

Assessing the Reliability of Heuristic Evaluation for Website Attractiveness and Usability

Alistair Sutcliffe

Centre for HCI Design, Department of Computation, UMIST, UK
a.g.sutcliffe@co.umist.ac.uk

Abstract

Web interfaces challenge traditional definitions of usability. A three-phase model for website evaluation is proposed, based on initial attractiveness, exploration/navigation and transaction. Usability is re-defined as trade-off between increasing the user's motivation to encourage exploration and purchasing in e-commerce, and the costs of usability errors. Heuristics for assessing the attractiveness of web user interfaces are proposed based on aesthetic design, general arousal created by content, corporate identity and brand, and the perceived utility matched to users' requirements. The heuristics are tested by evaluating three airline websites to demonstrate how different attractiveness and traditional usability trade-offs contribute to overall effectiveness.

1. Introduction

Traditional definitions of usability have emphasised utility and operational ease of use [4]. Usability evaluation methods concentrated on heuristics approaches to assessing design quality (e.g. Nielsen [11]) or observing user errors and inferring their causes in usability defects [19]. However, web interfaces are causing many to rethink such traditional definitions. Website designers are rightly concerned with aesthetic appeal and attracting users. If you cannot attract a user to stay on a website, it doesn't matter how well designed operational usability may be. In acknowledgement of these trends, variations of heuristic evaluation have appeared that assess "minimal and aesthetic" design [12] or affinity, i.e. bringing objects to life through good visual design [2].

However, no revised definition of usability has been proposed that accounts for attractiveness and user satisfaction with an interface, beyond simple questionnaires to capture users' ratings of such

variables. Furthermore, designers have little guidance for creating attractive user interfaces. Some advice can be found in the visual/graphics design community [6], [10]; however, interaction designers give advice in the form of examples and scenarios, leaving the practitioner to abstract generalisable laws and interpret them in a design context. Evaluation of websites has adopted two approaches: first, observation of users' errors when navigating websites; and secondly, expert style heuristic evaluation in which the quality of the interface is judged against a set of criteria- the heuristics. In the latter case the effectiveness of the heuristics has been evaluated for general user interfaces by Nielsen & Molich [13] who reported that 60% of observed user problems could be identified by expert inspection using heuristics. However, validation of heuristics when they are used as quality criteria rather than diagnostic guides is less sure.

This paper is motivated by two problems: first, how traditional definitions of usability need to be re-examined in light of experience of the web applications; and how attractiveness can be operationalised in terms of design guidance. Second how reliable heuristic judgement of website usability and attractiveness is. The paper is structured in three sections. Firstly, a model of usability and attractiveness is proposed for websites and e-commerce in particular. This is followed by investigating how one component of the new usability, attractiveness, can be evaluated. The third section provides a case study that assesses the reliability of heuristics for evaluating the usability and attractiveness of web interfaces. The paper concludes with a brief discussion.

2. A model of effectiveness for Web user interfaces

Usability evaluation has traditionally concentrated on the operational ease of use of an interface [4]. In web applications attention has focused on the ease of navigation, clarity of links, and the ease of information search (e.g. Spool et al, [16]). Heuristic evaluation methods have been extended for web interfaces; however, they still draw attention to design features that improve the predictability and intuitive operation of the interface. While these are important issues, overall effectiveness should be concerned with attractiveness as well. In the initial contract with a website, attractiveness will play a key role in determining dwell time (how long a user looks at the home page) and then increasing the motivation for exploration. Only when the user has decided to enter a site will operational usability and navigation become dominant issues. Attractiveness is a key design concern for browsing users, but it is also important for users who arrive at a site via a search engine. Given a choice of several similar sites, the one which makes a better initial impression is more likely to be explored. In e-commerce, the importance of initial attractiveness is obvious.

Attractiveness may be considered to be the result of matching the user's motivations and requirements with the design features on a website. Hence part of the design problem involves requirements analysis and user modelling, resulting in optional choice of content for a set of users. However, there are also more general issues such as the aesthetic appeal of a design, and eye-catching images that project brand identity and corporate image. Furthermore there are general design effects that capture our attention such as use of animations and sound [18]. Control of attention has a complex cognitive mechanism (see [18] for further details); however the design consequences are that user attention can be manipulated by choosing media to attract attention at the general level; only then does content have an important influence. Attractiveness is influenced not only by attention but also by aesthetic qualities of a design, our motivation, requirements, and excitement invoked by the interface. It is not surprising therefore that the connection to design can be mediated by many variables. The design stages that need guidance start with requirements and content selection, followed by selecting appropriate media to represent the content and design of the form of the media, and

finally the design of representational detail. These design stages are associated with the second set of variables, starting with choice of material to inform or stimulate. Design of form involves creating media for aesthetic purposes as well as selecting appropriate media for the message, and finally design of attention directing effects.

The model that anchors the design issues in a general process of interaction with websites is illustrated in figure 1. Assuming a website of interest to the user has been found, the objective of the first phase is to get the user to stay on the website for sufficient time to explore it further. Attracting the user's attention is vital at this stage. Attention has two components: selection of media and salience effects, and judicious choice of content [17]. For content, visual style and presentation of brand or corporate image are important. Attractiveness can be measured by logging dwell-time on sites, as well as user ratings of sites in questionnaires; however, it is an external manifestation of three complex variables:

- *Arousal*: how exciting/restful a website appears to the user. Arousal is linked to the excitement and interest of the user. Generally, moderate states of arousal are beneficial for problem solving and encouraging exploration [3], [14], both of which are necessary for the next phase of searching and navigation. Images of natural scenes, such as landscapes, tend to be restful and decrease arousal; in contrast, images of technology, unusual objects, sex, injury and violence are all arousing, although in different ways. Clearly sexually explicit material relates to sexual motivation; however, other arousing material has a more general effect of increasing excitement and attentiveness. Dynamic media, especially speech and video, also have a more arousing effect than static media, especially when they engage us in conversation.

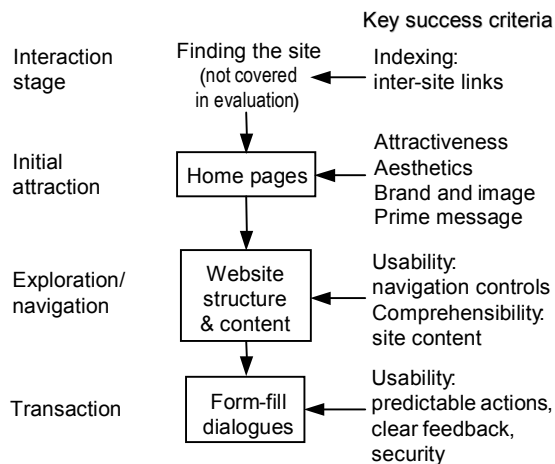


Figure 1. Attractiveness and usability criteria during the stages of interaction with e-commerce websites.

- *Motivation* is reflected in our will to act or hold a belief, and has a complex literature in psychology [8]. Motivation is an under-researched area in HCI that concerns our attitudes and predispositions to act, depending on variables such as need for goods, power, self esteem and less powerful factors of curiosity, learning and altruism. Arousal interacts with motivation by tuning our awareness, e.g. arousal increases curiosity. Measuring motivation is a complex subject because it runs into the problem of individual differences. Assuming that a group level of motivation can be identified, some guess can be made about how the different facets of motivation will be served by content and design features of user interfaces.
- *Perceived utility gain* is influenced by our motivation for goods, services and wealth in a direct manner, and less directly for power, self-esteem and altruism. Utility requires a detailed model of the user to match the user's need with the product/service.

The initial stage of attraction involves gaining the user's attention. Use of appropriate media is important, as are highlighting techniques. For instance, dynamic media (video and speech) are more attention directing than static media (text, still image) [17]. However, the effect can be overdone; for instance, too many animated banners compete with each other and rapidly become annoying, as many of us have experienced with search

engines. Video, audio and change in image by highlighting all focus attention [17]. Once the user's eye has been drawn to the web page, content-based attraction takes over. Projection of brand and organisational identity that promote trust [5], and information that conveys the potential utility of the website to the user will also contribute to holding the user's attention. This implies user modelling to specify their knowledge of brands as well as requirements.

More general attraction can be fostered by aesthetic design and use of media for arousal. Aesthetic attractiveness is a complex variable that is subject to individual differences, as summarised in the saying "beauty lies in the eye of the beholder". Nevertheless, some general principles of aesthetic design can be described and their application should result in a longer dwell-time on the site.

Once the user has been attracted to the website home page and has been persuaded to stay, the next phase begins. In most cases, finding the goods, service or information the user requires necessitates navigation. In exploration and navigation, the conventional quality of usability is paramount. Clear prompts, consistent layout and controls and observable effects all promote ease of use, which can be assessed by standard evaluation methods [9]. However, on websites information plays a key role that goes beyond conventional usability. Early hints on direction to follow towards the search target are important.

Usability problems can terminate user interaction as this stage, so careful design is vital. Critical incidents in which users are confused but can eventually guess what to do may be survivable if their motivation is high; however, errors from which the user cannot recover must be eliminated. Misleading cues for information searching will have a deleterious effect on users' patience. Motivation will have to be high to continue searching after they have followed false cues. User motivation will be subject to the conflicting forces of the promise of perceived utility on the plus side, and the cost of errors on the other. Sites with a close match between their product offering and users' requirements may be able to get away with poor usability, but most sites will not.

The final stage is the transaction when the user purchases the goods/service. Operational usability will be important as well as motivation to counteract any usability difficulties. In information intensive applications, presentation in appropriate media with a well-structured layout will be a key usability requirement [4]. In e-commerce, information

presentation has a key influence on purchasing behaviour; for instance, ranking product attributes (quality before cost) can sway users' choice [7]. Other techniques are use of speech and images of people to engender human response to computers by praising the user, being polite, showing interest and attention via gaze and posture, and use of authority figures or young women to persuade [1], [15].

Design of successful websites therefore has to recruit different design guidelines to fulfill different needs at each stage. Initial attention is replaced by arousal and content related attraction. This raises the user's motivation by the promise of the utility reward to come. The motivational capital has to be maintained during the exploration/navigation stage to counteract any difficulties and disappointments when searches result in dead ends. Once the search target has been reached, user motivation needs to be encouraged and usability errors eliminated to engender successful interaction.

User costs are composed of interaction steps, which in websites tend to be simple, so the number is vital. Long winded multiple step dialogues incur more cost, a lesson not lost on amazon.com who implemented one click shopping. The other element of cost is error, caused by usability problems, misleading cues and failed searches. The phases of interaction can be evaluated by the following measures and techniques:

- *Finding the website*: tests with different search engines using a selection of keywords. % of searches that correctly identify the website, with a relevance ranking $> x\%$.
- *Initial attraction*: dwell-time measured from user interaction logs and de-briefing interviews to investigate what features users noticed and what attracted or repelled them. Free recall memory tests to establish the topics and features users remembered about a site.
- *Exploration and navigation*: usability measures such as errors (% of searches correct) and task completion times. Expert judgement about conformance to design guidelines and heuristics. Cognitive walkthroughs of interaction to identify design flaws [19].
- *Transaction*: usability measures as above, plus de-briefing interviews to discover users' rating of website utility.

In the next section one aspect of the new usability, attractiveness, is investigated in more depth.

3. Evaluating attractiveness

Matching content to users' requirements and motivation is already a complex topic; when judgement on aesthetic design is added, the source of variation from individual differences will become even greater. In spite of these problems, it is better for HCI to attempt to understand the quality of attractiveness rather than leave such issues to designers' craft skills.

Attractiveness can be divided into generic qualities of a user interface such as aesthetic design, use of media to direct attention, and content related issues of linking visual style, brand image and messages to users' knowledge of the organisation and their requirements. Heuristics are proposed for both generic and content related design issues. The following heuristics extend existing advice on website design (e.g. Nielsen's [12] heuristics; IBM's [2] design principles) and can be used either as design advice or evaluation criteria. Proposing principles for aesthetic design is contentious because the graphics/visual design community follows an experiential approach rather than an engineering design philosophy, so articulating design principles is not encouraged. Nevertheless some researchers have partially formalised good design qualities and the following heuristics are based on their recommendations [6], [10].

The generic heuristics for attractiveness and aesthetic design are as follows:

- *Judicious use of colour*: colour use should be balanced and low saturation pastel colours should be used for backgrounds. Designs should not use more than 2-3 fully saturated intense colours.
- *Symmetry and style*: visual layout should be symmetrical, e.g. bilateral, radial organisation that can be folded over to show the symmetrical match. Use of curved shapes conveys an attractive visual style when contrasted with rectangles.
- *Structured and consistent layout*: use of grids to structure image components and portray a consistent order; grids need to be composed of rectangles which do not exceed a 5:3 height to width ratio.
- *Depth of field*: use of layers in an image stimulates interest and can be attractive by promoting a peaceful effect. Use of background image with low saturated colour provides depth for foreground components.
- *Choice of media to attract attention*: video, speech and audio all have an arousing effect and increase attention. Music can attract by setting the appropriate mood for a website.

- *Use of personality in media to attract and persuade:* this principle applies primarily to e-commerce websites when use of human image and speech can help to attract users and persuade them to buy goods by being polite and praising their choices [15].
- *Design of unusual or challenging images* that stimulate the users' imagination and increase attraction: unusual images often disobey normal laws of form and perspective.

Describing rules for each of these qualities is more difficult for some (unusual design) than others (structured layout). When used for design the heuristics are accompanied by examples to amplify their recommendations; in this paper they will be used for evaluation. Evaluators are requested to rate the interface on each heuristic using a 1 to 10 scale. As with general HCI heuristics, a greater number of evaluators will provide a more reliable opinion; however, experts may be unreliable in judging aesthetic qualities. Graphical design experts are prone to disagree about aesthetic qualities more than are ordinary users.

The content related heuristics are:

- *Consistent visual style.* This heuristic is on the borderline between the two sets. Visual style is generic in the sense that a website needs to be consistent in terms of layout and image, but the style also needs to reflect the corporate values. Hence a website targeted at the youth market should use arousing material, whereas a site targeted at older users may use more restful, natural images. For tranquillity, choosing natural world content is advisable; conversely the image of a modern, dynamic organisation is reinforced by technological subject matter (e.g. racing cars, jet aircraft, spacecraft) [15].
- *Visibility of identity and brand.* The effectiveness of this heuristic depends on the strength of the brand image and corporate identity. The design principle just recommends making the identity visible in a consistent manner.
- *Matching arousal to user's mood and motivation.* This heuristic focuses on the match between the user model and website content. Variations to be expected are between age and gender. Ultimately this is a complex topic dealt with in many books on marketing research.
- *Stimulating users' interest by secondary motivation.* Attractiveness can be increased by adding

functionality that is not geared to the site's primary purpose, but may attract for another motivation. Some examples are placing games and simulations on e-commerce sites for users' amusement.

- *Selecting content to suit users' requirements.* This should result from a sound requirements analysis, but poor content display may confound a thorough requirements analysis. Content related to users' requirements should be clearly stated, in unambiguous language, with clear cues on how to find it.

These heuristics need to be combined with existing usability principles to give an overall usability/attractiveness assessment. The generic heuristics apply more strongly to the initial attraction phase of website interaction, while both sets apply to initial attraction, exploration and transaction stages. If the site is rated well on the above heuristics, user motivation will be maintained so some usability errors may be tolerated.

In the next section the heuristics and model of website interaction are used in a case study evaluation of three websites with similar business objectives.

4. Evaluation case study

Three airline websites are assessed: EasyJet (EJ), Virgin Atlantic (VA) and British Airways (BA). Nine undergraduate students (6 male, 3 female) from those who were taking the HCI module assessed the website after a tutorial lecture had introduced the heuristics. The subjects rated each site on a 1-5 scale for each heuristic and were asked to report the rationale for their decision and the ease with which each heuristic could be interpreted. The rating scores were converted into net positive values to reflect the range of the evaluators' assessments. A worked example of this analysis is given in table 1. The frequency of the evaluators' ratings is multiplied by the +2 to -2 scale and the products summed to give a value for the heuristic.

Rating	1	2	3	4	5
Scale	-2	-1	0	+1	+2
Rating Freq/subject	2	0	4	3	0
Product	-4	0	0	3	0
Total Net positive value					-1

Table 1. Worked example of the net positive value for the rating of website persuasiveness by the nine subjects.

All three sites aim to provide information about the airline as well as online sales of flight tickets. The companies concerned have different corporate images, which to a lay observer may be characterised as blue chip reliability and quality (British Airways), modern and exciting (Virgin) and cheap and cheerful (EasyJet). The ratings of each site are given following the model as a cognitive walkthrough with a common scenario of buying a flight ticket. The overall assessment of the sites is given in table 2.

	EJ	VA	BA
Overall attractiveness	-6	7	-2
Persuade to buy	1	0	1
Usability and Navigation	8	4	5
Total	3	11	4

Table 2. Average scores for each site on attractiveness, persuasion and usability.

Overall, Virgin Atlantic received the higher scores, with EasyJet and British Airways being marginally more persuasive, even though two evaluators commented that EasyJet’s “buy now” message was obvious and on the home page. All three sites were favourably rated for usability with EasyJet in the lead, while Virgin Atlantic was the leader in attractiveness.

Assessment on the design quality heuristics (table 3) tells a more complex story. Two used colour for projecting corporate identity (red for Virgin, orange for EasyJet); however, this led to a low rating for Virgin. Both organisations’ colours are part of the corporate image, so judgement on this heuristic indicates a

possible clash of aesthetic appeal and brand projection. Symmetry/style was judged to be below average in all sites; however, several evaluators reported that this heuristic was the most difficult one to interpret.

	EJ	VA	BA
Use of colour	3	-4	6
Symmetry/aesthetic style	-3	-5	-6
Structured layout	3	2	-2
Depth of field	-4	3	-3
Choice of media	-9	3	-11
People and personality	-8	-11	-10
Unusual images	-11	7	-14
Totals	-29	-5	-40

Table 3. Aesthetic design qualities of the websites judged from front pages, rated on 1 to 5 scale, where 5 = excellent.

Virgin and EasyJet were rated more favourably than British Airways on well-structured and consistent pages, which seems to conflict with irregular appearance of the EasyJet’s pages; see figure 2.



Figure 2. EasyJet web front page, which has a somewhat irregular layout but to its credit is good at presenting its basic message: cheap flights and the economic incentive for booking on the web.

EasyJet and British Airways scored poorly on depth of field and choice of media to attract interest; however, Virgin's use of the jet windows metaphor did show some innovation, which is reflected in higher scores. None of the sites scored well on use of personality and people to engage users. Virgin scored well on use of animated media to attract attention, and overall created more exciting content by use of animation and design layout, as illustrated in figure 3.

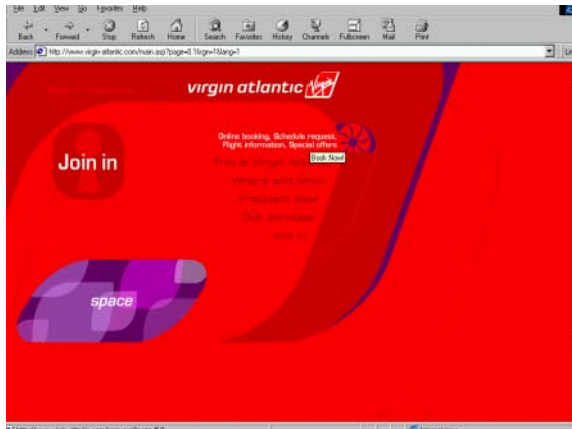


Figure 3. Virgin website, illustrating the contrast in graphic design with figure 2. The jet windows and keyhole metaphors contain animations to attract attention.

The conclusions from the Stage 1 analysis rated Virgin highest in terms of aesthetic design, reflected in the evaluators' comments that its appearance was clearly different from the more traditional block structure layout of the other two websites. However, the aesthetic qualities of all three sites were judged to be below average with British Airways being particularly bad.

The content assessment shows a less clear-cut picture, as summarised in table 4. In this case, judgement is made by browsing throughout the website to assess visual style and brand visibility, while primary, secondary motivation and contents requirements match were judged on the first two to three web pages encountered when following the flight booking scenario.

	EJ	VA	BA
Visual style	11	8	-2
Brand visibility	12	9	6
Mood and first motivation	0	2	0
Secondary motivation	-3	2	-9
Content and requirements	5	2	5
Totals	25	23	0

Table 4. Assessment using the content/attractiveness heuristics.

Both EasyJet and Virgin Atlantic were rated well on a consistent visual style and good brand visibility. In contrast, British Airways was noted to be more discreet about their corporate identity; see figure 4.

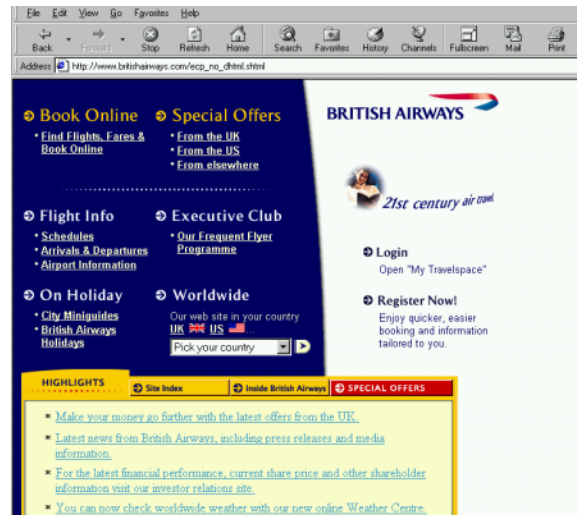


Figure 4. British Airways front page, with a low-key corporate identity (compared with Figure 3), but the structure is well laid out and the content meets with users' requirements for flight browsing/booking.

There was little to choose between the three sites in primary motivation, because all provided services for searching flight availability and booking, although EasyJet did make this functionality easier to access on the home page. Virgin scored slightly higher than British Airways and EasyJet by providing a clearer indication of secondary information on holidays, frequent flyer clubs, car rental, etc. on their front page.

The match of content and requirements for flight information and booking was similar for all three sites. In the totals for content attractiveness British Airways came off worst, whereas the clear brand image and corporate visual style paid off for EasyJet and Virgin. Three evaluators commented that the chromatic identity of Virgin was striking even if they didn't like it (see colour rating in aesthetic heuristics)

The final part of the assessment for Stage 2 (navigation) and Stage 3 (transaction) was judged by browsing extensively through the site and following a flight booking transaction. Heuristic evaluation criteria were used to judge the usability of navigation and transaction support. Space precludes reporting detailed results so a summary of the usability assessment for each site is given in table 5.

In the latter phases of interaction, EasyJet does not compare so favourably. The navigation controls and support are weak because no site map is provided, navigation bars have overlapping content, the side bar animations contain link cues but change, making navigation confusing; no back to top commands are given on the bottom of long pages. Transaction controls also suffer from similar defects, e.g. no exit, long and cumbersome scroll boxes. Virgin Atlantic and British Airways do better by supplying the essentials of good navigation (site maps, consistent navigation menus, back to top and exit/home), although navigation controls are more consistent and visible in the Virgin site. Both Virgin and British Airways use a clear task step metaphor (1-2-3-4-5) to guide the user through flight reservation, booking and payment. EasyJet adopts the same metaphor, but implements it less clearly.

Usability criterion	EJ	VA	BA
Navigation commands	1	3	2
Navigation support	2	3	2
Transaction prompts	3	4	5
Form fill layout	2	4	4
Transaction controls	1	4	5
Totals	9	18	18

Table 5. Usability assessment of the navigation and the transaction phase.

	EJ	VA	BA
Visibility of system status	2	1	-1
System RW match	6	6	6
Control and freedom	-1	0	3
Consistency and standards	15	6	6
Error prevention	2	0	-4
Recognition opposed to recall	8	-4	3
Flexibility and efficiency	1	-3	0
Aesthetic/minimalist design	3	6	4
Error handling	7	2	0
Help and documentation	-7	-8	4
Total	36	6	21

Table 6. Evaluators' rating using Nielsen's heuristics.

The evaluators' scores with Nielsen's heuristics (see table 6) showed less variation and more neutral assessments apart from the high consistency score given to EasyJet. Easyjet and British Airways come out more favourably on these heuristics; however, the high EasyJet rating on Consistency does not agree well with the Structured Layout heuristic in the aesthetic set, demonstrating some inconsistency in evaluator judgment.

The summary picture is shown in table 7. Overall, Virgin appears to be the best placed site with a first rank in the overall attractiveness, aesthetic heuristics and joint first on Usability/navigation. However, EasyJet ranked first on Nielsens' heuristics and Content related attractiveness.

	EJ	VA	BA
Overall appeal	3	11*	4
Aesthetic attraction	-29	-5 *	-40
Content attraction	25 *	23	0
Usability/Navigation	9	18 *	18 *
Nielsen heuristic	36 *	6	21

Table 7. Summary totals of all the evaluation heuristics and phases. * denotes first ranking site.

The evaluation shows clear strengths and weaknesses of each site at each stage. Overall, Virgin wins on attractiveness and aesthetic design as well as having a well-designed transaction/navigation interface. EasyJet is strong on initial content-driven attraction but suffers from basic usability defects in the exploration and transaction phases, e.g. no escape route. Motivation created during the initial phase may encourage users to tolerate usability error costs in later phases; however, overall, the Virgin site is likely to be the more effective because it combines reasonable motivation with aesthetic attractiveness and sound usability engineering.

Evaluators' assessment

Overall the attractiveness heuristics were positively rated and judged to be easy to use (npv = 5 for design quality and 11 for the content heuristics) and fared slightly better than the Nielsen heuristics which were also used (npv = 2). Only two heuristics attracted adverse comments: judgement of symmetry, which four out of nine evaluators found difficult to interpret; and depth of field (two evaluators). However, three or more evaluators commented that they felt their judgement was very subjective, in particular when judging motivation and visual content assessment, and usual/exciting content in the design heuristics. Two Nielsen heuristics (system match with real world and aesthetic/minimal design) were considered to be ambiguous by three of the evaluators. Flexibility and efficiency also required interpretation. The evaluators judged efficiency on response time and the ability to operate the interface quickly and with few errors, but the flexibility component was found to be difficult to judge. When inter-observer agreement was calculated by Kendall's coefficient of concordance only three of

the heuristics gave significant agreement at the $p = 0.05$ level: Depth of field, Use of Personality, and Unusual images. The lack of agreement is disappointing; however, Nielsen's heuristics fared no better with only two (Freedom and control, and Error prevention) passing at the $p = 0.05$ level, so this may be a common problem with heuristic style evaluations. Since the evaluators did agree on the overall rating of attractiveness ($p = 0.05$) and three of the general attractiveness heuristics, with another two just missing significance, there seems a reasonable prospect for improving this evaluation approach.

5. Discussion

This study is a preliminary investigation into the role of aesthetic and other factors in re-defining website usability. However, it does demonstrate how explicit consideration of aesthetics and motivation factors can alter judgement of website quality. Heuristic evaluation [11] placed EasyJet first. In contrast, explicit consideration of attractiveness suggests the Virgin Atlantic design has many advantages. The influence of aesthetics in the overall effectiveness requires further research to investigate correlations between performance measures of visit frequency and durations, users' purchases and subjective assessment of websites. The heuristics separate content from aesthetic based attraction so differences in their influence can be assessed, i.e. aesthetics may play a more important role for initial visits but content issues may be dominant for repeat visits. Attractiveness and aesthetic design are key factors in persuasive computing [1]; however, further research is required on articulating the design properties of attractiveness and aesthetics across different designs and user groups. Furthermore, we need to understand the interaction between user motivation and usability costs, the factors that augment motivation and the consequences for user interface effectiveness from motivation/cost trade-offs.

The new heuristics proposed in this study extend the range of evaluation from usability to more motivational aspects of design. The next step is to test and calibrate the judgements by further samples of evaluators: HCI experts, graphic designers, software engineers and users to see if each group produces different ratings. The second validation is to check expert assessment against end-user judgement and purchasing behaviour of the test websites. However, this data may be difficult to interpret because other factors are involved. For

instance, EasyJet advertises its Internet booking service extensively. However, first we need to reflect on the reliability of subjective judgement found in this study. The extended heuristics addressed issues of aesthetic judgement where individual difference in subjective assessment might be expected; what is worrying is that Nielsen's heuristics were not more reliable. While some variance between evaluators is to be expected, the comparative judgment of websites should be consistent if heuristics are to be used as an evaluative instrument. The subject population of HCI students were familiar with Nielsen's heuristics even though the extended set were new. The tentative conclusion is that heuristics should not be used for subjective rating style judgement, although this does not invalidate their use as a diagnostic aid to usability evaluation which was Nielsen's [11] original intention. The next step is to adopt grounded theory techniques from social science to develop more reliable questionnaire inventories [20] for assessing subjective qualities in websites, and then see how these correlate with observed user behaviour.

Acknowledgements

The author would like to thank Andreas Gregoriades for his help in preparing this paper.

References

- [1] Fogg, B. J. Persuasive computer: perspectives and research directions. *Proceedings: Human Factors in Computing Systems: CHI '98, Los Angeles CA, 18-23 April 1998*. ACM Press, New York, 1998, pp. 225-232.
- [2] IBM. *Ease of use: design principles*. http://www.ibm.com/ibm/easy/eou_ext.nsf/Publish/6. Accessed 20 November 2000.
- [3] Isen, A. M. Positive affect and decision making. In M. Lewis & J. Haviland (Eds.), *Handbook of emotions*. Guilford Press, New York, 2000.
- [4] ISO. *ISO 9241: Ergonomic requirements for office systems with visual display terminals (VDTs)*. International Standards Organisation, 1997.
- [5] Kollock, P. The production of trust in online markets. In E. J. Lawler, M. Macy, S. Thyne & H. A. Walker (Eds.), *Advances in group processes, vol. 16*. JAI Press, Greenwich CT, 1999.
- [6] Kristof, R. & Satran, A. *Interactivity by design: creating and communicating with new media*. Adobe Press, Mountain View CA, 1995.
- [7] Lohse, G. L. Usability and profits in the digital economy. In S. McDonald, Y. Waern & G. Cockton, (Eds). *People and Computers XIV: Usability or Else; Proceedings: BCS-HCI Conference*. Springer, Berlin, 2000, pp. 3-16.
- [8] Maslow, A. H., Frager, R., McReynolds, C., Cox, R. & Fadiman, J. *Motivation and personality* (3rd ed.). Addison Wesley-Longman, New York, 1987.
- [9] Monk, A. G. & Wright, P. *Improving your human-computer interface: a practical technique*. Prentice Hall, London, 1993.
- [10] Mullet, K. & Sano, D. *Designing visual interfaces: communication oriented techniques*. unSoft Press, Englewood Cliffs NJ, 1995.
- [11] Nielsen, J. *Usability engineering*. Academic Press, 1993.
- [12] Nielsen, J. *Designing Web usability: the practice of simplicity*. New Riders Publishing, 1999.
- [13] Nielsen, J. & Molich, R. Heuristic evaluation of user interfaces. *SIGCHI Bulletin* (April: special issue), 1990, 6.
- [14] Picard, R. W. *Affective computing*. MIT Press, Cambridge MA, 1997.
- [15] Reeves, B. & Nass, C. *The media equation: how people treat computers, television and new media like real people and places*. CLSI/Cambridge University Press, Stanford CA/Cambridge, 1996.
- [16] Spool, J. M., Scanlon, T., Snyder, C., Schroeder, W. & DeAngelo, T. *Web site usability: a designer's guide*. Morgan Kaufmann, San Francisco, 1999.
- [17] Sutcliffe, A. G. User-centred design for multimedia applications. *Proceedings ICMCS 99, Vol. 1: IEEE Conference on Multimedia Computing and Systems, Florence 7-11 June*, IEEE Computer Society Press, Los Alamitos CA, 1999, pp. 116-123.
- [18] Sutcliffe, A. G. *Multimedia and virtual reality: designing multisensory user interfaces*. Lawrence Erlbaum Associates, Mahwah NJ, 2001?
- [19] Sutcliffe, A. G., Ryan, M., Doubleday, A. & Springett, M. V. Model mismatch analysis: towards a deeper explanation of users' usability problems. *Behaviour and Information Technology*, 19(1), 2000, pp.43-55.
- [20] Watson, A. & Sasse, M. A. Evaluating audio and video quality in low cost multimedia conferencing systems. *Interacting with Computers*, 8(3), 1998, 255-275.