web sites'*. This statement summarizes the approach as used in Van Dyk (1999).

The previously discussed usability principles and properties can be presented as a number of overlapping guidelines sets for finding solutions in the surrounding usability problem space. Similarly Web-specific guidelines discussed in detail by Van Dyk (1999), can be presented as another overlapping set cluster that superimposes differently on the first (general) set cluster. The term superimpose is chosen carefully – it is unfortunate that the interaction between usability properties in the Web application domain is often regarded in a juxtapositional sense.

Figure 3 General and Web-specific usability guidelines in the usability problem space



This view, which is presented in figure 3, highlights another problem when choosing between general usability guidelines and Web-specific guidelines – they address usability issues from a different viewpoint. For example the important Web (and hypertext) aspect of navigation needs to be analysed in terms of visual clarity, user support/guidance, conceptual models, consistency, familiarity, predictability, synthesizability, control, observability and recoverability. (Note that the navigation set in figure 3 is intentionally drawn so as to indicate that the intersection of this set with learnability is larger than for the other two sets.) A similar observation can be made for most of the other Web-specific guidelines (Van Dyk, 1999).

The solution chosen was to conduct the evaluation first on the basis of the general usability guidelines and then additionally, in terms of the more important Web-specific guidelines. This will ensure that those Web-specific aspects that were not covered by the general usability guidelines (unlikely!) are included. (These orphaned guidelines would be the

complement of Navigation with the union of Flexibility, Robustness and Learnability in figure 3.)

In this regard Nielsen (URL: Useit2) refers to Dykstra (1993) and notes: *'In addition to the checklist of general heuristics to be considered for all dialogue elements, the evaluator obviously is also allowed to consider any additional usability principles or results that come to mind that may be relevant for any specific dialogue element. Furthermore, it is possible to develop category-specific heuristics that apply to a specific class of products as a supplement to the general heuristics. One way of building a supplementary list of category-specific heuristics is to perform competitive analysis and user testing of existing products in the given category and try to abstract principles to explain the usability problems that are found.'*

The previously mentioned aspect of criteria prioritisation was handled by noting that during the evaluation those usability aspects that are important for the Web (such as navigation), tended to be naturally examined and discussed in more detail than the other aspects. The relative size (area) of the guideline sets in figure 3 can be used to represent this aspect of prioritisation. Web-specificity was handled by being careful to include all of the more important Web-specific heuristic aspects – as identified in table 5 – during the evaluation.

Table 5 then lists the categories of usability heuristics that had been used for Standard Bank's (and the other banks studied) Web site evaluation. The navigation properties presented in this table were extracted from Whitaker's (1998:69) ten principles for navigation design.

Table 5 Experimental heuristics usability guidelines

General Usability Principles and Properties		**Web-specific Guidelines**	
Learnability	Synthesizability Consistency Familiarity Visual clarity User support and Guidance Predictability Generalizability Enjoyability Interaction Metaphors Conceptual Models	**Navigation**	Cues, prediction, landmarks, consistency, situational awareness, depth, user's purpose and proficiency, alternatives.
		Aesthetic and minimalist design	Dialogs, information relevance and visibility.
Flexibility	Dialog Initiative Multi-threading Task migratability Substitutivity Customizability Control	**Match between system and the real world**	User's language, familiar concepts, natural and logical ordering of information
Robustness	Observability Recoverability Responsiveness Task conformance WYSIWYG	**Consistency and standards**	Platform conventions, terminology.

4.2 Checklist-based usability evaluations.

These correspond to the Bastien and Scapin (1995:184) category of expert judgement methods, as examined in Van Dyk (1999). It is a subjective evaluation approach, and is

closely related to the heuristic evaluation methods as examined in the preceding section. It is based on training, field experience, and an examination of human factors data. This implies that the evaluator should be careful to avoid rating bias (also see Ravden and Johnson (1989:22,86), possibly introduced through aspects such as central tendency, leniency error, and cross-linking (Cooper and Emory, 1995). Some degree of bias will nonetheless unavoidably be present during the study.

4.3 The experimental usability evaluation checklist

The Ravden et al. (1989), checklist had been designed some time ago for a general HCI usability evaluation. A study of the specific questions contained therein shows that perhaps unsurprisingly – it has been proven to be a well-constructed design – almost all of the questions are nonetheless applicable to a Web-based evaluation. This is especially true for a genuine goal-oriented application such as Internet banking. It was therefore decided – once again based on the author's earlier experience (Van Dyk, 1996) in applying this list to a usability evaluation problem – to use the questionnaire with only minor modification.

Table 6 The Ravden usability questionnaire section guidelines

Section	Description
1	**VISUAL CLARITY**: Information displayed on the screen should be clear, well organized, unambiguous and easy to read.
2	**CONSISTENCY**: The way the system looks and works should be consistent at all times.
3	**COMPATIBILITY**: The way the system looks and works should be compatible with user conventions and expectations.
4	**INFORMATIVE FEEDBACK**: Users should be given clear, informative feedback on where they are in the system, what actions they have taken, whether these actions have been successful and what actions should be taken next.
5	**EXPLICITNESS**: The way the system works and is structured should be clear to the user
6	**APPROPRIATE FUNCTIONALITY**: The system should meet the needs and requirements of users when carrying out tasks.
7	**FLEXIBILITY AND CONTROL**: The interface should be sufficiently flexible in structure, in the way information is presented and in terms of what the user can do, to suit the needs and requirements of all users, and to allow them to feel in control of the system.
8	**ERROR PREVENTION AND CONTROL**: The system should be designed to minimize the possibility of user error, with inbuilt facilities for detecting and handling those, which do occur; users should be able to check their inputs and to correct errors, or potential error situations before the input is processed.
9	**USER GUIDANCE AND SUPPORT**: Informative, easy-to-use and relevant guidance and support should be provided, both on the computer (via an online help facility), and in hard-copy document form, to help the user understand and use the system.
10	**SYSTEM USABILITY PROBLEMS**: When using the system, did you experience any problem in terms of the listed questions?

Source: Ravden and Johnson (1989).

A single page 16-question section based on the Journal of Internet Banking and Commerce (URL: JIBC1) evaluation checklist (discussed in the next section), was added to the checklist evaluation. This additional checklist is specifically aimed at evaluating Internet banking applications and was the most comprehensive example found (Van Dyk, 1999). The section guidelines for the Ravden checklist are presented in table 6. Because the Ravden checklist itself is a lengthy ten-page questionnaire it is not presented here but can be found in Van Dyk (1999) as a scored table. The additional JIBC evaluation is presented in table 7.

Table 7 JIBC Internet banking evaluation checklist		
No.	**Criteria description**	**10 Point polar scale identifiers**
1	Overall effort/commitment to create a functional online service	Low – High
2	Service speed	Slow – Fast
3	How often the system dies when online	Rarely – Often
4	Convenience of performing simple transactions through the Net	Not at all – Highly convenient
5	Visible evidence of service's independent security certification	Non existent – Displayed on the Homepage
6	Documentation/passwords send to prospective clients at the start of the service	Confusing – Clear
7	Simplicity/intuitive character of online forms and work processes	Confusing – Clear
8	Illogical interface deviations from the printed statements of the same bank	Substantial – Insignificant
9	Ease of help access through email	No access – Clear access points on every page
10	Clarity/meaning of interactive messages	Meaningless – Clear
11	On site explanations of banking terms	Inadequate – Adequate
12	Ease of help access through the voice call center	Incompetent, busy – Competent, easy to reach
13	Misbehaving/superfluous Java applets and scripts on the site	All-the-time, everywhere – Moderate
14	Graphics that are slow to download and are difficult to the eyes	Flashy, garish, moving, non-functional – Moderate
15	Advertising that is slow to download and is difficult to the eyes	Flashy, garish, moving, destructive – Reasonable
16	The likelihood that the bank's own CEO is personally using this service	Highly unlikely – Very likely
Source: www.arraydev.com.		

4.4 Discussion

The heuristic usability evaluation of Standard Bank and the other banks indicated emphatically that general usability principles and properties can be applied with good effect to the hypermedia-hypertext designs as found on the World Wide Web. But for this type of specialized application, i.e. online banking, a restricted or qualified implementation of some of these usability properties is argued for. Usability properties such as flexibility, browsability, reachability, system versus user control, multithreading, user and system pre-emption and dialog initiative, can not be promoted to the same extent for these secure, high risk systems when compared to other (insecure or less risky) applications. All the properties that constitute the principle of learnability such as consistency, positive engagement, mental model support, metaphors, clarity, and predictability, do apply meaningfully and strongly to this type of application.

Based on these evaluations the suspicion grew that for the local banks technical developers rather than human computer interface experts were responsible for the technical, layout, and interface aspects of their banking sites. In contrast contributions from such human-oriented designers, and graphical presentation experts, are evident during a study of the Web sites of American banks.

The examination of the Internet-banking facilities offered by the four local banks was illuminating. Standard Bank improved the usability of their site considerably between May 1997 and April 1999 – especially for aspects such as reduced navigational complexity, enhanced task support, improved visual clarity, and faster response. Their site must rank with First National Bank as the most usable and functional of the local sites evaluated. Indications

are however (Parbhoo, 2000), that the Nedbank Web site is currently (January 2000), much improved and perhaps the most usable locally.

This (heuristic) approach to Web site evaluation should be considered exploratory in nature – a more structured evaluation, that for example includes a proper task analysis of the various (usually simple), tasks executed during online banking, can yield other valuable usability data related to task conformance and user conceptual support. The research should also be expanded to include a more detailed study of the now large number of overseas online banking organizations. What was found of particular significance and help during this evaluation was the observed and documented change in Standard Bank's site during the two-year period – the comparison of 'old' versus 'new' site structure yielded usability information that would have been impossible to obtain otherwise.

Even though the JIBC checklist is custom-designed for an evaluation of Financial Institutions such as Internet banks it exhibits several weaknesses. Firstly it is not a good reflection of the site usability – restricting the list to 16 questions unavoidably required discarding many important usability test factors. Secondly it has not been applied to a sufficiently large group of financial institutions to enable a valid conclusion to be made regarding the meaning of the usability scores. In addition it has not been widely applied to Internet banks and there seems to have been little response to the invitation to help to develop the questionnaire – it has remained unchanged over a long period. In its defence it should be noted that this was by far the most detailed usability-related checklist found on the Web, and that it is preferred to the simple four to five point, six category rating methods as employed by other sites (Van Dyk, 1999).

In contrast the Ravden checklist is considered to have once again proved it worth. It is possible to use it with minor changes to evaluate hypermedia applications such as Internet banking. It could be argued that some questions such those deals with monochrome displays are outdated. But this specific question also caters for disabled (colour-blind) user access – if the checklist is to be criticized then it should be on the moderate neglect shown by the checklist for this type of user.

The three lowest scores in from the Ravden checklist are for explicitness, user guidance, and flexibility – this result confirms the results obtained from the use of the JIBC checklist. User guidance suffered during this evaluation because of the absence of help during transaction stages, poor page layout, and the absence of hard-copy documentation. The highest scores are shown for error prevention and consistency. The first is (in part) due to the security-conscious design of the application, and the second because the application runs and behaves similar to a standard Windows application.

The Web community's professed quest for quick and simple solutions to complex usability problems will not succeed, and the solution does not lie in abbreviating the measurement instrument to a point beyond all usefulness and accuracy. Nielsen (1993:17, 175) is correct in noting that usability heuristic evaluation methods (which includes checklists), should be regarded as **discount approaches** – i.e. a comparatively large return is shown for a relatively

small investment in time, resources, and expertise. This does not mean that reaching the point where such a return will be shown will be easy, and achieving this will require at least a reasonable degree of effort, time, and knowledge.

5 Conclusion and lessons learned

Both heuristic and checklist-based usability evaluations are suitable measurement instruments for Web site and page usability. Existing (Web-sourced) evaluation services tend to be feature (task-support) oriented and neglect other more important usability aspects. Although guidelines designed specifically for Web site evaluation do have merit – especially for the usability-naïve Web designer – it is considered more rigorous to base an evaluation on general usability standards, principles, and properties. These should then be expanded by a carefully selected addition of Web-specific guidelines.

The prevalence of low usability Web sites is in part also due to the easy access of Web developers to the application domain – few significant barriers of entry exist. This implies that many usability-naive and software-development-inexperienced workers operate in this domain. In this context the various discount engineering methods such as proposed by Nielsen (1993) and Instone (1998) offer a solution. These are **low resource** (time, money, and skills) – **high yield** (usability improvement) methods, which are suitable for use by the inexperienced Web developer.

A large proportion of usability problems for the local banking Web sites originate from using technology-wise developers for building the interaction interface and page layout and design, in addition to the backroom technical systems and functions. Local banks should follow the example of overseas banks and use multi-disciplinary development teams with a broad range of skills. Human, psychology, graphical design, and document design skills should be added to their (in-house and outsourced) technical development and maintenance teams.

The post-evaluation measures as described by Levi and Conrad (1996) are impressive. Their assignment of severity ratings to usability violations appears to work well, and yields important additional qualitative data. It also partly solves a difficult problem during heuristic inspections namely the issue of comparison against an objective qualitative baseline or standard.

Nielsen (2000) emphasises an important reason why usability is so important for the (commercial) Web – he speaks of an ***inversion of the usability experience*** for the Web when compared to other intangible (such as software) but also tangible products. Web users experience usability at a much earlier stage and bad impressions will impact more severely on reducing their brand loyalty and relocating their buying preference.

Usability metric

It is concluded that no suitable single Web-specific usability metric exists that will satisfy both qualitative and quantitative evaluation requirements – a combination of evaluation approaches is required. A measurement instrument combination similar to that used for this

research project has since been applied during the development of the NBS (Natal Building Society) Internet banking site with some success (Parbhoo, 2000) – this evaluation reassuringly also yielded useful pointers to usability problems that may have else gone unnoticed.

Implications for future research and practice

This study makes no claim to have solved the difficult problem of reliable comparable Web usability ratings, but a study of the existing Web-based evaluation services in combination with traditional usability property-based evaluations offered important guidelines as to where future development work could be directed to, and importantly can also demonstrate what does *not* work during these evaluations.

The Internet and its hosted applications remain a remarkably immature technology and application domain, given the explosive growth in its user count. This conjuncture of an immature technology, a distributed delivery and interaction, an immense heterogeneous client base, and a usability-naïve Web developer community, will continue to produce large numbers of both unique and common usability problems, but with an impact that surpasses by far anything else that has previously been encountered.

6 List of references

Bastien J M C and Scapin DL. 1995. **Evaluating a User Interface with Ergonomic Criteria**. International Journal of Human-Computer Interaction 7(2).

Bevan N**. 1997**. URL: NPL1. http://www.npl.co.uk/npl/sections/us/frames/fweb.html

Cockburn A and Jones S. 1996. **Which way now? Analyzing and Easing Inadequacies in WWW Navigation.** International Journal of Human-Computer Studies. 45.

Cooper D R and Emory C.W. 1995. **Business Research Methods.** 5th Ed. RD Irwin.

Diaper D. 1989. **Task Analysis for Human-Computer Interaction.** Ellis Horwood Ltd.

Diaper D and Addison M. 1991. **User Modeling: The Task Oriented Modeling (TOM) Approach to the Designer's Model.** In People and Computers VI. (Ed.) Diaper D and Hammond N. Cambridge University Press.

Diaper D and Addison M. 1992. **Task Analysis and Systems Analysis for Software Development.** Interacting with Computers. 4(1).

Dix A, Finlay J, Abowd G and Beale R. 1998 (2nd Ed.). **Human-Computer Interaction**. Hemel Hampstead: Prentice-Hall International (UK).

Dykstra, D J. 1993. **A Comparison of Heuristic Evaluation and Usability Testing: The Efficacy of a Domain-Specific Heuristic Checklist.** Ph.D. Department of Industrial Engineering, Texas A&M University, College Station, TX.

Financial Mail. 1997. **Red Carpets for the Rich**. 10 October 1997.

Hennigan M and Gourvennec Y. 1998. **Internet Banking in Europe.** URL: http://www.unisys.com.

Instone K. 1998. **Usability Engineering for the Web**. URL: http://www.w3j.com/5/s3.instone.html

Johnson P, Diaper D and Long J. 1984**. Tasks, Skills and Knowledge: Task Analysis for Knowledge Based Descriptions.** Human Computer Interaction – INTERACT '84.

Shackel B. (Ed.) Elsevier-Science Publishers.

Johnson P. 1989. **HCI models in Software Design: Task Oriented Models of Interactive Software Systems.** In Software Engineering Environments: Research and Practice. Bennett K H (Ed). Ellis Horwood.

Journal of Internet Banking and Commerce. 1998. **URL: JIBC1**. http://www.arraydev.com/commerce/JIBC/

Kirwan B. and Ainsworth L K. (Ed.) 1992. **A Guide to Task Analysis.** Taylor and Francis.

Lagoutte V. 1996. **The Direct Banking Challenge**. Research project. B.A. (Hons) (European Business Administration). Middlesex University.

Levi M D and Conrad F G. 1996. **A Heuristic Evaluation of A World Wide Web Prototype.** Interactions. July-August 1996.

Maguire M. 1997. **RESPECT User Requirements Framework Handbook.** HUSAT Research Institute, The Elms, Elms Grove, Loughborough, Leicester, UK.

Mayhew D. 1998. **Introduction to Human Factors and the Web**. In Forsythe C, Grose E and Ratner J. (Ed). 1998. **Human Factors and Web Development**. Lawrence Erlbaum Associates.

Newman W M and Lamming M G. 1995. **Interactive System Design.** Wokingham: Addison Wesley.

Nielsen J. 1993. **Usability Engineering**. AP Professional Academic Press.

Nielsen J. 1998. **How to Conduct a Heuristic Evaluation.** URL: **Useit2**. http://www.useit.com/papers/heuristic/heuristic_evaluation.html

Nielsen J. 2000. **Designing Web Usability: The Practice of Simplicity**. New Riders Publishing.

Parbhoo S. 2000. **An Evaluation of Usability Issues for Internet Banking**. Unpublished research report. University of South Africa.

Ravden S J and Johnson G I. 1989. **Evaluating Usability of Human-Compute Interfaces: A Practical Method.** John Wiley and Sons.

Sbic. Standard Bank. 2000. URL: Standard Bank Home page: http://www.sbic.co.za

Schneider G P and Perry J T. 2000. **Electronic Commerce**. Course Technology.

Thimbleby H. 1997. **Gentler: A Tool for Systematic Web Authoring.** International Journal of Human-Computer Studies. 47(1). URL: http://ijhcs.open.ac.uk/thimbleby/thimbleby-01.html

Van Dyk T J. 1996. **A Usability Evaluation of the File and Resource Browser-Manager Components of three Common Desktop Graphical User Interfaces**. Unpublished research report. University of South Africa.

Van Dyk T J. 1999. **Usability and Internet-based Banking**. M.Sc. dissertation. University of South Africa.

Whitaker L A. 1998. **Human Navigation**. In Forsythe C, Grose E and Ratner J. (Ed). 1998. **Human Factors and Web Development**. Lawrence Erlbaum Associates.

Woodson W E, Tillman B and Tillman P. 1992. **Human Factors Design Handbook.** (2nd Ed.). McGraw Hill.