

International Encyclopedia of Rehabilitation

Copyright © 2010 by the Center for International Rehabilitation Research Information and Exchange (CIRRIE).

All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system without the prior written permission of the publisher, except as permitted under the United States Copyright Act of 1976.

Center for International Rehabilitation Research Information and Exchange (CIRRIE)
515 Kimball Tower
University at Buffalo, The State University of New York
Buffalo, NY 14214
E-mail: ub-cirrie@buffalo.edu
Web: <http://cirrie.buffalo.edu>

This publication of the Center for International Rehabilitation Research Information and Exchange is supported by funds received from the National Institute on Disability and Rehabilitation Research of the U.S. Department of Education under grant number H133A050008. The opinions contained in this publication are those of the authors and do not necessarily reflect those of CIRRIE or the Department of Education.

The Design and Evaluation of Assistive Technology Products and Devices Part 2: Evaluation of Assistive Products

Marion A. Hersh

**Dept. of Electronics and Electrical Engineering,
University of Glasgow, Glasgow, G12 8LT, Scotland.
Tel: +44 141 330 4906. Email: m.hersh@elec.gla.ac.uk**

This article is the second of three parts and discusses the design of assistive products. Part 1 considers design and Part 3 the outcomes of assistive product use.

Introduction: Definition and Aims of Evaluation

There is a relatively limited body of literature on the evaluation of assistive technology and this has tended to focus on outcome measures related to the use of assistive products, in particular measures associated with quality of life or functioning. However, even in the area of outcome measures, there is only a limited body of literature, for a number of reasons, including the belief that the benefits are obvious, a focus on the performance of the product, rather than its interaction with the user and relatively little demand from stakeholders (Fuhrer 2001). While the outcomes of the use of assistive products are important, outcome measurement is only one aspect of evaluation. In general, evaluation should take place throughout the product lifecycle. At the same time there is a need for a parallel evaluation process of individual and groups of users' experiences with and attitudes to the product and the two processes should feed into and support each other with information. However, there will generally be some degree of correlation between the resources, including time, put into evaluation and the importance, complexity and cost of the product.

The evaluation of assistive products has generally focused on the outcomes of use by a particular user or group of users, whereas the focus of evaluation for other products has been the likely market success of new products. However, despite these differences, many of the reasons for evaluating assistive product (outcomes) are similar to those for other (consumer) products, including supporting marketing decisions, increasing accountability, adding to knowledge about assistive products and developing new products, which requires accurate data on which devices work for which (groups of) people under which circumstances (Fuhrer 2001). Other reasons, which again hold for both assistive and other products, include ensuring that products comply with legislation, including any legal requirements for evaluations, have satisfactorily resolved any health and safety issues, have the claimed functionality and perform satisfactorily. An appropriate evaluation (by a competent body) and the completion of associated documentation will be required for products to be kite marked as meeting appropriate standards. In some cases, the prerequisite conditions for eligibility for funding for product development may have conditions include the requirement to carry out an evaluation. For instance carrying out efficacy trials to determine whether new devices achieve the expected benefits for users is a condition of federal funding for the development of assistive technology in the USA (Fuhrer 2001).

The extent and type of evaluation that is appropriate for a particular assistive product depends on a number of product factors. In particular, complicated electronic assistive products, such as an eye-gaze computer pointing system for people who have little or no limb movement

require a more detailed evaluation than simple manual assistive products, such as tactile playing cards for blind people. However, the successful performance of assistive products such as gardening tools with long handles or cutlery and kitchen utensils with easy to hold grips depends on good ergonomic design. The evaluation of safety features of the design is particularly important for a number of assistive products, such as wheelchairs.

Part 1 of the paper (Hersh 2010a) discussed the design of assistive products. This part will consider the evaluation of assistive products and Part 3 (Hersh 2010b) the evaluation of the outcomes of the use of assistive products by disabled and elderly people. The paper is organised as follows. Section 2 considers the classification of approaches to evaluation. The stage-gate approach to evaluation, stages in the evaluation process and discussion of evaluation criteria are discussed in Section 3. Sections 4 and 5 consider the evaluation of usability and accessibility, and functional evaluation respectively. Section 6 discusses production simulation, the evaluation of emotional reactions to product and Section 7 the use of decision support tools to support evaluation. Conclusions are presented in Section 8.

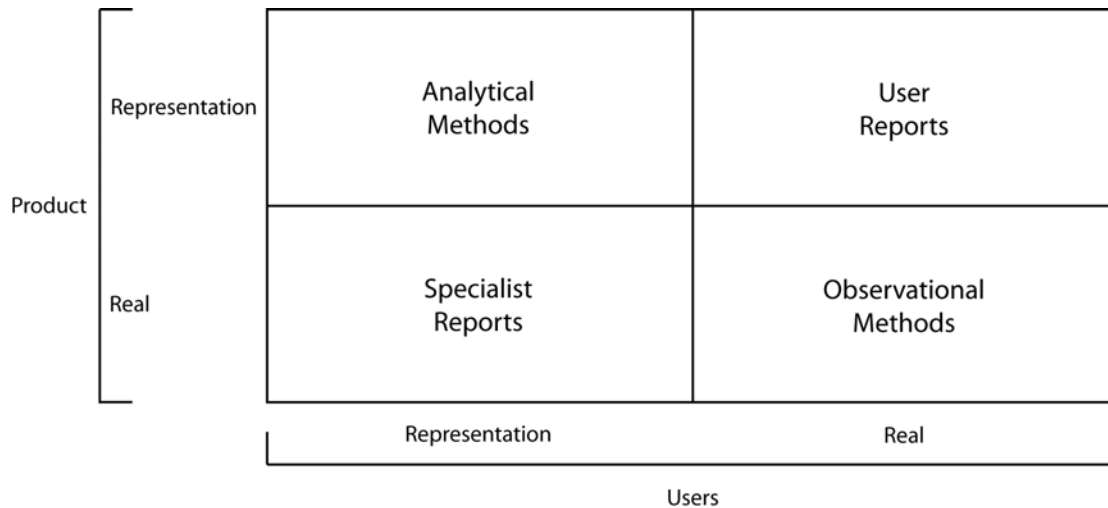
Classification of Approaches to Evaluation

There are a number of different ways of classifying evaluation and evaluation methods. For instance, evaluation can be purely evaluative or diagnostic. In the first case, the aim is solely to investigate performance and conformity with specifications, leading to a presentation evaluation statement. In the second case, evaluation additionally considers the causes of this performance and, in particular, will try to identify the reasons for any unsatisfactory performance and support the development of measures to correct problems, leading to a diagnostic evaluation statement (Whitefield et al. 1991).

Generalising a categorisation of evaluation methods for human computer interaction (Whitefield et al. 1991) to other classes of assistive products gives four classes of methods, as illustrated in Figure 1. One division is based on whether the evaluation involves actual users or user representations, including human factors specialists or other professionals. The other division is based on whether the real product, including prototypes and simulations of it, or representations such as models or user memories of the product are used. This gives the following four categories:

- Analytical methods: specialists manipulate system models to predict performance.
- User reports: actual users work with representations or their memories of particular products. This includes, survey methods, such as interviews and questionnaires.
- Specialist reports: human factors specialists or other professionals assess real products, prototypes or simulations.
- Observational methods: real users assess real products, prototypes or simulations.

Figure 1: Classes of evaluation methods



There have been a number of studies of the factors which are most important in product success and/or product screening decisions, as well as the factors which influence decisions on whether or not to discontinue a product. The factors that affect decisions on discontinuing products can also be considered as factors that influence decisions not to proceed with a new product and those that impact on decisions not to proceed with a new product can also be seen as factors that influence decisions not to discontinue an existing product. Therefore all these factors should all be included in product evaluation at appropriate life cycle stages. They include (Avlonitis 1993; de Brentani 1986; Schmidt et al. 2009; Song and Parry 1996):

- Product advantage and expected market reaction.
- Expected financial potential, including market share, sales and market growth and expected profitability.
- The implications for the firm’s financial, physical and human resources.
- Fit to current business, organisation, technological, marketing and managerial skills.
- The potential or availability of new products and/or alternative opportunities, as well as the product range policy.

The initial development and evaluation stage after the identification of opportunities or design concepts and including concept testing has been found to be particularly important in determining the subsequent market performance and separating successful and unsuccessful products (Schmidt et al. 2009). This is at least partly a consequence of the importance of getting the design right from the start and the additional costs and reduced effectiveness of changes made at a later stage. For instance, 80% of product costs are defined in the design phase (Jagou 1993). In addition, the increasing investment in and commitment to a product as it proceeds (Schmidt and Calantone 2002) makes significant modifications or termination more difficult (Schmidt et al. 2009). Therefore particular attention should be given to evaluation at this stage to limit the risks associated with product development and inappropriate decisions.

The Stage-Gate Approach to Evaluation

Although originally aimed at the introduction of new products (Cooper 1990; Hart et al. 2003; Tzokas et al. 2004), the principle of the stage-gate approach can be extended to evaluation at all stages of the product life cycle. The approach is process-based and divided

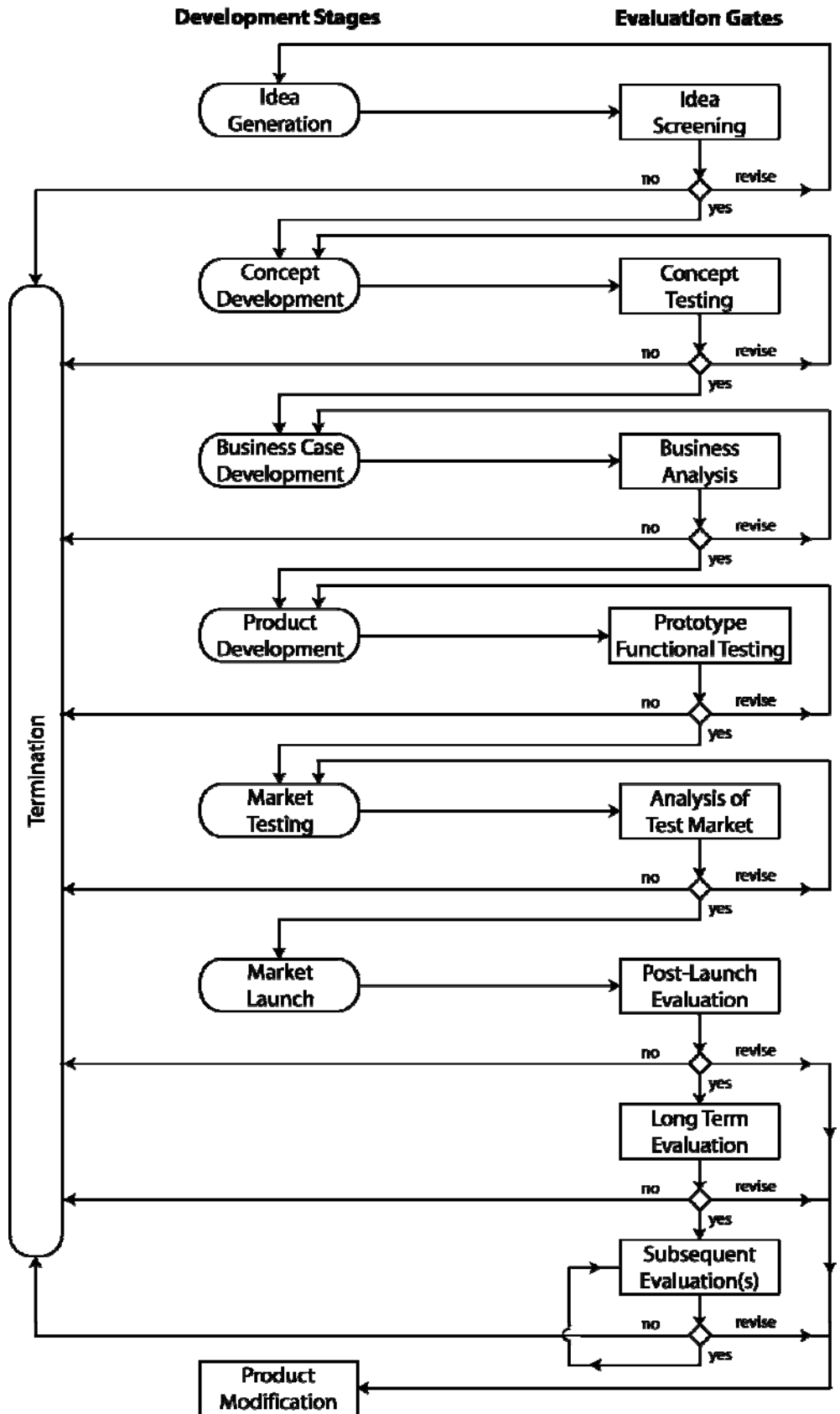
into a number of different stages, with the activities associated with each stage often taking place in parallel. Moving between stages is controlled by gates or evaluative stages. Each gate has deliverables or inputs from the previous stage, a set of criteria to be met before the project can move to the next stage and outputs, generally decisions on whether to proceed, terminate, hold or repeat the project design process and, in the case of a proceed decision, an action plan for the next stage. At each evaluation stage 'gatekeepers' review the quality of the deliverables from the previous stage, make decisions on whether or not the project passes the gate, approve the action plan for the next stage and allocate resources. It has been suggested that the criteria to be used at the different gates should include technical, financial and market based criteria (Hart et al. 2003).

The evaluators could act as 'gatekeepers' or decision makers on subsequent product progress, revision or discontinuing or producing an evaluation report for decision by management. There are advantages in including people with decision making authority amongst the evaluators, with regards to improving efficiency and ensuring that the evaluation is fully understood. There is some disagreement on whether or not senior managers should act as gatekeepers. This has the advantage of ensuring that the gatekeepers have decision making authority (Cooper 1990), but their involvement could give undue weight to a particular point of view and distort decision making (Hart et al. 2003)

The details of the aims and criteria, as well as the supporting structures and the degree of formality with which the evaluation is carried out will depend on the particular organisation and product. The number of stages required will depend on a number of factors, including the type of product and the organisation developing it. Examination of the literature, for instance (Cooper 1990; Ozer 1999; Tzokas et al. 2004) and adding 'subsequent evaluation' stage(s) gives the following evaluation stages, as shown in Figure 2:

1. Idea screening: evaluation of new product ideas.
2. Concept testing: evaluation of end-user reactions to a new product idea, identification of important features, evaluation of market potential and size.
3. Business analysis: evaluation of the technical feasibility of product, market potential and financial contribution to the firm.
4. Prototype functional testing: evaluation by users or experts to determine whether the prototype meets technical, functional and manufacturing requirements.
5. Analysis of test market: decision on the market potential of the prototype.
6. Post-launch evaluation: evaluation of the market success of the new product after 25% of the estimated market lifetime.
7. Long term evaluation: evaluation of the product market success after 75% of the estimated market lifetime.
8. Subsequent evaluation(s): subsequent evaluation(s) of product performance if it is continued beyond the expected market lifetime.

Figure 2: Development and evaluation stages over the product life cycle (after Tzokas et al. 2004)

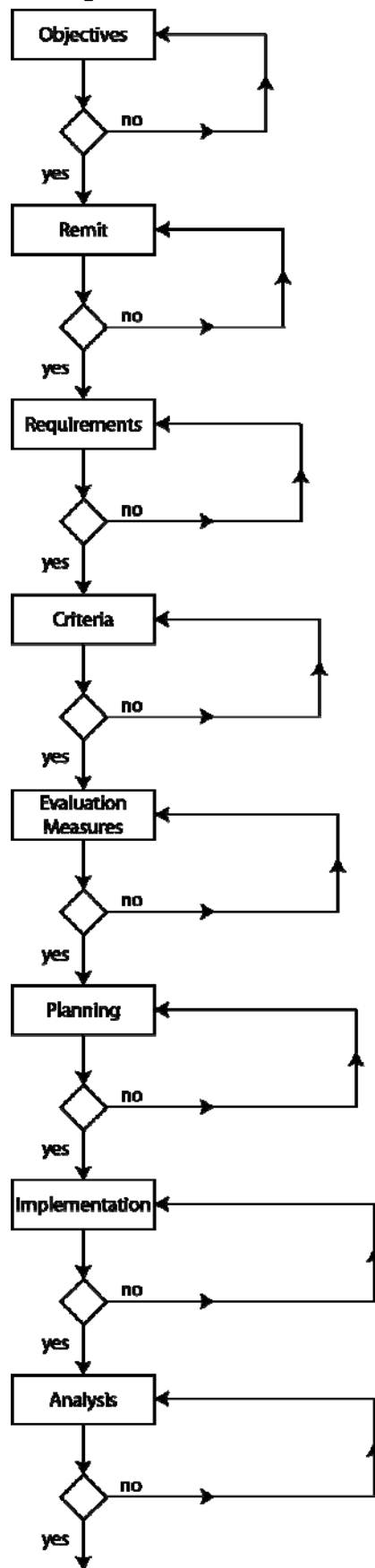


Stages in Product Evaluation

The first part of this Section suggested that evaluation should take place at a number of different evaluation stages throughout the product life cycle. However, there is still the issue of how evaluation should be carried out at each of these stages. There are a number of different approaches, which are, however, all based on dividing the process of evaluation into a number of different steps or tasks to be carried out, though there are differences in both the number and designations of these steps. These approaches include the STACE framework for the evaluation of commercial off-the-shelf software (Kunda and Brooks, 1999), the PECA process (Cornella-Dorda et al. 2004) which is derived in part from the international standard ISO 14598 Information Technology – Software Product Evaluation and a process for usability evaluation (Kwakh and Han, 2002). Examination of common and unique steps in these approaches can be used to derive the steps in product evaluation listed below and illustrated in Figure 3. It will often be appropriate to carry out the process iteratively with backtracking rather than purely sequentially (without backtracking) to take account of the need for additional criteria and/or data or unexpected discoveries. Although both the STACE and PECA processes have been derived from an industrial perspective and the development of a product with at least a small run, the evaluation steps listed below are also relevant to the one-off products and non-industrial approaches to product design and development discussed in (Hersh 2010a).

1. Objectives: Definition of the objectives.
2. Remit: Determining the remit of the evaluation, including the particular features to be evaluated, the depth and rigour of the evaluation, any constraints on the evaluation, including from prior decisions, the team members and their roles and obtaining a statement of commitment from evaluators and management.
3. Requirements: Determining end-user, other stakeholder and system requirements, including by consultation with end-users and other stakeholders, examination of system documents, domain knowledge and market studies.
4. Criteria: Using the requirements obtained in step 3 and standard checklists to produce criteria to be used in evaluation. This may involve a hierarchical decomposition to produce criteria which can either be measured or investigated qualitatively.
5. Evaluation measures: Determining quantitative and qualitative measures for evaluating the criteria and the evaluation techniques that are going to be used to do this.
6. Planning: Planning the evaluation, including determining alternatives, setting up the evaluation scenario or other experiments, as appropriate, and preparing appropriate people, materials and apparatus.
7. Implementation: Carrying out the evaluation, including by collecting data, the investigation of scenarios, end-user, other stakeholder and/or expert evaluation of product mock-ups or prototypes and/or evaluating the alternatives, a particular product or design against the criteria.
8. Analysis: Analysing, interpreting, reporting and reflecting on the results.

Figure 3: Product evaluation steps



Evaluation Criteria

Although the titles of the evaluation stage gates discussed in Section 3 and illustrated in figure 2 indicate their (probable) main focus of evaluation, the evaluation carried out at each stage will generally be broader than this particular focus and in some cases the stage name will not indicate the main focus. In addition, it will be necessary to determine appropriate evaluation criteria. These criteria should generally include both qualitative and quantitative criteria. However, all criteria should be measurable in some way, to enable determination of whether and to what extent they are met, even if the process of measurement is subjective, for instance asking users their opinions. Successful evaluation should take account of the context, uncertainty and facts rather than the opinions of the designers (Comella-Dorda et al. 2004), though it should also consider the subjective views of stakeholders. In the case of assistive technology stakeholders will include all categories of users (disabled and elderly people, personal assistants and possibly family and friends), various professionals, such as occupational therapists and teachers, working with the disabled or elderly person (some of whom may be users of the technology) and funders.

Sources of evaluation criteria include user and other stakeholder requirements and expectations, constraints of different types, checklists of product features and previous evaluations of other systems. Consistent evaluation environments, accurate test equipment and repetition of tests are required to avoid bias or other errors in quantitative data and clear explanations of the meaning of the different ratings to avoid errors in qualitative data. When some criteria are more important than others, weights can be used to indicate the relative priority or importance of the different criteria. For instance the Analytical Hierarchy Process (AHP) (Saaty 1990, 2008) and the P/G% method (Nagel 1988) offer a structured approach to doing this. It can also be useful to divide the criteria into constraints and desirable criteria, where constraints must be satisfied and the desirable criteria are used to evaluate the set of designs or products which meet the constraints and, if required, to choose between them.

Data collection generally results in unorganised raw data which needs to be analysed. Qualitative data which can be rank ordered can generally be converted to a numerical scale. This data can then be integrated with other data and different types of data processed to facilitate comparison using normalisation, of which there are four main types. Their advantages and disadvantages are discussed in (Hersh 1999).

Weights can be used to prioritise criteria and/or to consolidate the data into either a small number of categories or a single category to facilitate analysis. Reducing the number of indicators in this way can facilitate judgements of the relative performance of the different products or designs (Fu and Cebon 2003). However, it does this at the cost of reducing the amount of available information and the use of weights may not be sufficient to allow full consideration of the relative importance of the different criteria. Sensitivity analysis can be used to investigate the effect of changes in assumptions, variable values or weights on the results of the evaluation. In particular this can show whether the evaluation performance is robust to changes in assumptions, values and weights or whether quite small changes can lead to significant differences in the results (Chineck 2007).

Consideration of the literature, for instance (Comella-Dorda 2004; Kunda and Brooks 1999; Raviart et al. 1998) and drawing on the discussion of design issues in Part 1 (Hersh 2010a) of the paper can be used to draw up a list of factors to be assessed to form the basis of criteria. However, it should be noted that in general only a subset of these factors will be evaluated at each stage.

The factors include the following:

Compliance and good practice

- Health and safety
- Compliance with legislation, regulations and standards
- Ethical issues

Technical issues

- Technical performance
- Range and type of functionality
- Efficiency and efficacy
- Data security
- The standards met by interface components
- Reliability, robustness and dependability.

End-user issues

- Usability, accessibility and acceptability for the target user groups, all potential users or the whole population
- User performance and experiences with the product
- End-user experience with similar products and technologies
- Ergonomics
- User satisfaction.

Resource and financial issues

- Direct, indirect and hidden costs
- Financial performance factors
- The costs of training and support.
- The reliability, capability and performance of the assistive product vendor
- The feasibility and costs of upgrading
- Resource utilisation.

Evaluation of Usability and Accessibility

Accessibility is often considered to be an umbrella term for the parameters that affect human functioning in an environment, whereas usability has been defined in the International Organization for Standardization standard ISO 92 4-11 as the extent to which a product can be used by specified users to achieve specified goals in a specified context of use with effectiveness, efficiency and satisfaction'. Both usability and accessibility are discussed in more detail in Part 1 (Hersh 2010a).

Evaluation of Usability

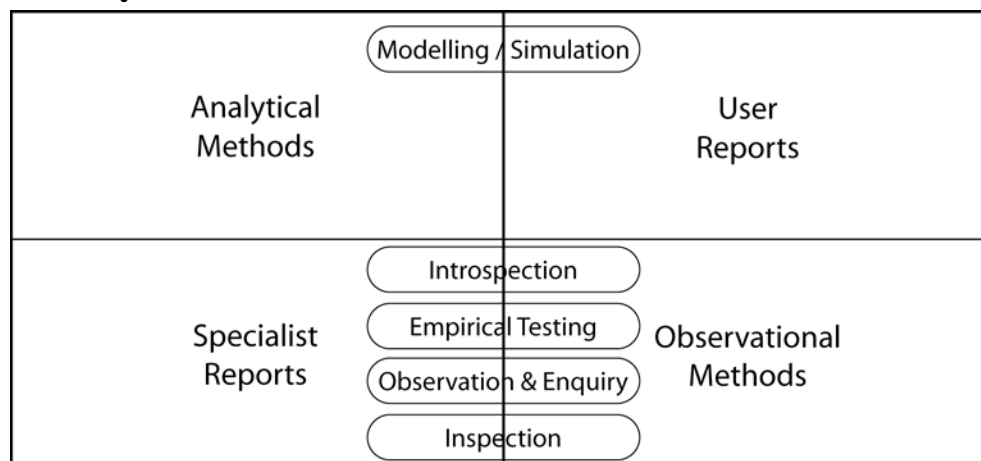
Evaluation of usability should take account of the user, the task and the context or environment as well as the product. Studies of usability generally focus on the system interface, including user dialogue components (Kwakh and Han 2002), since users generally interact with the product via the user interface. However, usability is relevant to all stages of product use that a user might be involved in, including installation and maintenance (Nielsen 1993) and therefore usability evaluation should cover these other stages in addition to product

use. Evaluation of usability for novice, expert and occasional users has been proposed (Nielsen). However, it would be useful to further distinguish between expert and intermediate users, so that expert users can be defined as those who have developed some degree of real expertise, whereas intermediate users have some experience and skills, but do not have the extent of skills, experience and understanding of the domain to be considered ‘expert’. A study of the literature found 36 different usability measures and 35 different usability evaluation techniques which could be used in practice (Kwakh and Han 2002). The techniques were classified into the following five groups:

- Observation and enquiry, including time line analysis and focus groups;
- Empirical testing, including benchmark tests and user trials;
- Introspection, including cognitive walkthrough and thinking aloud;
- Inspection, including feature evaluation and heuristic evaluation;
- Modelling/simulation, including model based evaluation and computer simulation.

As discussed in Section 2, the different methods can also be classified according to whether they involve actual users or other people and whether they involve tests of the actual product or representations of it. The relationship between two different types of classification of usability evaluation methods is illustrated in Figure 4.

Figure 4: The relationship between different types of classification of methods for usability evaluation



Different types of methods may also be appropriate at different points in the life cycle. In particular, the use of product representations allows usability to be tested early in the design lifecycle before the user interface (or other features) have been implemented. It may be easier for experts than end-users to evaluate such representations and determine whether or not there are likely to be usability problems. One approach that can be used from an early stage in product design, involves a small number of evaluators independently evaluating the conformity of the product interface(s) with recognised usability principles or heuristics, which describe common properties of usability interfaces. Three to five evaluators are required to identify all the usability problems, which should be listed separately, even if they affect the same dialogue element (such as the boxes with options for opening or saving files in word processing programmes) to avoid repeating problems when a dialogue element is modified or replaced. The severity of these problems can then be evaluated by providing the evaluators with descriptions of all the usability problems and the ratings averaged (Nielsen).

The involvement of users in usability evaluation from an early a stage as possible is likely to be particularly important for assistive products, since experts can experience difficulties in putting themselves in the users' place (Sweeney et al. 1993), unless the experts are disabled or elderly people with similar experiences to the target users. User evaluation of usability generally involves (Sweeney et al. 1993) users completing bench-marked tasks and measurements of performance times, accuracy and error levels and/or types being compared to target performance criteria. Usability testing by end-users raises the issue of how best to obtain end-user feedback without distorting the results, for instance by users becoming self-conscious or having part of their attention diverted by the need to focus on the comments they are recording.

Both direct observation and the examination of video recordings may appear intrusive and/or make the end-user self-conscious. Recordings have the advantage of being less likely to miss relevant user behaviour, providing a record which can be replayed for fuller analysis and allowing analysis by several evaluators. In addition, an edited version can be used to demonstrate problems to the design team. However, analysis of the recordings is likely to be time-consuming. Recording users' think-aloud comments while they are using the product or system can present useful data on their cognitive processes and hence on both the associated cognitive load and any particular difficulties users encountered. However, this is intrusive and may distract their attention from the product, thereby distorting performance. Obtaining subsequent comments from users on their cognitive processes and opinions either through interviews and questionnaires or user diaries avoids the problem of user being distracted, but may lead to distortion of the results due to the forgetting and distortion of memories that occurs even a very short time after the event.

Evaluation of Accessibility

Most of the work on evaluating accessibility has been domain specific with a particular focus on web accessibility. Generic evaluation methods which are applicable across different domains do not seem to have been developed. Web accessibility will now be discussed briefly as the area where evaluation procedures are the best developed. In addition, it is both directly and indirectly relevant, since many assistive products have websites which need to be accessible and it is possible to draw some general conclusions from the evaluation of web accessibility which are relevant to other types of accessibility.

Evaluation of web accessibility is generally based on guidelines, the most comprehensive of which are probably due to the World Wide Web Consortium (W3C) Web Accessibility Initiative Web Content Accessibility Guidelines (WCAG), with version 2 the current version (WCAG 2.0 [updated 2008]). WCAG 2 is organised into four principles. Each principle has a list of guidelines that address the principle and a number of success criteria grouped into three levels of conformance, representing increasing levels of accessibility for the associated guidelines (Hersh and Johnson, 2008). Although WCAG 2 evaluation can be used to determine accessibility at levels one, two or three, there is a tendency to consider web sites either accessible or inaccessible.

It has been suggested that a good web site accessibility metric should satisfy the following conditions (Parmanto and Zeng, 2005):

- Provision of a quantitative score that provides a continuous range of values from completely inaccessible to totally accessible.

- Good discriminating power to enable investigation of changes in accessibility over time or differences in accessibility between different websites.
- Consideration of the size and complexity of web sites to allow fair comparisons of the accessibility of different types of sites.
- Support for aggregation and second order statistics, such as standard deviation, to allow large scale web accessibility studies.

Automated indicators have the advantage of speed and objectivity and enable a large number of web sites to be evaluated quickly. However, using such indicators is not the same as user evaluation (Parmanto and Zeng, 2005) and some groups of disabled users may experience accessibility barriers on a site that has received the top or a high accessibility rating. In addition, the guidelines do not necessarily cover all the criteria of relevance for all users. For instance, the guidelines do not say anything about the avoidance of scrolling text, which can be an accessibility barrier to some people with autistic spectrum conditions. To give a fuller picture of the usability barriers that could be encountered as well as to support a more detailed comparison of different web sites, there is a need for both metrics which are applicable to different aspects of accessibility as well as metrics which treat accessibility as a whole. It is also useful to investigate the accessibility barriers likely to be encountered by particular groups of end-users.

Techniques for evaluating the accessibility of assistive products are in general less well-developed than those for evaluating the accessibility of web sites. Thus, there is a need for the development of automated tests which draw on existing guidelines, regulations and checklists of the prerequisite conditions to be met for user accessibility. In addition, it should be recognised that as far as possible automated tests should complement rather than be a substitute for tests by actual end-users, giving a need for the development of methodologies for end-user evaluation of accessibility.

Financial or Economic Evaluation

There are technical differences between the terms financial and economic evaluation, which relate to the items which should and should not be included. While the full details will not be considered here, it will be noted that economic evaluation may include items which are not direct expenditure or income in financial terms, such as in-kind services. The term economic evaluation will generally be used in the remainder of this section.

Two different types of economic evaluation are relevant to assistive products:

- Economic evaluation from the perspective of firms or other organisations developing and producing assistive products.
- Economic or cost-outcome evaluation of the provision of assistive products to individuals.

This section will consider the first type of economic evaluation and the second type will be discussed in Part 3 (Hersh, 2010b) of the paper, which deals with the evaluation of the outcomes of the use of assistive products. Although some tools can be used in both applications, what is being evaluated is quite different. It should be noted that, particularly in the case of products with high benefits to a small number of people, a particular product may have a very positive evaluation for the individual, but not be financially viable for a firm.

From the perspective of a firm or other organisation, financial analysis is necessary to determine whether a particular assistive product is financially viable and/or is likely to be profitable. This will require all financial costs to be covered by income, including subsidies and grants. In addition, where products are developed commercially a return on investment is generally required. A number of economic evaluation criteria can be used to investigate the financial viability of products. Most of them use discount rates to take account of the time value of money i.e. changes due to interest rates and inflation and risk. However, care has to be taken with the choice of discount rate to ensure that projects with high capital costs and an income stream over a long period are not automatically excluded.

- Pay-back period: The [amount](#) of time to recover the capital expenditure and any expenditure during the payback period. It does not consider the time value of money or expenditure after the payback period, so only gives a partial picture. Expenditure should only be made if the payback period is less than some value, such as a year.
- Net present value (NPV): It can be used to estimate the profit over a given period. Present values of income and expenditure over the period are calculated using a discount rate and the difference between them is the NPV. Non financial costs and benefits can be included if monetary values can be calculated for them. Expenditure should only be made if the NPV is positive. NPVs can also be used to compare different projects with the same lifespan, with higher values indicating better financial performance.
- Equivalent annual cost (EAC): The cost per year of owning and operating an asset over its entire lifespan. It is calculated by dividing the NPV by the present value of an annuity factor or multiplying the NPV by the loan repayment factor. An annuity factor is the present value of an income stream in a given currency that generates one unit of that currency each period for a specific number of periods. It can be used to compare investment projects with unequal durations, where the use of NPVs is generally inappropriate.
- Benefit:cost ratio (BCR): The ratio of the benefits to costs. All benefits and costs should be expressed in monetary terms using discounted present values. A variable discount rate with higher rates for cash flows considerably in the future is sometimes used, but this can discount the value of long term benefits to zero. It attempts to summarise the overall value for money of a project. It can also be used to compare projects (with the same lifespan) with higher values of BCR to be preferred. A BCR of one indicates that costs and benefits will be equal over the period considered, with values greater than one indicating that benefits will be greater than costs.
- Internal rate of return (IRR) or discounted cash flow rate of return: The interest rate at which the costs of an investment lead to its benefits. It is used to measure and compare the profitability of investments. The IRR should be greater than a given minimum value for the project to be accepted. Unlike the other indicators in this section, which are indicators of the value of the benefits of an investment, IRR is a rate quantity and is therefore an indicator of the efficiency or yield of an investment. It does not consider the costs of capital and should therefore not be used to compare projects with different durations.

As discussed in the literature on new product introduction, for instance (Cooper 1990; Griffin and Page 1993; Hart et al. 2003) and discontinuance of existing products, for instance (Avlonitis 1993), economic evaluation is an important, though by no means the only factor. It is also clearly prudent that economic evaluation is one of the factors that is considered at important points in the product life cycle. This section has given a brief overview of some of the methods that can be used to do this. However, length considerations prevent detailed

discussion of the advantages and disadvantages of the different approaches and their areas of application.

Other Approaches to Evaluation

Product Simulation

The use of prototypes and computer or human simulations of product functions to obtain end-user and stakeholder feedback enables end-user feedback to be obtained from an early stage in the design and development process before the final product is available. The use of simulations, in particular, can reduce costs. However, evaluations of this type require very clear specifications of the factors being investigated to ensure that a realistic representation of these factors is included in the simulation or prototype. This is important to avoid biasing the evaluation by, for instance, giving the impression that the assistive product works much better than is in fact the case or is possible with the current state of the underlying technology. For instance, a study of user preferences for speech and other modes of input used human simulation of the speech interface by listening and typing the user's speech output (Manaris et al. 2002). Speech recognition technology is by no means fully mature and its performance generally depends on a number of factors, including prior system training, the clarity of the speech input and noise and other environmental features. Since users' preferences are likely to be affected by how well the system works in practice, simulation in this case may have given misleading results and tests with the real system would be required to determine true user preferences.

Evaluation of Emotional Reactions to Products

A number of techniques, such as product personality profiling, mood boards and visual product evaluation, have been used to investigate users' emotional reactions to products (McDonagh et al. 2002). These techniques are based on the use of images and would therefore require modification for use with blind, visually impaired and deafblind people. In addition, all the techniques have drawbacks. In particular, product personality profiling draws on stereotypes and it may be difficult to interpret the results; mood boards (collages of images) can easily be misunderstood and the images may be too literal; and visual product evaluation requires good quality pictures and the forms and questionnaires used can restrict user responses. However, despite these drawbacks, the principle of investigating the emotional reaction of users to products is an important one and can usefully be extended to assistive products, an area where it has often been ignored.

Evaluation of Data Security

Data security is important for a number of different assistive products, including those which use user profiles to support product customisation or record user data, for instance to provide a service at a distance. Data security issues should therefore be included in the evaluation for these products. There are a number of different standards for information security (Eloff and von Solms 2000) which could form the basis of evaluation procedures. However, many of them seem to be aimed at information management systems and companies and other organisations rather than the information security issues relating to the use of certain products.

Decision Support and Object Based Modelling and Evaluation Systems

Decision Support Systems

Since one of the aims of evaluation is supporting decision making, for instance between particular designs or products, it is useful to consider decision support tools. Most decision support tools are based on the choice of the best alternative from a set through comparison of all the alternatives on one or more criteria, generally in terms of maximising satisfaction of one or more of these criteria. However, there is another approach to decision making, called satisficing which involves the sequential comparison of alternatives until a satisfactory one is found. Initially maximising may seem a better approach since it obtains the best alternative from all the available options. However, it may lead to the choice of the best, but still unsatisfactory, option from a number of poor alternatives, whereas satisficing could lead to attempts to improve the set of options by changing the problem constraints (Hersh 1999).

The evaluation of assistive products generally involves a number of different factors or criteria and will therefore usually require the use of multicriteria decision support systems. There are a number of different types. Multi-criteria problems can frequently be formulated in terms of a set of objectives, a set of criteria, a set of alternatives and a set of attributes which are used to determine the extent to which a particular alternative satisfies the criteria or objectives. The values and preferences of the decision maker and other stakeholders will often affect the choice of objectives and criteria and the prioritisation between them (Hersh 2006).

Decision problems have been divided into the following four stages (Simon 1977):

- Intelligence, involving information gathering.
- Design, involving problem formulation and the determination of families of alternatives, criteria and attributes.
- Choice, which generally involves the choice of an alternative from a set, but could also involve the choice of a particular type of approach or solution in multi-stage decision problems.
- Review, involving assessment of both the decision process and the solution obtained.

Both satisficing and maximising approaches can be used in multi-criteria problems. When a maximising approach is used in multi-criteria problems, there is not a unique solution and the multi-criteria algorithm can be used to identify an overall optimum solution, a solution set containing all the good or satisfactory alternatives and/or a rank order of all (or a subset) of the alternatives. The issue of multiple solutions is not relevant in the case of satisficing, as satisficing approaches will choose the first solution which meets all the criteria. Decision support methods have generally focused on either the choice of the 'best' alternative or ranking the alternatives from a given set using a pre-defined set of criteria and objectives. Using an appropriate algorithm and implementing it correctly will generally result in a good solution to the problem as formulated. This solution will only be useful in real terms if the important issues in the problem have been well represented and the solution can be implemented in practice.

There are a number of different multi-criteria methods, which can be categorised as follows (Hersh 2006), where more details and a case study of their use is provided:

- Outranking methods, including ELECTRE AND PROMETHEE.
- Aggregation of criteria methods, including P/G% and goal programming.
- The Analytic Hierarchy Process (AHP).
- Multi-attribute utility theory.

The following four tools support the development and use of decision support systems (Shim et al. 2002):

- Data warehouses, which are integrated collections of data on a particular topic which need to be updated over time (Immon 1992). They have taken advantage of improved database technologies. A number of firms have developed data warehouses, often resulting in increased interest in using decision support systems to analyse the data.
- On-line analytic processing tools, which transform the data to provide fast interactive access to various perspectives on it of relevance to the organisation.
- Data mining or database exploration tools obtain patterns in data and infer patterns from them and can be used for more sophisticated data analysis than on-line processing tools.
- World Wide Web and associated technologies, leading to the development of web-based decision support tools that deliver decision support information using a Web browser. These tools have facilitated access to decision relevant information and model driven decision support systems to geographically dispersed users at a reduced cost.

Decision support tools are potentially very useful in supporting the evaluation of products at different life cycle stages. They provide a framework to structure decision making on the comparative performance of different products and components on given criteria, as well as the evaluation of whether or not a given product meets the requisite criteria within the required tolerance. In addition, use of decision support methods provides a basis for explaining and justifying decisions to colleagues and management (Hammond et al. 1984). The decision support framework also provides support for ensuring that all relevant data has been collected and included in the problem. However, care may be required when formulating the decision problem from the real world situation to ensure that information is not degraded and in the conversion of qualitative to quantitative data (Hersh 1999). In addition, the use of decision support tools is probably most useful in evaluations involving a large number of criteria or when there is an organisational need to justify decision making, but less relevant in simpler problems involving a small number of criteria.

The use of expert systems can enhance decision support through encoding relevant expert knowledge and making it available throughout the organisation. Expert or knowledge-based support systems combine features of decision support systems and knowledge based approaches, such as expert systems (Liberatorou and Stylianos 1995). A limited number of expert systems have been developed to be used in project evaluation (Balachandra 1989; Liberatorou and Stylianos 1995; Wilkinson 1991) or screening new product ideas (Ram and Ram 1988). A number of expert support systems have been developed, but few of them have been applied to product evaluation or related areas and these systems generally have domain specific requirements. One of the few exceptions has been developed to provide a framework for the evaluation of product development projects. It has been applied to the evaluation of floor products (Liberatorou and Stylianos 1995). However, it does not seem to have been developed further. It is based on the following components:

- Knowledge and data acquisition module.

- Knowledge and data modelling, using Analytic Hierarchy Process, utility theory or other methods.
- End-user interface, based on a front-end expert system.
- Case history database to which an inductive expert system or neural network can be added to analyse the factors that lead to acceptance or termination of particular products.

The adaptation of decision and expert support tools for use in the evaluation of assistive (and other) products will probably require the development of specifically tailored easy to use software packages. A number of decision support systems for use in marketing have been developed, with overviews of what is available including (Matsatisnis and Siskos 1999; Wierenga 1992). Most of these systems are still at the prototype stage, have limited knowledge bases and cover only one phase of the new project development process. In addition, many of the earlier systems are not suited to qualitative data and have generally only been evaluated by the researchers in laboratory conditions (Matsatisnis and Siskos 1999). There has also been some development of decision support systems applicable to particular industries, such as the MOVIEMOD system for pre-release evaluation of films (Eliashberg et al. 2000). While some of these systems may be relevant to (assistive) product evaluation, considerable further development work will be required before they can be used with confidence. This will need to include a number of case studies and performance evaluation in real conditions.

Object-Based Modelling and Evaluation Systems

Object-based modelling and evaluation approaches (Pahng et al. 1998) decompose problems into modules that interact with each other and have embedded evaluation modules. These modules can be used to observe design quality from different viewpoints, compare different alternatives against criteria, such as cost, technical performance and ergonomic factors, and construct objective functions to be used in optimisation. The DOME (distributed object based modeling and evaluation) approach includes ‘lenses’ or specialised modules which provide multiple objective design evaluation services and include embedded probabilistic acceptability models. SPE-ED (Smith and Williams 1997) is an object oriented design evaluation tool which can be used to model software systems under development, evaluate software performance and identify potential hardware and software bottlenecks. Users use a graphical user interface and view the model as a scenario or execution graph of the software processing step, raising issues of accessibility to blind designers. SPE-ED outputs the end-to-end response time, the elapsed time and time at each computer device for each processing step. The use of object-based modelling and evaluation approaches is relatively new and the methods are still under development.

As in the case of decision support systems, these tools have some potential. However, their development has not yet reached the stage where they are suitable for application. Both further development of the underlying methodologies and the development of user friendly software packages, which have been fully tested in real world conditions, will be required. The price of such packages as well as the perceived benefits will affect their adoption, particularly by the SMEs which are often involved with the development of assistive products. It is therefore likely that, if adopted in the assistive technology area, the early adopters will be in areas such as hearing aid development, where there is a larger market.

Computer-aided software measurement and evaluation tools, including decision support and object-based modelling and evaluation systems, can be used to support and streamline the design and evaluation process. However, the tools will need to be chosen appropriately to

ensure that they are suitable for the specific software and hardware platforms. In addition, knowledge of the parameters of the software environment and an analysis of the effort and costs involved are required for the effective use of such tools (Dumke and Grigoleit 1997).

Conclusions

This article has continued the discussion in Part 1 on the design of assistive products with consideration of their evaluation. The final paper, Part 3, will consider evaluation of the outcomes of assistive product use by disabled and elderly people. There has been limited work on the evaluation of assistive products, most of which has focused on evaluation of the outcomes of their use. Thus this paper has had to draw on the literature on evaluation of consumer and other products. While it is considered that there are many similarities in the evaluation of different types of products, there is a need for further research to investigate this and for further development of methodologies specifically for the evaluation of assistive products.

The paper has drawn on and moved beyond the existing literature to propose the following:

- The use of a stage-gate approach to assistive product evaluation based on evaluation taking place throughout the product life cycle.
- An eight step process for planning and carrying out the evaluation at each stage.

This approach is well grounded in the literature on the assessment of consumer and other products. However, there is still a need for further research to investigate and validate the applicability of both the proposed stage-gate approach and the eight step process for carrying out the evaluation at each stage specifically for assistive products.

There has been some discussion of the criteria and techniques used to support evaluation at each stage. The criteria were categorised into the following four areas:

- Compliance and good practice, including health and safety, compliance with legislation and ethical issues.
- Technical issues, including technical performance, functionality, standards for interface components and reliability, robustness and dependability.
- End-user issues, including usability and acceptability, ergonomics, user satisfaction and performance and experience with the product.
- Resource and financial issues, including financial performance factors, resource utilisation and direct, indirect and hidden costs.

Further research will be required to investigate whether there are further criteria which should be taken into account when evaluating assistive products and whether all the criteria proposed in this paper are relevant in this case.

The techniques discussed include the use of decision and expert support systems, as well as object-based modelling and evaluation systems. While these support systems have some potential, most of the developments are still at the prototype stage and few of them have been tested outside the laboratory. In addition, to date none of these systems has been developed specifically for use in the evaluation of assistive products or been tested in this area. Before these systems are widely adopted, further research and development and testing, including in real product development and evaluation situations, will probably be required. There will

also be a need for the development of databases, data warehouses, knowledge bases and possibly other components which are specific to assistive products. In addition, the attractiveness and take-up of these tools are likely to depend on the availability of low cost or open source user friendly software packages, for which full technical support is available.

Acknowledgements

I would like to thank Mike Johnson and the two reviewers, Stephen Bauer and Renzo Andrich, for their useful and sometimes challenging comments and suggestions. I would also like to thank Peter McKenna for drawing the figures.

References

- Avlonitis GJ. 1993. Project Dropstrat: what factors do managers consider in deciding whether to drop a project. *European Journal of Marketing* 27(4):35-57.
- Balachandra R 1989. *Early Warning Signals for R&D Projects*. Lexington Books.
- Brentani de U. 1986. Do firms need a custom-designed new product screening model. *Journal of Product Innovation Management* 3:108-119.
- Chineck JW. 2007. Chapter 6: sensitivity analysis. In: Chineck JW. *Practical optimization: a gentle introduction*. [cited 2010 May 10]. Available from: <http://www.freetechbooks.com/practical-optimization-a-gentle-introduction-t761.html>.
- Cooper RG. 1990. Stage-gate systems: a new tool for managing new products. *Business Horizons* May-June:44-54.
- Comella-Dorda S et al. 2004. A process for COTS software product evaluation. Technical Report CMU/SEI-2003-TR-017.
- Dumke RR, Grigoleit D. 1997. Efficiency of CAME tools in software quality evaluation. *Software Quality Journal* 6:137-169.
- Eliashbeg J, Jonker J-J, Mohanbir SS, Wierenga B. 2000. MOVIEMOD: an implementable decision-support system for prerelease market-evaluation of motion pictures. *Marketing Science* 19(3):226-243.
- Eloff MM, von Solms SH. 2000. Information security management: an approach to combine process certification and product evaluation. *Computers and Security* 19:698-709.
- Fu TT, Cebon D. Economic evaluation and the design of vehicle suspensions. 2003. *International Journal of Vehicle Design* 31(2):125-161.
- Fuhrer MJ. 2001. Assistive technology outcomes research challenges met and yet unmet. *American Journal of Physical Medicine & Rehabilitation* 80(7):528-535.
- Griffin A, Page AL. 1993. An interim report on measuring product development success and failure. *Journal of Product Innovation Management* 10:291-308.
- Hammond KR, et al. 1984. *The relative efficiency of intuitive and analytical cognition*. Boulder (CO): Center for Research on Judgment and Policy.

- Hart S, Hultink EJ, Tzokas N, Commandeur HR. 2003. Industrial companies' evaluation criteria in new product development gates. *Product Innovation Management* 20:22-36.
- Hersh MA. 1999. Sustainable decision making: the role of decision support systems. *IEEE Transactions on Systems, Man and Cybernetics* 29(3):395-408.
- Hersh MA. 2006. *Mathematical modelling for sustainable development*. Springer Verlag. ISBN 3-540-24216-3.
- Hersh MA. 2010. The Design and Evaluation of Assistive Technology Products and Devices Part 1: Design. In: JH Stone, M Blouin, editors. *International Encyclopedia of Rehabilitation*. Available online: <http://cirrie.buffalo.edu/encyclopedia/en/article/309/>
- Hersh MA. 2010b. The Design and Evaluation of Assistive Technology Products and Devices Part 3: Outcomes. In: JH Stone, M Blouin, editors. *International Encyclopedia of Rehabilitation*.
- Hersh MA, Johnson MA. 2008. *Assistive technology for visually impaired and blind people*. Springer Verlag. 978-1-84628-866-1.
- Immon WH. 1992. *Building the data warehouse*. Wellesley (MA): QED Information Sciences.
- Jagou P. 1993. *Concurrent engineering*. Paris: Hermès.
- Kunda D, Brooks L. 1999. Applying Social-technical approach for COTS selection. *Proceedings of the 4th UKAIS Conference*. University of York: McGraw Hill.
- Kwakh J, Han SH. 2002. A methodology for evaluating the usability of audiovisual consumer electronic products. *Applied Ergonomics* 33:419-431.
- Liberatorou MT, Stylianou AC. 1995. Expert support systems for new product development decision making: a modeling framework and applications. *Management Science* 41(8):1296-1316.
- McDonagh D, Bruseberg A, Haslam C. 2002. Visual product evaluation: exploring users' emotional relationships with product. *Applied Ergonomics* 33:231-240.
- Manaris B, McGyvers V, Lagoudakis M. 2002. A listening keyboard for users with motor impairments – a usability study. *International Journal of Speech Technology* 5:371-388.
- Matsasinis NF and Siskos Y. 1999. MARKEX: an intelligent decision support system for product development decisions. *European Journal of Operational Research* 113:336-354.
- Nagel SS. 1988 *Evaluation analysis with microcomputers*. Greenwich (CT): JAI Press.
- Nielsen J. How to conduct a heuristic evaluation. [cited 2010 June 24]. Available from: http://www.useit.com/papers/heuristic/heuristic_evaluation.html.
- Nielsen J. 1993. *Usability engineering*. New York: Academic Press Ltd.

- Ozer M. 1999. A survey of new product evaluation models. *Journal of Product Innovation Management* 16:77-94.
- Pahng F, Senin N, Wallace D. 1998. Distribution modelling an evaluation of product design problems. *Computer-Aided Design* 30(6):411-423.
- Parmanto B, Zeng X. 2005. Metric for web accessibility evaluation. *Journal of the American Society for Information Science and Technology* 56(13):1394-1404.
- Ram S, Ram S. 1988. INNOVATOR: an expert system for new product launch decisions. *Applied Artificial Intelligence* 2(2):129-148.
- Raviart D, Sénéchal O, Tahon C. 1998. Methodology for a physico-economic product evaluation, starting in the design phase. *Control Engineering Practice* 6:489-498.
- Saaty TL. 1990. How to make a decision: the Analytic Hierarchy Process. *European Journal of Operational Research* 48:9-26.
- Saaty TL. 2008. Decision making with the Analytic Hierarchy Process. *International Journal of Services Sciences* 1(1):83-98.
- Schmidt JR, Sarangee, KR, Montoya MM. 2009. Exploring new product development project review practices. *Journal of Product Innovation Management* 26:520-535.
- Schmidt JR, Calantone RJ. 2002. Escalation of commitment during new product development. *Journal of the Academy of Marketing Science* 30:103-118.
- Shim JP, Warkentin M, Courtner JF, Power DJ, Sharda R, Carlsson C. 2002. Past, present and future of decision support technology. *Decision Support Systems* 33:111-126.
- Simon HA. 1977. *A new science of management decision*. Revised edition. Englewood Cliffs: Prentice-Hall.
- Smith CU, Williams LG. 1997. Performance evaluation of object oriented systems with SPE-ED. In: R Marie, et al., editors. *Computer performance evaluation: Modelling Techniques and Tools*, No. 1245. Springer Verlag.
- Song XM, Parry ME. 1996. What separates Japanese new product winners from losers. *Journal of Product Innovation Management* 13:422-439.
- Sweeney M, Maguire M, Shackel B. 1993. Evaluating user-computer interaction: a framework. *International Journal of Man-Machine Studies* 38:689-711.
- Tzokas N, Hultink EJ, Hart S, 2004. Navigating the new product development process. *Industrial Marketing Management* 33:619-626.
- WCAG 2.0 Web content accessibility guidelines. [Internet]. [updated 2008 Dec 11]. [cited 2010 June 24]. Available from: <http://www.w3.org/TR/WCAG20/>.
- Whitefield A, Wilson F, Dowell J. 1991. A framework for human factors evaluation. *Behaviour and Information Technology* 10(1):65-79.

Wierenga B. 1992. Knowledge-based systems in marketing: purpose, performance, perceptions and perspective. Rotterdam School of Management, Erasmus University. Working Paper 112.

Wilkinson A. 1991. Developing an expert system on project evaluation. R&D Management 21:19-29.