Better or Just Different? On the Benefits of Designing Interactive Systems in terms of Critical Parameters

William M. Newman
Rank Xerox Research Centre
61 Regent Street
Cambridge CB2 1AB, UK
+44 1223 341500; newman@cambridge.rxrc.xerox.com

ABSTRACT
Critical parameters are quantitative measures of performance that may be used to determine the overall ability of a design to serve its purpose. Although critical parameters figure in almost every field of design where there is a demand for progressive improvement, they do not appear to figure significantly in the design of interactive systems. As a result, systems are designed that are recognizably different from other systems but not necessarily better at doing the job intended. This paper discusses the role of critical parameters in design, and illustrates their lack of use in interactive system design by presenting a number of examples drawn from the HCI literature. It identifies a consequent need for research to establish critical parameters for applications and to build models of the performance of designs against these parameters. Some ideas are presented on how critical parameters might be established for specific applications, and the paper concludes by summarising some of the benefits that might be gained from moving in this direction.

KEYWORDS: Design, critical parameters, performance measurement.

INTRODUCTION
Does Version 2.0 need to be better than Version 1.0, or just different? Is the design of interactive systems a matter of making progressive improvements, or simply enabling the customer to tell products apart? I have begun to ponder on these questions recently, and to wonder if we place too much emphasis on deliberately designing systems to be different. Are we becoming too intent on following what David Kelley calls “the designer’s dream that goes beyond what exists”? [14] This line of thought has led me to ask myself whether, if we were to choose instead to focus on making systems progressively better, we could lay our hands on the methodological tools to do so. My conclusions, which I present in this paper, are that we do not, and that there is an important problem here for the HCI research community to solve.

The problem lies in the lack of identified critical parameters governing the design of interactive systems. I refer here to the established parameters by which designers measure whether an artefact or system serves its purpose, and compare one design with another. Critical parameters apply not just to an individual design project but to any attempt to apply technology to a particular purpose, such as designing an airport (where the main critical parameter is the airport’s capacity in flight operations per hour or per day) or developing a turbojet engine (where the main parameter is thrust in pounds). These are what Herbert Simon has called figures of merit, permitting comparison between designs in terms of “better” or “worse” [26]. They are application-specific, measuring how well the technology serves a specific purpose, not how well it performs in general. They are defined at the outset of design as performance targets, rather than later as evaluation parameters, and thus have a particularly strong guiding influence on design.

The design of interactive systems is unique, I believe, in its lack of use of critical parameters in spite of a recognised need to achieve progressive design improvements. There are many other fields of design where critical parameters are not applied as a matter of course, such as jewelry and fashion design, but in these fields the demand is not for improvements on past designs but for new designs that are different. In interactive system design we are in a peculiar situation where there is a constant demand for improved usability, but where little attempt is made to establish application-specific parameters by which improvements could be measured. Instead evaluation is performed, often as a separate activity, against parameters that are set by the evalua-
Critical Parameters in HCI: examples

In the course of this paper I will be citing a number of published accounts of design, and I will start with two that illustrate my main point. The first is the well-known study by Gray, John and Atwood of workstations designed to support Toll and Assistance Operators (TAOs), a study also known as “Project Ernestine” [8]. Here a well-defined and widely-accepted performance parameter exists, in the form of the average time taken by operators to handle any particular type of call, such as a collect call. This parameter is undeniably critical to the application: reducing call-handling times has been the primary justification for developing successive generations of operators’ workstations. One of Gray et al.’s contributions was to construct a CPM-GOMS model that enabled accurate predictions of performance to be made. But my main point is that designing in terms of such parameters, with or without the means for performance predictions, is rare in interactive system design.

A more typical approach to performance is found in another well-known example, the design of the Xerox Star user interface [25]. This pioneering design effort resulted in the first commercially available personal computer with a graphical desktop display. The development of Star was prompted largely by the opportunity for a radically new design approach, presented by the advent of bitmapped personal workstations, the mouse, windowed displays, laser printers and WYSIWYG editing. Performance targets did not figure in the design, a point that has been made in the retrospective account by Johnson et al. [13]. Instead the challenge was, as they put it, to pull all of the ideas together to produce a coherent functional design. Subsequent designs for graphical desktops have maintained this emphasis on functionality. Thus in Sullivan’s study of the usability of Windows® 95 it appears that critical parameters still do not govern the design of desktop environments. Emphasis remains on making functional innovations and eradicating functional problems [28].

In this paper I present the lack of application-specific critical parameters as a problem for designers and a challenge for researchers. I explain the role of critical parameters in design, and point out the lack of reference to them in the HCI literature. I develop some ideas on how critical parameters might be established for specific applications, and enumerate some of the benefits that might be gained by doing this.

The Roles of Critical Parameters in Design

A critical parameter, such as the capacity in vehicles/hour of a traffic intersection, provides the designer with a primary unit of performance against which to predict or measure the design’s ability to meet a set target. It provides a measure in terms of the purpose of the artefact or system (e.g., achieving a certain traffic flow rate), rather than in terms of its functional design (width or number of traffic lanes). Critical parameters therefore provide a direct and manageable measure of the design’s ability to serve its purpose. Furthermore, once critical parameters become established they can act as a focus for the development of design methods. The cumulative effect is to assist design in terms of performance targets.

Critical parameters figure constantly in those fields of design that have a culture of progressive improvement. They are defined explicitly by those who write about design, e.g., Rogers [23] and Vincenti [31], and implicitly in virtually every engineering textbook. Rogers discusses at some length the designer’s need to predict critical performance values in designs. Vincenti, in his insightful accounts of pioneering aeronautical design work, explores the topic of critical parameters and how they are established through research.

In these and other accounts of design we can see critical parameters making a number of crucial contributions to the progressive improvement of designs. Most crucial, perhaps, is the role identified earlier, the provision of a basis for measuring how well the design serves its purpose. These parameters are called “critical” because the success or failure of the design hinges critically on whether target values are met. Thus success in making improvements to a traffic intersection will depend critically on whether a certain set of flow rates and delay times are achieved. At the same time there may be an increase or decrease in the number of squeegee teams working at the intersection, or of tyre marks on the road surface, but these parameters are not held to be critical to design success.

Critical parameters tend to be invariant for a particular class of design problem, e.g., the design of suspension bridges, of automobile fuel pumps, or of airport runways. So universal are these parameters that they rarely need to be introduced or discussed; their use is taken for granted. Vincenti quotes the parameter invariably used to set targets for flying qualities in aircraft—stick force per g acceleration—and points out that today’s aeronautical engineers find it difficult to believe that any other parameter was ever used in this role [31].

Critical parameters are not always obvious, or easily identified. Indeed Vincenti describes how, in the case of flying qualities, years of research were needed to establish stick force per g as the primary critical parameter. As I point out later, the same kind of research investment may be needed to establish critical parameters for interactive system design.
The definition of critical parameters does not stifle innovation, indeed innovation may be made easier. Once established and taken for granted, critical parameters create an opportunity for innovation: a variety of forms of solution can be devised with the aim of achieving higher performance against the parameters, or maintaining performance while reducing cost. Consider traffic intersections, for example. The two main critical parameters—flow rates and delay—permit solutions ranging from the simple crossroads through signaled intersections to cloverleaf and flyover junctions. Radical inventions may also occur: Frank Whittle's invention of the turbojet engine was prompted by the desire to overcome apparent limits in aircraft performance against two critical parameters, speed and range [5].

Critical parameters define the units of measurement when performance requirements are drawn up. The task of setting these requirements is greatly simplified if the parameters are already taken for granted, as critical parameters usually are. This may explain why, in most fields of engineering other than software engineering, requirements do not figure as a major methodological issue.

One further role of critical parameters, already mentioned, is to encourage the development of methods for designing to meet targets. Since performance targets are based on critical parameters, which in turn provide a measure of design success, and since critical parameters are relatively invariant, it is well worthwhile to invest in developing analytical models for predicting the performance of designs. The availability of these models makes it easier to test the design before building a prototype. Thus the establishment of critical parameters paves the way for research that in turn builds a design discipline that is reliable, analytical and yet innovative.

CRITICAL PARAMETERS IN THE DESIGN OF INTERACTIVE SYSTEMS
To what extent do critical parameters figure in the design of interactive systems? I have claimed earlier that they play little or no part, and in this section I will explain why I make this claim. I base it partly on published accounts of how interactive systems are designed, and partly on my own experience. I will quote examples from both.

Critical Parameters in published accounts of design
As a starting point, I will take the above-mentioned roles and characteristics of critical parameters, and look for references to them in the literature on interactive system design. In this survey I draw a distinction between the design of interactive systems, applied to supporting human endeavour, and of user-interface components such as menus, pointing devices and gesture recognisers. There is a well-established culture of measuring the performance of new user-interface components, but interactive systems are a different matter entirely. Not only are critical parameters ignored here, but nothing else is offered that could play their roles instead.

First and foremost, it appears that the design of interactive systems rarely takes place in the context of measuring performance. Rather the usual design goal is to overcome functional shortcomings of existing support systems. This leads to adopting a new functional form of solution, which in turn involves the designers in a succession of design iterations to deal with functional gaps, side-effects and technical opportunities. Several interesting examples of this approach to design can be found in the sequences of papers on systems for audio indexing [1, 6, 27], for choreography [24, 3], and for remote collaboration via transparent interfaces [29, 11, 12]. In these and many other published examples the purpose of the system is clear but there is no mention of measuring how well this purpose is achieved.

Second, when designers do identify the performance of interactive systems as an issue, they can rarely take advantage of established, taken-for-granted performance parameters. Performance problems are often identified via contact with users, a source that does not necessarily provide clear, precise identification of performance parameters. Evidence of this can be seen in many design projects described in the literature, e.g., in [10] and [4]. In such situations, defining appropriate performance parameters becomes part of the design task and this, as I have mentioned earlier, can be a difficult and time-consuming research problem in itself—see, for example, Lee and Whalen's discussion of evaluation parameters for facial recognition systems [15].

A third issue I have mentioned above is the definition of performance requirements. This too is something that seems scarcely ever to happen in documented cases of interactive system design. Indeed it was the question of how interactive system requirements are specified that led me in 1992 to start an extensive study of the literature, which resulted eventually in an analysis of the products of HCI research [18]. I was interested at the time in comparing how interactive system designers and engineering designers specify requirements, but I was unable to find any real points of comparison. It appeared that the definition of performance requirements was not part of the HCI culture. I have since come to realise that the lack of identified performance parameters may be to blame.

Finally, there is a severe and widely recognised lack of analytical models for use in simulating the performance of interactive system designs [21, 19]. Various reasons are given for this. One possible explanation lies in the need to identify application-specific performance parameters in order to build useful predictive models.
Critical parameters are not totally absent from the interactive systems literature, however. I have mentioned one instance of their use, Project Ernestine, in the Introduction; and there are others. For example, a paper by Ballas et al. addresses the problem of automation deficit in aircraft autopilots [2]. This is a term for the delay between when the autopilot is switched off and when the pilot regains full effectiveness in manual flying. Automation deficit, measured in seconds, is regarded as a critical parameter for autopilot design. Ballas et al. describe how they developed a number of tactical display designs and tested their impact on automation deficit. More recently Pirolli et al. have experimented with different methods of extracting usable structures from the Web, working in terms of a critical parameter of information gain per unit time [22]. And Hartson et al. describe technologies for performing usability evaluation remotely, in which the main design parameter is the cost of evaluation [9]. Interestingly, in each case at least two different design approaches are presented and compared, bearing out a point I make earlier: the availability of critical parameters can encourage designers to try a variety of solutions.

**Some personal experiences**

As a sporadic designer, I find it insightful to contrast my own experiences with published accounts by others. An early example was the Officetalk-Zero project, whose goal was the design and construction of an integrated office system [16]. The first version of the system barely worked, and served primarily as a feasibility model for the second version, OZ-2, shown in Figure 1. This prototype was then tested in a field trial at a Xerox branch office in California. The main finding was that the system was not robust enough to be used operationally. A second finding, perhaps more interesting in the context of this paper, was that the system was too slow for the users to get their work done even when it was performing reliably. Tasks that they could previously complete in an hour or two were taking them most of the day. There was to be no third version of Officetalk-Zero, although its design did influence the Xerox Star [13]. What I find interesting twenty years later is that our preliminary studies had given us a clear indication of how long our users’ tasks were taking. We had an opportunity, therefore, to establish critical parameters for the application design. If we had done so, we could then have taken steps to avoid the disappointments of the field trial.

A more recent project, still ongoing, is the DigitalDesk [32, 33]. The aim here has been to investigate the use of an over-the-desk video camera and projector to support the user’s interaction with electronic and paper documents (Figure 2a). The DigitalDesk’s technology has been tried out on a succession of applications. In one of these it was used for freehand drawing (Figure 2b), an application that proved very popular but did not lead to a quantitative comparison with alternative methods, either paper-based or electronic. I am not sure even now how this comparison could have been conducted, since there are no obvious critical parameters for the freehand drawing application.

![Figure 1: The Officetalk-Zero system of 1978 [16].](image1)

![Figure 2: (a) The DigitalDesk [32]; (b) PaperPaint, its application to freehand drawing](image2)
A second DigitalDesk application was to provide translations of foreign words, projected onto the desk alongside the foreign-language paper document (Figure 3). Here we were able to conduct an informal comparison with conventional paper dictionaries, showing that the DigitalDesk could reduce the lookup time significantly [17]. Lookup time may indeed be regarded as a critical parameter for this application, but we have not carried out the necessary research to verify this.

Designing without critical parameters: the implications
These examples illustrate the almost universal tendency to ignore critical parameters when designing interactive systems. Space does not allow me to extend the survey to cover a larger sample. However, my impression from conducting more extensive reviews of the HCI literature is that it contains very few references to critical parameters or even to the concept of improving on the performance of existing designs [18].

If indeed critical parameters are being overlooked to this degree, then there are some serious implications for designers and researchers. The design of interactive systems may be made significantly more difficult by having to work without the benefit of identified critical parameters. Problems must inevitably arise at the very outset, when requirements are being drawn up. If instead critical parameters could be defined in advance, requirements definition would then be largely a matter of setting target values for these parameters. As it is, performance requirements for interactive systems are often very hard to define, and instead a set of functional requirements must be drawn up, a laborious task. Furthermore, the general lack of identified critical parameters must make it hard to develop predictive models of system performance; without these models it becomes difficult for the designer to determine how well the design is likely to serve its purpose. Changes to the system’s functionality cannot easily be translated into predicted changes in performance. One further problem is the constant need during design to decide which usability parameters should be measured, and what levels of performance should be considered satisfactory. It is difficult to avoid a degree of arbitrariness or bias in the choice of measures. It appears to me, overall, that many design projects must be experiencing difficulties as a result of lack of identified critical parameters.

A PROPOSAL: INCREASING THE OPPORTUNITIES TO DESIGN WITH CRITICAL PARAMETERS
The obvious way out of these difficulties is to enable designers to base their work on critical parameters. This is easier said than done. Critical parameters are difficult to establish, and designers are not easily persuaded to change their methods. What I suggest is needed is a relatively low-key research effort to increase our knowledge of critical parameters and how they can assist in design. I have a modest proposal on how this research might be got under way.

Where critical parameters lie
The thrust of the proposed research is towards identifying critical parameters for a range of applications. By “applications” I do not mean interactive systems, but human activities that interactive systems might support. For illustrative purposes I will choose four such applications and explore how we might establish their critical parameters:

- The authoring of documents in a professional institution;
- The gathering of research materials in support of writing articles and dissertations;
- Document-related activities of engineering managers;
- Transcription of audio tapes recorded during police interviews.

My reason for choosing these four applications is that each of them is commonly supported by the same kind of interactive system, namely a word processor. However, the critical parameters of each application would appear to be entirely different. I am not aware that any of the four sets of critical parameters have ever been identified, even though all four of these applications are extremely common and even universal within the relevant working communities. But studies of these applications suggest how we might go about identifying the critical parameters of each.

In international institutions, such as the United Nations or the World Bank, professional staff tend to spend a large amount of their time authoring documents whose purpose is to aid the decision-making processes of the institution. Studies indicate that this authoring work is generally carried out at the last minute: writing is timed to commence a certain number of days or hours before the document is due, allowing just enough time to complete it [20]. A variety of methods are used to ensure that the authoring is completed in time. For example, a busy author may delegate the entire task to a deputy, or may split the document into several parts to be worked on in parallel by members of a team. The process tends to involve a succession of steps of different kinds: outlining, drafting, editing, reviewing, correcting, etc. For these reasons it is hard to identify critical parameters in any of the individual steps. However, the overall time to complete the document appears to represent a highly critical parameter. If the standard time to complete a five-page text for a speech is two days, working flat-out, then a valid design problem may be to devise a technology that will reduce this time, without affecting the quality of the resulting document.
In the humanities departments of universities, researchers spend large amounts of their time reading source documents. Their ultimate aim is to develop material for papers and dissertation chapters they will write, incorporating their own interpretations of the source documents and supporting these with verbatim quotes. Two favoured modes of work are to read photocopies of the source documents, highlight and annotating the copies; or to take notes and copy out passages verbatim. A number of potentially critical parameters can be observed in the researcher’s work: they include the time taken to copy a passage verbatim from a source document, per line of text; and the overall time taken to read and take notes on a document, per page of the document. Any technology that could achieve a measurable reduction in such parameters would almost certainly be regarded as useful by researchers. Figure 4 shows a recently developed version of the DigitalDesk that might achieve this [30].

In engineering management, a large proportion of the working day is spent in a continuous round of meetings with team members or with other managers. Managers need to refer to documents during these meetings, and the time between meetings is often spent collecting information for the next meeting [7]. Managers resort to standard techniques for faxing, printing, reading email, making telephone inquiries, etc. An important consideration for the manager is the amount of time available between meetings, for this affects how much information can be gathered. Each standard technique incurs its own time penalty, but if this penalty can be reduced the manager can perhaps squeeze more information-gathering into the gap between meetings or shorten the gap. Therefore the time to perform each type of information gathering may be considered a critical parameter.

In police work, interviews with suspects and witnesses must by law be recorded on audio tape, and every tape must subsequently be transcribed. In the UK, police officers must often type the transcriptions themselves, a task that can take up to ten times as long as the original interview. Here the need for effective supporting technology is obvious, as is the critical parameter—the amount of time taken to transcribe each hour of audio.

**Identifying critical parameters**

My reason for quoting these four applications is to point out, firstly, that it may be feasible to identify critical parameters in each one. It is important to note that these parameters vary considerably from one application to the next; yet all four applications are supportable by the same technology, the word processor. Thus I am suggesting that there will sometimes be a range of different sets of critical parameters for an interactive system, each set corresponding to a different application. This perhaps explains why the design of highly generic systems, such as word processors and desktop environments, is so difficult and why improvements in their design are not always apparent to users.

Identifying critical parameters may prove to be a time-consuming research task, but it offers long-term payoffs. Work activities such as authoring, collecting research material, holding engineering meetings and transcribing interviews are widespread and thoroughly well-established in work practice. Although methods for performing them may change, the overall parameters governing their performance may change much less rapidly. For example, in the institution studied in [20], last-minute authoring appeared to have been standard practice for at least thirty years, perhaps longer. Likewise, notetaking by the present generation of PhD students appears to follow the patterns of previous generations. Research effort invested in identifying the critical parameters of such work is unlikely to go to waste.

**CONCLUSION**

I will conclude with a summary of the benefits that I believe might follow from paying more attention to critical parameters in the design of interactive systems.

A primary advantage would be the opportunity to harness the creative abilities of designers towards building new systems offering significant advantages over existing systems. I have great sympathy for designers who want to make their mark in interactive systems design, but who cannot provide convincing evidence that their design represents a major advance. I have often been in this position myself. In the absence of critical parameters the decision on whether an
A similar advantage lies in wait, I believe, for behavioral scientists, in the form of two interdependent seams of research that could have a major impact on interactive system design. The establishment of application-specific critical parameters, through studies and analyses, would form one seam; the development of analytical models to support predictions of performance would form another. The latter is an area where progress is urgently needed, for the analytical modelling tools available to the designer are few and far between. Until now, research efforts in this area have been hampered by the difficulty of developing useful generic models of interaction. If critical parameters could be established for individual applications, specific models of these applications could be developed, like the specific model of TAO task performance developed in Project Ernestine [8]. Each of these two lines of research would generate results of immediate value to designers.

A third major benefit can be gained by providing stronger links between interactive system design and the computer science of embedded hardware and software components. The performance of an interactive system is bound to be influenced by the performance of its components. But how well do we understand the nature of this influence? Can we assume, for example, that improvements to the performance of components (e.g., to the speed of the communication line to a TAO workstation) will improve performance overall? The answer is that we cannot: the performance relationships are highly non-linear. For example, the Project Ernestine research showed that “improvements” to the communication line had contributed to a degradation in speed of task performance [8]. There is an opportunity here to inform computer scientists how they can help improve interactive systems through improvements to embedded technologies.

Thus the study of critical parameters offers potential benefits for system designers, HCI researchers and computer scientists. In particular, the HCI community might in this way make more rapid progress towards helping the designers of interactive systems. They might help bring about an increase in the number of designs that are measurably better, without necessarily impeding the search for designs that are different.

REFERENCES
8


