

Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology

By: Fred D. Davis
Computer and Information Systems
Graduate School of Business
Administration
University of Michigan
Ann Arbor, Michigan 48109

Abstract

Valid measurement scales for predicting user acceptance of computers are in short supply. Most subjective measures used in practice are unvalidated, and their relationship to system usage is unknown. The present research develops and validates new scales for two specific variables, perceived usefulness and perceived ease of use, which are hypothesized to be fundamental determinants of user acceptance. Definitions for these two variables were used to develop scale items that were pretested for content validity and then tested for reliability and construct validity in two studies involving a total of 152 users and four application programs. The measures were refined and streamlined, resulting in two six-item scales with reliabilities of .98 for usefulness and .94 for ease of use. The scales exhibited high convergent, discriminant, and factorial validity. Perceived usefulness was significantly correlated with both self-reported current usage ($r = .63$, Study 1) and self-predicted future usage ($r = .85$, Study 2). Perceived ease of use was also significantly correlated with current usage ($r = .45$, Study 1) and future usage ($r = .59$, Study 2). In both studies, usefulness had a significantly greater correlation with usage behavior than did ease of use. Regression analyses suggest that perceived ease of use may actually be a causal antece-

dent to perceived usefulness, as opposed to a parallel, direct determinant of system usage. Implications are drawn for future research on user acceptance.

Keywords: User acceptance, end user computing, user measurement

ACM Categories: H.1.2, K.6.1, K.6.2, K.6.3

Introduction

Information technology offers the potential for substantially improving white collar performance (Curley, 1984; Edelman, 1981; Sharda, et al., 1988). But performance gains are often obstructed by users' unwillingness to accept and use available systems (Bowen, 1986; Young, 1984). Because of the persistence and importance of this problem, explaining user acceptance has been a long-standing issue in MIS research (Swanson, 1974; Lucas, 1975; Schultz and Slevin, 1975; Robey, 1979; Ginzberg, 1981; Swanson, 1987). Although numerous individual, organizational, and technological variables have been investigated (Benbasat and Dexter, 1986; Franz and Robey, 1986; Markus and Bjorn-Anderson, 1987; Robey and Farrow, 1982), research has been constrained by the shortage of high-quality measures for key determinants of user acceptance. Past research indicates that many measures do not correlate highly with system use (DeSanctis, 1983; Ginzberg, 1981; Schewe, 1976; Srinivasan, 1985), and the size of the usage correlation varies greatly from one study to the next depending on the particular measures used (Baroudi, et al., 1986; Barki and Huff, 1985; Robey, 1979; Swanson, 1982, 1987). The development of improved measures for key theoretical constructs is a research priority for the information systems field.

Aside from their theoretical value, better measures for predicting and explaining system use would have great practical value, both for vendors who would like to assess user demand for new design ideas, and for information systems managers within user organizations who would like to evaluate these vendor offerings.

Unvalidated measures are routinely used in practice today throughout the entire spectrum of design, selection, implementation and evaluation activities. For example: designers within vendor organizations such as IBM (Gould, et al., 1983), Xerox (Brewley, et al., 1983), and Digital Equip-

ment Corporation (Good, et al., 1986) measure user perceptions to guide the development of new information technologies and products; industry publications often report user surveys (e.g., Greenberg, 1984; Rushinek and Rushinek, 1986); several methodologies for software selection call for subjective user inputs (e.g., Goslar, 1986; Klein and Beck, 1987); and contemporary design principles emphasize measuring user reactions throughout the entire design process (Anderson and Olson 1985; Gould and Lewis, 1985; Johansen and Baker, 1984; Mantei and Teorey, 1988; Norman, 1983; Shneiderman, 1987). Despite the widespread use of subjective measures in practice, little attention is paid to the quality of the measures used or how well they correlate with usage behavior. Given the low usage correlations often observed in research studies, those who base important business decisions on unvalidated measures may be getting misinformed about a system's acceptability to users.

The purpose of this research is to pursue better measures for predicting and explaining use. The investigation focuses on two theoretical constructs, perceived usefulness and perceived ease of use, which are theorized to be fundamental determinants of system use. Definitions for these constructs are formulated and the theoretical rationale for their hypothesized influence on system use is reviewed. New, multi-item measurement scales for perceived usefulness and perceived ease of use are developed, pretested, and then validated in two separate empirical studies. Correlation and regression analyses examine the empirical relationship between the new measures and self-reported indicants of system use. The discussion concludes by drawing implications for future research.

Perceived Usefulness and Perceived Ease of Use

What causes people to accept or reject information technology? Among the many variables that may influence system use, previous research suggests two determinants that are especially important. First, people tend to use or not use an application to the extent they believe it will help them perform their job better. We refer to this first variable as *perceived usefulness*. Second, even if potential users believe that a given application is useful, they may, at the same time,

believe that the systems is too hard to use and that the performance benefits of usage are outweighed by the effort of using the application. That is, in addition to usefulness, usage is theorized to be influenced by *perceived ease of use*.

Perceived usefulness is defined here as "the degree to which a person believes that using a particular system would enhance his or her job performance." This follows from the definition of the word useful: "capable of being used advantageously." Within an organizational context, people are generally reinforced for good performance by raises, promotions, bonuses, and other rewards (Pfeffer, 1982; Schein, 1980; Vroom, 1964). A system high in *perceived usefulness*, in turn, is one for which a user believes in the existence of a positive use-performance relationship.

Perceived ease of use, in contrast, refers to "the degree to which a person believes that using a particular system would be free of effort." This follows from the definition of "ease": "freedom from difficulty or great effort." Effort is a finite resource that a person may allocate to the various activities for which he or she is responsible (Radner and Rothschild, 1975). All else being equal, we claim, an application perceived to be easier to use than another is more likely to be accepted by users.

Theoretical Foundations

The theoretical importance of perceived usefulness and perceived ease of use as determinants of user behavior is indicated by several diverse lines of research. The impact of perceived usefulness on system utilization was suggested by the work of Schultz and Slevin (1975) and Robey (1979). Schultz and Slevin (1975) conducted an exploratory factor analysis of 67 questionnaire items, which yielded seven dimensions. Of these, the "performance" dimension, interpreted by the authors as the perceived "effect of the model on the manager's job performance," was most highly correlated with self-predicted use of a decision model ($r = .61$). Using the Schultz and Slevin questionnaire, Robey (1979) finds the performance dimension to be most correlated with two objective measures of system usage ($r = .79$ and $.76$). Building on Vertinsky, et al.'s (1975) expectancy model, Robey (1979) theorizes that: "A system that does not help people perform their jobs is not likely to be received favorably

in spite of careful implementation efforts" (p. 537). Although the perceived use-performance contingency, as presented in Robey's (1979) model, parallels our definition of perceived usefulness, the use of Schultz and Slevin's (1975) performance factor to operationalize performance expectancies is problematic for several reasons: the instrument is empirically derived via exploratory factor analysis; a somewhat low ratio of sample size to items is used (2:1); four of thirteen items have loadings below .5, and several of the items clearly fall outside the definition of expected performance improvements (e.g., "My job will be more satisfying," "Others will be more aware of what I am doing," etc.).

An alternative expectancy-theoretic model, derived from Vroom (1964), was introduced and tested by DeSanctis (1983). The use-performance expectancy was not analyzed separately from performance-reward instrumentalities and reward valences. Instead, a matrix-oriented measurement procedure was used to produce an overall index of "motivational force" that combined these three constructs. "Force" had small but significant correlations with usage of a DSS within a business simulation experiment (correlations ranged from .04 to .26). The contrast between DeSanctis's correlations and the ones observed by Robey underscore the importance of measurement in predicting and explaining use.

Self-efficacy theory

The importance of perceived ease of use is supported by Bandura's (1982) extensive research on self-efficacy, defined as "judgments of how well one can execute courses of action required to deal with prospective situations" (p. 122). Self-efficacy is similar to perceived ease of use as defined above. Self-efficacy beliefs are theorized to function as proximal determinants of behavior. Bandura's theory distinguishes self-efficacy judgments from outcome judgments, the latter being concerned with the extent to which a behavior, once successfully executed, is believed to be linked to valued outcomes. Bandura's "outcome judgment" variable is similar to perceived usefulness. Bandura argues that self-efficacy and outcome beliefs have differing antecedents and that, "In any given instance, behavior would be best predicted by considering both self-efficacy and outcome beliefs" (p. 140).

Hill, et al. (1987) find that both self-efficacy and outcome beliefs exert an influence on decisions

to learn a computer language. The self efficacy paradigm does not offer a general measure applicable to our purposes since efficacy beliefs are theorized to be situationally-specific, with measures tailored to the domain under study (Bandura, 1982). Self efficacy research does, however, provide one of several theoretical perspectives suggesting that perceived ease of use and perceived usefulness function as basic determinants of user behavior.

Cost-benefit paradigm

The cost-benefit paradigm from behavioral decision theory (Beach and Mitchell, 1978; Johnson and Payne, 1985; Payne, 1982) is also relevant to perceived usefulness and ease of use. This research explains people's choice among various decision-making strategies (such as linear compensatory, conjunctive, disjunctive and elimination-by-aspects) in terms of a cognitive trade-off between the effort required to employ the strategy and the quality (accuracy) of the resulting decision. This approach has been effective for explaining why decision makers alter their choice strategies in response to changes in task complexity. Although the cost-benefit approach has mainly concerned itself with unaided decision making, recent work has begun to apply the same form of analysis to the effectiveness of information display formats (Jarvenpaa, 1989; Kleinmuntz and Schkade, 1988).

Cost-benefit research has primarily used objective measures of accuracy and effort in research studies, downplaying the distinction between objective and subjective accuracy and effort. Increased emphasis on subjective constructs is warranted, however, since (1) a decision maker's choice of strategy is theorized to be based on subjective as opposed to objective accuracy and effort (Beach and Mitchell, 1978), and (2) other research suggests that subjective measures are often in disagreement with their objective counterparts (Abelson and Levi, 1985; Adelbratt and Montgomery, 1980; Wright, 1975). Introducing measures of the decision maker's own perceived costs and benefits, independent of the decision actually made, has been suggested as a way of mitigating criticisms that the cost/benefit framework is tautological (Abelson and Levi, 1985). The distinction made herein between perceived usefulness and perceived ease of use is similar to the distinction between subjective decision-making performance and effort.

Adoption of innovations

Research on the adoption of innovations also suggests a prominent role for perceived ease of use. In their meta-analysis of the relationship between the characteristics of an innovation and its adoption, Tornatzky and Klein (1982) find that compatibility, relative advantage, and complexity have the most consistent significant relationships across a broad range of innovation types. Complexity, defined by Rogers and Shoemaker (1971) as "the degree to which an innovation is perceived as relatively difficult to understand and use" (p. 154), parallels perceived ease of use quite closely. As Tornatzky and Klein (1982) point out, however, compatibility and relative advantage have both been dealt with so broadly and inconsistently in the literature as to be difficult to interpret.

Evaluation of information reports

Past research within MIS on the evaluation of information reports echoes the distinction between usefulness and ease of use made herein. Larcker and Lessig (1980) factor analyzed six items used to rate four information reports. Three items load on each of two distinct factors: (1) *perceived importance*, which Larcker and Lessig define as "the quality that causes a particular information set to acquire relevance to a decision maker," and the extent to which the information elements are "a necessary input for task accomplishment," and (2) *perceived usability*, which is defined as the degree to which "the information format is unambiguous, clear or readable" (p. 123). These two dimensions are similar to perceived usefulness and perceived ease of use as defined above, respectively, although Larcker and Lessig refer to the two dimensions collectively as "perceived usefulness." Reliabilities for the two dimensions fall in the range of .64-.77, short of the .80 minimal level recommended for basic research. Correlations with actual use of information reports were not addressed in their study.

Channel disposition model

Swanson (1982, 1987) introduced and tested a model of "channel disposition" for explaining the choice and use of information reports. The concept of channel disposition is defined as having

two components: attributed information quality and attributed access quality. Potential users are hypothesized to select and use information reports based on an implicit psychological trade-off between information quality and associated costs of access. Swanson (1987) performed an exploratory factor analysis in order to measure information quality and access quality. A five-factor solution was obtained, with one factor corresponding to information quality (Factor #3, "value"), and one to access quality (Factor #2, "accessibility"). Inspecting the items that load on these factors suggests a close correspondence to perceived usefulness and ease of use. Items such as "important," "relevant," "useful," and "valuable" load strongly on the value dimension. Thus, value parallels perceived usefulness. The fact that relevance and usefulness load on the same factor agrees with information scientists, who emphasize the conceptual similarity between the usefulness and relevance notions (Saracevic, 1975). Several of Swanson's "accessibility" items, such as "convenient," "controllable," "easy," and "unburdensome," correspond to perceived ease of use as defined above. Although the study was more exploratory than confirmatory, with no attempts at construct validation, it does agree with the conceptual distinction between usefulness and ease of use. Self-reported information channel use correlated .20 with the value dimension and .13 with the accessibility dimension.

Non-MIS studies

Outside the MIS domain, a marketing study by Hauser and Simmie (1981) concerning user perceptions of alternative communication technologies similarly derived two underlying dimensions: ease of use and effectiveness, the latter being similar to the perceived usefulness construct defined above. Both ease of use and effectiveness were influential in the formation of user preferences regarding a set of alternative communication technologies. The human-computer interaction (HCI) research community has heavily emphasized ease of use in design (Branscomb and Thomas, 1984; Card, et al., 1983; Gould and Lewis, 1985). For the most part, however, these studies have focused on objective measures of ease of use, such as task completion time and error rates. In many vendor organizations, usability testing has become a standard phase in the product development cycle, with

large investments in test facilities and instrumentation. Although objective ease of use is clearly relevant to user performance given the system is used, subjective ease of use is more relevant to the users' decision whether or not to use the system and may not agree with the objective measures (Carroll and Thomas, 1988).

Convergence of findings

There is a striking convergence among the wide range of theoretical perspectives and research studies discussed above. Although Hill, et al. (1987) examined learning a computer language, Larcker and Lessig (1980) and Swanson (1982, 1987) dealt with evaluating information reports, and Hauser and Simmie (1981) studied communication technologies, all are supportive of the conceptual and empirical distinction between usefulness and ease of use. The accumulated body of knowledge regarding self-efficacy, contingent decision behavior and adoption of innovations provides theoretical support for perceived usefulness and ease of use as key determinants of behavior.

From multiple disciplinary vantage points, perceived usefulness and perceived ease of use are indicated as fundamental and distinct constructs that are influential in decisions to use information technology. Although certainly not the only variables of interest in explaining user behavior (for other variables, see Cheney, et al., 1986; Davis, et al., 1989; Swanson, 1988), they do appear likely to play a central role. Improved measures are needed to gain further insight into the nature of perceived usefulness and perceived ease of use, and their roles as determinants of computer use.

Scale Development and Pretest

A step-by-step process was used to develop new multi-item scales having high reliability and validity. The conceptual definitions of perceived usefulness and perceived ease of use, stated above, were used to generate 14 candidate items for each construct from past literature. Pretest interviews were then conducted to assess the semantic content of the items. Those items that best fit the definitions of the constructs were

retained, yielding 10 items for each construct. Next, a field study (Study 1) of 112 users concerning two different interactive computer systems was conducted in order to assess the reliability and construct validity of the resulting scales. The scales were further refined and streamlined to six items per construct. A lab study (Study 2) involving 40 participants and two graphics systems was then conducted. Data from the two studies were then used to assess the relationship between usefulness, ease of use, and self-reported usage.

Psychometricians emphasize that the validity of a measurement scale is built in from the outset. As Nunnally (1978) points out, "Rather than test the validity of measures after they have been constructed, one should ensure the validity by the plan and procedures for construction" (p. 258). Careful selection of the initial scale items helps to assure the scales will possess "content validity," defined as "the degree to which the score or scale being used represents the concept about which generalizations are to be made" (Bohrnstedt, 1970, p. 91). In discussing content validity, psychometricians often appeal to the "domain sampling model," (Bohrnstedt, 1970; Nunnally, 1978) which assumes there is a domain of content corresponding to each variable one is interested in measuring. Candidate items representative of the domain of content should be selected. Researchers are advised to begin by formulating conceptual definitions of what is to be measured and preparing items to fit the construct definitions (Anastasi, 1986).

Following these recommendations, candidate items for perceived usefulness and perceived ease of use were generated based on their conceptual definitions, stated above, and then pretested in order to select those items that best fit the content domains. The Spearman-Brown Prophecy formula was used to choose the number of items to generate for each scale. This formula estimates the number of items needed to achieve a given reliability based on the number of items and reliability of comparable existing scales. Extrapolating from past studies, the formula suggests that 10 items would be needed for each perceptual variable to achieve reliability of at least .80 (Davis, 1986). Adding four additional items for each construct to allow for item elimination, it was decided to generate 14 items for each construct.

The initial item pools for perceived usefulness and perceived ease of use are given in Tables

1 and 2, respectively. In preparing candidate items, 37 published research papers dealing with user reactions to interactive systems were reviewed in order to identify various facets of the constructs that should be measured (Davis, 1986). The items are worded in reference to "the electronic mail system," which is one of the two test applications investigated in Study 1, reported below. The items within each pool tend to have a lot of overlap in their meaning, which is consistent with the fact that they are intended as measures of the same underlying construct. Though different individuals may attribute slightly different meaning to particular item statements, the goal of the multi-item approach is to reduce any extraneous effects of individual items, allowing idiosyncrasies to be cancelled out by

other items in order to yield a more pure indicant of the conceptual variable.

Pretest interviews were performed to further enhance content validity by assessing the correspondence between candidate items and the definitions of the variables they are intended to measure. Items that don't represent a construct's content very well can be screened out by asking individuals to rank the degree to which each item matches the variable's definition, and eliminating items receiving low rankings. In eliminating items, we want to make sure not to reduce the representativeness of the item pools. Our item pools may have excess coverage of some areas of meaning (or substrata; see Bohrnstedt, 1970) within the content domain and not enough of

Table 1. Initial Scale Items for Perceived Usefulness

1. My job would be difficult to perform without electronic mail.
2. Using electronic mail gives me greater control over my work.
3. Using electronic mail improves my job performance.
4. The electronic mail system addresses my job-related needs.
5. Using electronic mail saves me time.
6. Electronic mail enables me to accomplish tasks more quickly.
7. Electronic mail supports critical aspects of my job.
8. Using electronic mail allows me to accomplish more work than would otherwise be possible.
9. Using electronic mail reduces the time I spend on unproductive activities.
10. Using electronic mail enhances my effectiveness on the job.
11. Using electronic mail improves the quality of the work I do.
12. Using electronic mail increases my productivity.
13. Using electronic mail makes it easier to do my job.
14. Overall, I find the electronic mail system useful in my job.

Table 2. Initial Scale Items for Perceived Ease of Use

1. I often become confused when I use the electronic mail system.
2. I make errors frequently when using electronic mail.
3. Interacting with the electronic mail system is often frustrating.
4. I need to consult the user manual often when using electronic mail.
5. Interacting with the electronic mail system requires a lot of my mental effort.
6. I find it easy to recover from errors encountered while using electronic mail.
7. The electronic mail system is rigid and inflexible to interact with.
8. I find it easy to get the electronic mail system to do what I want it to do.
9. The electronic mail system often behaves in unexpected ways.
10. I find it cumbersome to use the electronic mail system.
11. My interaction with the electronic mail system is easy for me to understand.
12. It is easy for me to remember how to perform tasks using the electronic mail system.
13. The electronic mail system provides helpful guidance in performing tasks.
14. Overall, I find the electronic mail system easy to use.

others. By asking individuals to rate the similarity of items to one another, we can perform a cluster analysis to determine the structure of the substrata, remove items where excess coverage is suggested, and add items where inadequate coverage is indicated.

Pretest participants consisted of a sample of 15 experienced computer users from the Sloan School of Management, MIT, including five secretaries, five graduate students and five members of the professional staff. In face-to-face interviews, participants were asked to perform two tasks, prioritization and categorization, which were done separately for usefulness and ease of use. For prioritization, they were first given a card containing the definition of the target construct and asked to read it. Next, they were given 13 index cards each having one of the items for that construct written on it. The 14th or "overall" item for each construct was omitted since its wording was almost identical to the label on the definition card (see Tables 1 and 2). Participants were asked to rank the 13 cards according to how well the meaning of each statement matched the given definition of ease of use or usefulness.

For the categorization task, participants were asked to put the 13 cards into three to five categories so that the statements within a category were most similar in meaning to each other and dissimilar in meaning from those in other categories. This was an adaptation of the "own categories" procedure of Sherif and Sherif (1967). Categorization provides a simple indicant of similarity that requires less time and effort to obtain than other similarity measurement procedures such as paid comparisons. The similarity data was cluster analyzed by assigning to the same cluster items that seven or more subjects placed in the same category. The clusters are considered to be a reflection of the domain substrata for each construct and serve as a basis of assessing coverage, or representativeness, of the item pools.

The resulting rank and cluster data are summarized in Tables 3 (usefulness) and 4 (ease of use). For perceived usefulness, notice that items fall into three main clusters. The first cluster relates to job effectiveness, the second to productivity and time savings, and the third to the importance of the system to one's job. If we eliminate the lowest-ranked items (items 1, 4, 5 and 9), we see that the three major clusters each have at least two items. Item 2, "control

over work" was retained since, although it was ranked fairly low, it fell in the top 9 and may tap an important aspect of usefulness.

Looking now at perceived ease of use (Table 4), we again find three main clusters. The first relates to physical effort, while the second relates to mental effort. Selecting the six highest-priority items and eliminating the seventh provides good coverage of these two clusters. Item 11 ("understandable") was reworded to read "clear and understandable" in an effort to pick up some of the content of item 1 ("confusing"), which has been eliminated. The third cluster is somewhat more difficult to interpret but appears to be tapping perceptions of how easy a system is to learn. Remembering how to perform tasks, using the manual, and relying on system guidance are all phenomena associated with the process of learning to use a new system (Nickerson, 1981; Roberts and Moran, 1983). Further review of the literature suggests that ease of use and ease of learning are strongly related. Roberts and Moran (1983) find a correlation of .79 between objective measures of ease of use and ease of learning. Whiteside, et al. (1985) find that ease of use and ease of learning are strongly related and conclude that they are congruent. Studies of how people learn new systems suggest that learning and using are not separate, disjoint activities, but instead that people are motivated to begin performing actual work directly and try to "learn by doing" as opposed to going through user manuals or online tutorials (Carroll and Carrithers, 1984; Carroll, et al., 1985; Carroll and McKendree, 1987).

In this study, therefore, ease of learning is regarded as one substratum of the ease of use construct, as opposed to a distinct construct. Since items 4 and 13 provide a rather indirect assessment of ease of learning, they were replaced with two items that more directly get at ease of learning: "Learning to operate the electronic mail system is easy for me," and "I find it takes a lot of effort to become skillful at using electronic mail." Items 6, 9 and 2 were eliminated because they did not cluster with other items, and they received low priority rankings, which suggests that they do not fit well within the content domain for ease of use. Together with the "overall" items for each construct, this procedure yielded a 10-item scale for each construct to be empirically tested for reliability and construct validity.

Table 3. Pretest Results: Perceived Usefulness

Old Item #	Item	Rank	New Item #	Cluster
1	Job Difficult Without	13		C
2	Control Over Work	9	2	
3	Job Performance	2	6	A
4	Addresses My Needs	12		C
5	Saves Me Time	11		B
6	Work More Quickly	7	3	B
7	Critical to My Job	5	4	C
8	Accomplish More Work	6	7	B
9	Cut Unproductive Time	10		B
10	Effectiveness	1	8	A
11	Quality of Work	3	1	A
12	Increase Productivity	4	5	B
13	Makes Job Easier	8	9	C
14	Useful	NA	10	NA

Table 4. Pretest Results: Perceived Ease of Use

Old Item #	Item	Rank	New Item #	Cluster
1	Confusing	7		B
2	Error Prone	13		
3	Frustrating	3	3	B
4	Dependence on Manual	9	(replace)	C
5	Mental Effort	5	7	B
6	Error Recovery	10		
7	Rigid & Inflexible	6	5	A
8	Controllable	1	4	A
9	Unexpected Behavior	11		
10	Cumbersome	2	1	A
11	Understandable	4	8	B
12	Ease of Remembering	8	6	C
13	Provides Guidance	12	(replace)	C
14	Easy to Use	NA	10	NA
NA	Ease of Learning	NA	2	NA
NA	Effort to Become Skillful	NA	9	NA

Study 1

A field study was conducted to assess the reliability, convergent validity, discriminant validity, and factorial validity of the 10-item scales resulting from the pretest. A sample of 120 users within IBM Canada's Toronto Development Laboratory were given a questionnaire asking them to rate the usefulness and ease of use of two systems available there: PROFS electronic mail and the XEDIT file editor. The computing environment consisted of IBM mainframes accessible through 327X terminals. The PROFS electronic mail system is a simple but limited messaging facility for brief messages. (See Panko, 1988.) The XEDIT editor is widely avail-

able on IBM systems and offers both full-screen and command-driven editing capabilities. The questionnaire asked participants to rate the extent to which they agree with each statement by circling a number from one to seven arranged horizontally beneath anchor point descriptions "Strongly Agree," "Neutral," and "Strongly Disagree." In order to ensure subject familiarity with the systems being rated, instructions asked the participants to skip over the section pertaining to a given system if they never use it. Responses were obtained from 112 participants, for a response rate of 93%. Of these 112, 109 were users of electronic mail and 75 were users of XEDIT. Subjects had an average of six months' experience with the two systems studied. Among

the sample, 10 percent were managers, 35 percent were administrative staff, and 55 percent were professional staff (which included a broad mix of market analysts, product development analysts, programmers, financial analysts and research scientists).

Reliability and validity

The perceived usefulness scale attained Cronbach alpha reliability of .97 for both the electronic mail and XEDIT systems, while perceived ease of use achieved a reliability of .86 for electronic mail and .93 for XEDIT. When observations were pooled for the two systems, alpha was .97 for usefulness and .91 for ease of use.

Convergent and discriminant validity were tested using multitrait-multimethod (MTMM) analysis (Campbell and Fiske, 1959). The MTMM matrix contains the intercorrelations of items (methods) applied to the two different test systems (traits), electronic mail and XEDIT. Convergent validity refers to whether the items comprising a scale behave as if they are measuring a common underlying construct. In order to demonstrate convergent validity, items that measure the same trait should correlate highly with one another (Campbell and Fiske, 1959). That is, the elements in the monotrait triangles (the submatrix of intercorrelations between items intended to measure the same construct for the same system) within the MTMM matrices should be large. For perceived usefulness, the 90 monotrait-heteromethod correlations were all significant at the .05 level. For ease of use, 86 out of 90, or 95.6%, of the monotrait-heteromethod correlations were significant. Thus, our data supports the convergent validity of the two scales.

Discriminant validity is concerned with the ability of a measurement item to differentiate between objects being measured. For instance, within the MTMM matrix, a perceived usefulness item applied to electronic mail should not correlate too highly with the same item applied to XEDIT. Failure to discriminate may suggest the presence of "common method variance," which means that an item is measuring methodological artifacts unrelated to the target construct (such as individual differences in the style of responding to questions (see Campbell, et al., 1967; Silk, 1971)). The test for discriminant validity is that an item should correlate more highly with other items intended to measure the same trait than with either the same item used to measure a

different trait or with different items used to measure a different trait (Campbell and Fiske, 1959). For perceived usefulness, 1,800 such comparisons were confirmed without exception. Of the 1,800 comparisons for ease of use there were 58 exceptions (3%). This represents an unusually high level of discriminant validity (Campbell and Fiske, 1959; Silk, 1971) and implies that the usefulness and ease of use scales possess a high concentration of trait variance and are not strongly influenced by methodological artifacts.

Table 5 gives a summary frequency table of the correlations comprising the MTMM matrices for usefulness and ease of use. From this table it is possible to see the separation in magnitude between monotrait and heterotrait correlations. The frequency table also shows that the heterotrait-heteromethod correlations do not appear to be substantially elevated above the heterotrait-monomethod correlations. This is an additional diagnostic suggested by Campbell and Fiske (1959) to detect the presence of method variance.

The few exceptions to the convergent and discriminant validity that did occur, although not extensive enough to invalidate the ease of use scale, all involved negatively phrased ease of use items. These "reversed" items tended to correlate more with the same item used to measure a different trait than they did with other items of the same trait, suggesting the presence of common method variance. This is ironic, since reversed scales are typically used in an effort to reduce common method variance. Silk (1971) similarly observed minor departures from convergent and discriminant validity for reversed items. The five positively worded ease of use items had a reliability of .92 compared to .83 for the five negative items. This suggests an improvement in the ease of use scale may be possible with the elimination or reversal of negatively phrased items. Nevertheless, the MTMM analysis supported the ability of the 10-item scales for each construct to differentiate between systems.

Factorial validity is concerned with whether the usefulness and ease of use items form distinct constructs. A principal components analysis using oblique rotation was performed on the twenty usefulness and ease of use items. Data were pooled across the two systems, for a total of 184 observations. The results show that the

Table 5. Summary of Multitrait-Multimethod Analyses

Correlation Size	Construct							
	Perceived Usefulness				Perceived Ease of Use			
	Same Trait/ Diff. Method		Different Trait		Same Trait/ Diff. Method		Different Trait	
	Elec. Mail	XEDIT	Same Meth.	Diff. Meth.	Elec. Mail	XEDIT	Same Meth.	Diff. Meth.
-.20 to -.11				6			1	1
-.10 to -.01			3	25	2		1	5
.00 to .09			2	27	2		5	32
.10 to .19			5	25	9		1	40
.20 to .29				7	14	2	2	11
.30 to .39					9			1
.40 to .49					3	9		
.50 to .59	4				3	11		
.60 to .69	14	4			3	13		
.70 to .79	20	11			3	8		
.80 to .89	7	26				2		
.90 to .99		4						
# Correlations	45	45	10	90	45	45	10	90

usefulness and ease of use items load on distinct factors (Table 6). The multitrait-multimethod analysis and factor analysis both support the construct validity of the 10-item scales.

Scale refinement

In applied testing situations, it is important to keep scales as brief as possible, particularly when multiple systems are going to be evaluated. The usefulness and ease of use scales were refined and streamlined based on results from Study 1 and then subjected to a second round of empirical validation in Study 2, reported below. Applying the Spearman-Brown prophecy formula to the .97 reliability obtained for perceived usefulness indicates that a six-item scale composed of items having comparable reliability would yield a scale reliability of .94. The five positive ease of use items had a reliability of .92. Taken together, these findings from Study 1 suggest that six items would be adequate to achieve reliability levels above .9 while maintaining adequate validity levels. Based on the results of the field study, six of the 10 items for each construct were selected to form modified scales.

For the ease of use scale, the five negatively worded items were eliminated due to their apparent common method variance, leaving items 2, 4, 6, 8 and 10. Item 6 ("easy to remember

how to perform tasks"), which the pretest indicated was concerned with ease of learning, was replaced by a reversal of item 9 ("easy to become skillful"), which was specifically designed to more directly tap ease of learning. These items include two from cluster C, one each from clusters A and B, and the overall item. (See Table 4.) In order to improve representative coverage of the content domain, an additional A item was added. Of the two remaining A items (#1, Cumbersome, and #5, Rigid and Inflexible), item 5 is readily reversed to form "flexible to interact with." This item was added to form the sixth item, and the order of items 5 and 8 was permuted in order to prevent items from the same cluster (items 4 and 5) from appearing next to one another.

In order to select six items to be used for the usefulness scale, an item analysis was performed. Corrected item-total correlations were computed for each item, separately for each system studied. Average Z-scores of these correlations were used to rank the items. Items 3, 5, 6, 8, 9 and 10 were top-ranked items. Referring to the cluster analysis (Table 3), we see that this set is well-representative of the content domain, including two items from cluster A, two from cluster B and one from cluster C, as well as the overall item (#10). The items were permuted to prevent items from the same cluster from appearing next to one another. The result-

Table 6. Factor Analysis of Perceived Usefulness and Ease of Use Questions: Study 1

Scale Items	Factor 1 (Usefulness)	Factor 1 (Ease of Use)
Usefulness		
1 Quality of Work	.80	.10
2 Control over Work	.86	-.03
3 Work More Quickly	.79	.17
4 Critical to My Job	.87	-.11
5 Increase Productivity	.87	.10
6 Job Performance	.93	-.07
7 Accomplish More Work	.91	-.02
8 Effectiveness	.96	-.03
9 Makes Job Easier	.80	.16
10 Useful	.74	.23
Ease of Use		
1 Cubersome	.00	.73
2 Ease of Learning	.08	.60
3 Frustrating	.02	.65
4 Controllable	.13	.74
5 Rigid & Inflexible	.09	.54
6 Ease of Remembering	.17	.62
7 Mental Effort	-.07	.76
8 Understandable	.29	.64
9 Effort to Be Skillful	-.25	.88
10 Easy to Use	.23	.72

ing six-item usefulness and ease of use scales are shown in the Appendix.

Relationship to use

Participants were asked to self-report their degree of current usage of electronic mail and XEDIT on six-position categorical scales with boxes labeled "Don't use at all," "Use less than once each week," "Use about once each week," "Use several times a week," "Use about once each day," and "Use several times each day." Usage was significantly correlated with both perceived usefulness and perceived ease of use for both PROFS mail and XEDIT. PROFS mail usage correlated .56 with perceived usefulness and .32 with perceived ease of use. XEDIT usage correlated .68 with usefulness and .48 with ease of use. When data were pooled across systems, usage correlated .63 with usefulness and .45 with ease of use. The overall usefulness-use correlation was significantly greater than the ease of use-use correlation as indicated by a test of dependent correlations ($t_{181} = 3.69$, $p < .001$) (Cohen and Cohen, 1975). Usefulness and ease of use were significantly correlated with each other for electronic mail (.56), XEDIT

(.69), and overall (.64). All correlations were significant at the .001 level.

Regression analyses were performed to assess the joint effects of usefulness and ease of use on usage. The effect of usefulness on usage, controlling for ease of use, was significant at the .001 level for electronic mail ($b = .55$), XEDIT ($b = .69$), and pooled ($b = .57$). In contrast, the effect of ease of use on usage, controlling for usefulness, was non-significant across the board ($b = .01$ for electronic mail; $b = .02$ for XEDIT; and $b = .07$ pooled). In other words, the significant pairwise correlation between ease of use and usage vanishes when usefulness is controlled for. The regression coefficients obtained for each individual system within each study were not significantly different ($F_{3, 178} = 1.95$, n.s.). As the relationship between independent variables in a regression approach perfect linear dependence, multicollinearity can degrade the parameter estimates obtained. Although the correlations between usefulness and ease of use are significant, according to tests for multicollinearity they are not large enough to compromise the accuracy of the estimated regression coefficients since the standard errors of the estimates are low (.08 for both usefulness and

ease of use), and the covariances between the parameter estimates are negligible ($-.004$) (Johnston, 1972; Mansfield and Helms, 1982). Based on partial correlation analyses, the variance in usage explained by ease of use drops by 98% when usefulness is controlled for. The regression and partial correlation results suggest that usefulness mediates the effect of ease of use on usage, i.e., that ease of use influences usage indirectly through its effect on usefulness (J.A. Davis, 1985).

Study 2

A lab study was performed to evaluate the six-item usefulness and ease of use scales resulting from scale refinement in Study 1. Study 2 was designed to approximate applied prototype testing or system selection situations, an important class of situations where measures of this kind are likely to be used in practice. In prototype testing and system selection contexts, prospective users are typically given a brief hands-on demonstration involving less than an hour of actually interacting with the candidate system. Thus, representative users are asked to rate the future usefulness and ease of use they would expect based on relatively little experience with the systems being rated. We are especially interested in the properties of the usefulness and ease of use scales when they are worded in a prospective sense and are based on limited experience with the target systems. Favorable psychometric properties under these circumstances would be encouraging relative to their use as early warning indicants of user acceptance (Ginzberg, 1981).

The lab study involved 40 voluntary participants who were evening MBA students at Boston University. They were paid \$25 for participating in the study. They had an average of five years' work experience and were employed full-time in several industries, including education (10 percent), government (10 percent), financial (28 percent), health (18 percent), and manufacturing (8 percent). They had a range of prior experience with computers in general (35 percent none or limited; 48 percent moderate; and 17 percent extensive) and personal computers in particular (35 percent none or limited; 48 percent moderate; and 15 percent extensive) but were unfamiliar with the two systems used in the study.

The study involved evaluating two IBM PC-based graphics systems: Chart-Master (by Decision Resources, Inc. of Westport, CN) and Pendraw (by Pencept, Inc. of Waltham, MA). Chart-Master is a menu-driven package that creates numerical business graphs, such as bar charts, line charts, and pie charts based on parameters defined by the user. Through the keyboard and menus, the user inputs the data for, and defines the desired characteristics of, the chart to be made. The user can specify a wide variety of options relating to title fonts, colors, plot orientation, cross-hatching pattern, chart format, and so on. The chart can then be previewed on the screen, saved, and printed. Chart-Master is a successful commercial product that typifies the category of numeric business charting programs.

Pendraw is quite different from the typical business charting program. It uses bit-mapped graphics and a "direct manipulation" interface where users draw desired shapes using a digitizer tablet and an electronic "pen" as a stylus. The digitizer tablet supplants the keyboard as the input medium. By drawing on a tablet, the user manipulates the image, which is visible on the screen as it is being created. Pendraw offers capabilities typical of PC-based, bit-mapped "paint" programs (see Panko, 1988), allowing the user to perform freehand drawing and select from among geometric shapes, such as boxes, lines, and circles. A variety of line widths, color selections and title fonts are available. The digitizer is also capable of performing character recognition, converting hand-printer characters into various fonts (Ward and Blesser, 1985). Pencept had positioned the Pendraw product to compete with business charting programs. The manual introduces Pendraw by guiding the user through the process of creating a numeric bar chart. Thus, a key marketing issue was the extent to which the new product would compete favorably with established brands, such as Chart-Master.

Participants were given one hour of hands-on experience with Chart-Master and Pendraw, using workbooks that were designed to follow the same instructional sequence as the user manuals for the two products, while equalizing the style of writing and eliminating value statements (e.g., "See how easy that was to do?"). Half of the participants tried Chart-Master first and half tried Pendraw first. After using each package, a questionnaire was completed.

Reliability and validity

Cronbach alpha was .98 for perceived usefulness and .94 for perceived ease of use. Convergent validity was supported, with only two of 72 monotrait-heteromethod correlations falling below significance. Ease of use item 4 (flexibility), applied to Chart-Master, was not significantly correlated with either items 3 (clear and understandable) or 5 (easy to become skillful). This suggests that, contrary to conventional wisdom, flexibility is not always associated with ease of use. As Goodwin (1987) points out, flexibility can actually impair ease of use, particularly for novice users. With item 4 omitted, Cronbach alpha for ease of use would increase from .94 to .95. Despite the two departures to convergent validity related to ease of use item 4, no exceptions to the discriminant validity criteria occurred across a total of 720 comparisons (360 for each scale).

Factorial validity was assessed by factor analyzing the 12 scale items using principal components extraction and oblique rotation. The resulting two-factor solution is very consistent with distinct, unidimensional usefulness and each of use scales (Table 7). Thus, as in Study 1, Study 2 reflects favorably on the convergent, discriminant, and factorial validity of the usefulness and ease of use scales.

Relationship to use

Participants were asked to self-predict their future use of Chart-Master and Pendraw. The

questions were worded as follows: "Assuming Pendraw would be available on my job, I predict that I will use it on a regular basis in the future," followed by two seven-point scales, one with likely-unlikely end-point adjectives, the other, reversed in polarity, with improbable-probable end-point adjectives. Such self-predictions, or "behavioral expectations," are among the most accurate predictors available for an individual's future behavior (Sheppard, et al., 1988; Warshaw and Davis, 1985). For Chart-Master, usefulness was significantly correlated with self-predicted usage ($r = .71, p < .001$), but ease of use was not ($r = .25, n.s.$) (Table 8). Chart-Master had a non-significant correlation between ease of use and usefulness ($r = .25, n.s.$). For Pendraw, usage was significantly correlated with both usefulness ($r = .59, p < .001$) and ease of use ($r = .47, p < .001$). The ease of use-usefulness correlation was significant for Pendraw ($r = .38, p < .001$). When data were pooled across systems, usage correlated .85 ($p < .001$) with usefulness and .59 ($p < .001$) with ease of use (see Table 8). Ease of use correlated with usefulness .56 ($p < .001$). The overall usefulness-use correlation was significantly greater than the ease of use-use correlation, as indicated by a test of dependent correlations ($t_{77} = 4.78, p < .001$) (Cohen and Cohen, 1975).

Regression analyses (Table 9) indicate that the effect of usefulness on usage, controlling for ease of use, was significant at the .001 level for Chart-Master ($b = .69$), Pendraw ($b = .76$) and overall ($b = .75$). In contrast, the effect of ease of use on usage, controlling for usefulness, was

Table 7. Factor Analysis of Perceived Usefulness and Ease of Use Items: Study 2

Scale Items	Factor 1 (Usefulness)	Factor 2 (Ease of Use)
Usefulness		
1 Work More Quickly	.91	.01
2 Job Performance	.98	-.03
3 Increase Productivity	.98	-.03
4 Effectiveness	.94	.04
5 Makes Job Easier	.95	-.01
6 Useful	.88	.11
Ease of Use		
1 Easy to Learn	-.20	.97
2 Controllable	.19	.83
3 Clear & Understandable	-.04	.89
4 Flexible	.13	.63
5 Easy to Become Skillful	.07	.91
6 Easy to Use	.09	.91

Table 8. Correlations Between Perceived Usefulness, Perceived Ease of Use, and Self-Reported System Usage

	Correlation		
	Usefulness & Usage	Ease of Use & Usage	Ease of Use & Usefulness
Study 1			
Electronic Mail (n = 109)	.56***	.32***	.56***
XEDIT (n = 75)	.68***	.48***	.69***
Pooled (n = 184)	.63***	.45***	.64***
Study 2			
Chart-Master (n = 40)	.71***	.25	.25
Pendraw (n = 40)	.59***	.47***	.38**
Pooled (n = 80)	.85***	.59***	.56***
Davis, et al. (1989) (n = 107)			
Wave 1	.65***	.27**	.10
Wave 2	.70***	.12	.23**

*** p<.001

** p<.01

* p<.05

Table 9. Regression Analyses of the Effect of Perceived Usefulness and Perceived Ease of Use on Self-Reported Usage

	Independent Variables		R ²
	Usefulness	Ease of Use	
Study 1			
Electronic Mail (n = 109)	.55***	.01	.31
XEDIT (n = 75)	.69***	.02	.46
Pooled (n = 184)	.57***	.07	.38
Study 2			
Chart-Master (n = 40)	.69***	.08	.51
Pendraw (n = 40)	.76***	.17	.71
Pooled (n = 80)	.75***	.17*	.74
Davis, et al. (1989) (n = 107)			
After 1 Hour	.62***	.20***	.45
After 14 Weeks	.71***	-.06	.49

*** p<.001

** p<.01

* p<.05

non-significant for both Chart-Master ($b = .08$, n.s.) and Pendraw ($b = .17$, n.s.) when analyzed separately and borderline significant when observations were pooled ($b = .17$, $p < .05$). The regression coefficients obtained for Pendraw and Chart-Master were not significantly different ($F_{3,74} = .014$, n.s.). Multicollinearity is ruled out since the standard errors of the estimates are low (.07 for both usefulness and ease of use) and the covariances between the parameter estimates are negligible ($-.004$).

Hence, as in Study 1, the significant pairwise correlations between ease of use and usage drop dramatically when usefulness is controlled for, suggesting that ease of use operates

through usefulness. Partial correlation analysis indicates that the variance in usage explained by ease of use drops by 91% when usefulness is controlled for. Consistent with Study 1, these regression and partial correlation results suggest that usefulness mediates the effect of ease of use on usage. The implications of this are addressed in the following discussion.

Discussion

The purpose of this investigation was to develop and validate new measurement scales for perceived usefulness and perceived ease of use, two distinct variables hypothesized to be deter-

minants of computer usage. This effort was successful in several respects. The new scales were found to have strong psychometric properties and to exhibit significant empirical relationships with self-reported measures of usage behavior. Also, several new insights were generated about the nature of perceived usefulness and ease of use, and their roles as determinants of user acceptance.

The new scales were developed, refined, and streamlined in a several-step process. Explicit definitions were stated, followed by a theoretical analysis from a variety of perspectives, including: expectancy theory; self-efficacy theory; behavioral decision theory; diffusion of innovations; marketing; and human-computer interaction, regarding why usefulness and ease of use are hypothesized as important determinants of system use. Based on the stated definitions, initial scale items were generated. To enhance content validity, these were pretested in a small pilot study, and several items were eliminated. The remaining items, 10 for each of the two constructs, were tested for validity and reliability in Study 1, a field study of 112 users and two systems (the PROFS electronic mail system and the XEDIT file editor). Item analysis was performed to eliminate more items and refine others, further streamlining and purifying the scales. The resulting six-item scales were subjected to further construct validation in Study 2, a lab study of 40 users and two systems: Chart-Master (a menu-driven business charting program) and Pendraw (a bit-mapped paint program with a digitizer tablet as its input device).

The new scales exhibited excellent psychometric characteristics. Convergent and discriminant validity were strongly supported by multitrait-multimethod analyses in both validation studies. These two data sets also provided strong support for factorial validity: the pattern of factor loadings confirmed that *a priori* structure of the two instruments, with usefulness items loading highly on one factor, ease of use items loading highly on the other factor, and small cross-factor loadings. Cronbach alpha reliability for perceived usefulness was .97 in Study 1 and .98 in Study 2. Reliability for ease of use was .91 in Study 1 and .94 in Study 2. These findings mutually confirm the psychometric strength of the new measurement scales.

As theorized, both perceived usefulness and ease of use were significantly correlated with self-reported indicants of system use. Perceived use-

fulness was correlated .63 with self-reported current use in Study 1 and .85 with self-predicted use in Study 2. Perceived ease of use was correlated .45 with use in Study 1 and .69 in Study 2. The same pattern of correlations is found when correlations are calculated separately for each of the two systems in each study (Table 8). These correlations, especially the usefulness-use link, compare favorably with other correlations between subjective measures and self-reported use found in the MIS literature. Swanson's (1987) "value" dimension correlated .20 with use, while his "accessibility" dimension correlated .13 with self-reported use. Correlations between "user information satisfaction" and self-reported use of .39 (Barki and Huff, 1985) and .28 (Baroudi, et al., 1986) have been reported. "Realism of expectations" has been found to be correlated .22 with objectively measured use (Ginzberg, 1981) and .43 with self-reported use (Barki and Huff, 1985). "Motivational force" was correlated .25 with system use, objectively measured (DeSanctis, 1983). Among the usage correlations reported in the literature, the .79 correlation between "performance" and use reported by Robey (1979) stands out. Recall that Robey's expectancy model was a key underpinning for the definition of perceived usefulness stated in this article.

One of the most significant findings is the relative strength of the usefulness-usage relationship compared to the ease of use-usage relationship. In both studies, usefulness was significantly more strongly linked to usage than was ease of use. Examining the joint direct effect of the two variables on use in regression analyses, this difference was even more pronounced: the usefulness-usage relationship remained large, while the ease of use-usage relationship was diminished substantially (Table 8). Multicollinearity has been ruled out as an explanation for the results using specific tests for the presence of multicollinearity. In hindsight, the prominence of perceived usefulness makes sense conceptually: users are driven to adopt an application primarily because of the functions it performs for them, and secondarily for how easy or hard it is to get the system to perform those functions. For instance, users are often willing to cope with some difficulty of use in a system that provides critically needed functionality. Although difficulty of use can discourage adoption of an otherwise useful system, no amount of ease of use can compensate for a

system that does not perform a useful function. The prominence of usefulness over ease of use has important implications for designers, particularly in the human factors tradition, who have tended to overemphasize ease of use and overlook usefulness (e.g., Branscomb and Thomas, 1984; Chin, et al., 1988; Shneiderman, 1987). Thus, a major conclusion of this study is that perceived usefulness is a strong correlate of user acceptance and should not be ignored by those attempting to design or implement successful systems.

From a causal perspective, the regression results suggest that ease of use may be an antecedent to usefulness, rather than a parallel, direct determinant of usage. The significant pairwise correlation between ease of use and usage all but vanishes when usefulness is controlled for. This, coupled with a significant ease of use-usefulness correlation is exactly the pattern one would expect if usefulness mediated between ease of use and usage (e.g., J.A. Davis, 1985). That is, the results are consistent with an ease of use --> usefulness --> usage chain of causality. These results held both for pooled observations and for each individual system (Table 8). The causal influence of ease of use on usefulness makes sense conceptually, too. All else being equal, the easier a system is to interact with, the less effort needed to operate it, and the more effort one can allocate to other activities (Radner and Rothschild, 1975), contributing to overall job performance. Goodwin (1987) also argues for this flow of causality, concluding from her analysis that: "There is increasing evidence that the effective functionality of a system depends on its usability" (p. 229). This intriguing interpretation is preliminary and should be subjected to further experimentation. If true, however, it underscores the theoretical importance of perceived usefulness.

This investigation has limitations that should be pointed out. The generality of the findings remains to be shown by future research. The fact that similar findings were observed, with respect to both the psychometric properties of the measures and the pattern of empirical associations, across two different user populations, two different systems, and two different research settings (lab and field), provides some evidence favoring external validity.

In addition, a follow-up to this study, reported by Davis, et al. (1989) found a very similar pat-

tern of results in a two-wave study (Tables 8 and 9). In that study, MBA student subjects were asked to fill out a questionnaire after a one-hour introduction to a word processing program, and again 14 weeks later. Usage intentions were measured at both time periods, and self-reported usage was measured at the later time period. Intentions were significantly correlated with usage (.35 and .63 for the two points in time, respectively). Unlike the results of Studies 1 and 2, Davis, et al. (1989) found a significant direct effect of ease of use on usage, controlling for usefulness, after the one-hour training session (Table 9), although this evolved into a non-significant effect as of 14 weeks later. In general, though, Davis, et al. (1989) found usefulness to be more influential than ease of use in driving usage behavior, consistent with the findings reported above.

Further research will shed more light on the generality of these findings. Another limitation is that the usage measures employed were self-reported as opposed to objectively measured. Not enough is currently known about how accurately self-reports reflect actual behavior. Also, since usage was reported on the same questionnaire used to measure usefulness and ease of use, the possibility of a halo effect should not be overlooked. Future research addressing the relationship between these constructs and objectively measured use is needed before claims about the behavioral predictiveness can be made conclusively. These limitations notwithstanding, the results represent a promising step toward the establishment of improved measures for two important variables.

Research implications

Future research is needed to address how other variables relate to usefulness, ease of use, and acceptance. Intrinsic motivation, for example, has received inadequate attention in MIS theories. Whereas perceived usefulness is concerned with performance as a consequence use, intrinsic motivation is concerned with the reinforcement and enjoyment related to the process of performing a behavior per se, irrespective of whatever external outcomes are generated by such behavior (Deci, 1975). Although intrinsic motivation has been studied in the design of computer games (e.g., Malone, 1981), it is just beginning to be recognized as a potential mechanism underlying user acceptance of end-user

systems (Carroll and Thomas, 1988). Currently, the role of affective attitudes is also an open issue. While some theorists argue that beliefs influence behavior only via their indirect influence on attitudes (e.g., Fishbein and Ajzen, 1975), others view beliefs and attitudes as co-determinants of behavioral intentions (e.g., Triandis, 1977), and still others view attitudes as antecedents of beliefs (e.g., Weiner, 1986). Counter to Fishbein and Ajzen's (1975) position, both Davis (1986) and Davis, et al. (1989) found that attitudes do not fully mediate the effect of perceived usefulness and perceived ease of use on behavior.

It should be emphasized that perceived usefulness and ease of use are people's subjective appraisal of performance and effort, respectively, and do not necessarily reflect objective reality. In this study, beliefs are seen as meaningful variables in their own right, which function as behavioral determinants, and are not regarded as surrogate measures of objective phenomena (as is often done in MIS research, e.g., Ives, et al., 1983; Srinivasan, 1985). Several MIS studies have observed discrepancies between perceived and actual performance (Cats-Baril and Huber, 1987; Dickson, et al., 1986; Gallupe and DeSanctis, 1988; McIntyre, 1982; Sharda, et al., 1988). Thus, even if an application would objectively improve performance, if users don't perceive it as useful, they're unlikely to use it (Alavi and Henderson, 1981). Conversely, people may overrate the performance gains a system has to offer and adopt systems that are dysfunctional. Given that this study indicates that people act according to their beliefs about performance, future research is needed to understand why performance beliefs are often in disagreement with objective reality. The possibility of dysfunctional impacts generated by information technology (e.g., Kottemann and Remus, 1987) emphasizes that user acceptance is not a universal goal and is actually undesirable in cases where systems fail to provide true performance gains.

More research is needed to understand how measures such as those introduced here perform in applied design and evaluation settings. The growing literature on design principles (Anderson and Olson, 1985; Gould and Lewis, 1985; Johansen and Baker, 1984; Mantei and Teorey, 1988; Shneiderman, 1987) calls for the use of subjective measures at various points throughout the development and implementation process, from the earliest needs assessment

through concept screening and prototype testing to post-implementation assessment. The fact that the measures performed well psychometrically both after brief introductions to the target system (Study 2, and Davis, et al., 1989) and after substantial user experience with the system (Study 1, and Davis, et al., 1989) is promising concerning their appropriateness at various points in the life cycle. Practitioners generally evaluate systems not only to predict acceptability but also to diagnose the reasons underlying lack of acceptance and to formulate interventions to improve user acceptance. In this sense, research on how usefulness and ease of use can be influenced by various externally controllable factors, such as the functional and interface characteristics of the system (Benbasat and Dexter, 1986; Bewley, et al., 1983; Dickson, et al., 1986), development methodologies (Alavi, 1984), training and education (Nelson and Cheney, 1987), and user involvement in design (Baroudi, et al. 1986; Franz and Robey, 1986) is important. The new measures introduced here can be used by researchers investigating these issues.

Although there has been a growing pessimism in the field about the ability to identify measures that are robustly linked to user acceptance, the view taken here is much more optimistic. User reactions to computers are complex and multifaceted. But if the field continues to systematically investigate fundamental mechanisms driving user behavior, cultivating better and better measures and critically examining alternative theoretical models, sustainable progress is within reach.

Acknowledgements

This research was supported by grants from the MIT Sloan School of Management, IBM Canada Ltd., and The University of Michigan Business School. The author is indebted to the anonymous associate editor and reviewers for their many helpful suggestions.

References

- Abelson, R.P. and Levi, A. "Decision Making and Decision Theory," in *The Handbook of Social Psychology*, third edition, G. Lindsay and E. Aronson (eds.), Knopf, New York, NY, 1985, pp. 231-309.

- Adelbratt, T. and Montgomery, H. "Attractiveness of Decision Rules," *Acta Psychologica* (45), 1980, pp. 177-185.
- Alavi, M. "An Analysis of the Prototyping Approach to Information Systems Development," *Communications of the ACM* (27:6), June 1984, pp. 556-563.
- Alavi, M. and Henderson, J.C. "An Evolutionary Strategy for Implementing a Decision Support System," *Management Science* (27:11), November 1981, pp. 1309-1323.
- Anastasi, A. "Evolving Concepts of Test Validation," *Annual Review of Psychology* (37), 1986, pp. 1-15.
- Anderson, N.S. and Olson, J.R. (eds.) *Methods for Designing Software to Fit Human Needs and Capabilities: Proceedings of the Workshop on Software Human Factors*, National Academy Press, Washington, D.C., 1985.
- Bandura, A. "Self-Efficacy Mechanism in Human Agency," *American Psychologist* (37:2), February 1982, pp. 122-147.
- Barki, H. and Huff, S. "Change, Attitude to Change, and Decision Support System Success," *Information and Management* (9:5), December 1985, pp. 261-268.
- Baroudi, J.J., Olson, M.H. and Ives, B. "An Empirical Study of the Impact of User Involvement on System Usage and Information Satisfaction," *Communications of the ACM* (29:3), March 1986, pp. 232-238.
- Beach, L.R. and Mitchell, T.R. "A Contingency Model for the Selection of Decision Strategies," *Academy of Management Review* (3:3), July 1978, pp. 439-449.
- Benbasat, I. and Dexter, A.S. "An Investigation of the Effectiveness of Color and Graphical Presentation Under Varying Time Constraints," *MIS Quarterly* (10:1), March 1986, pp. 59-84.
- Bewley, W.L., Roberts, T.L., Schoit, D. and Verplank, W.L., "Human Factors Testing in the Design of Xerox's 8010 'Star' Office Workstation," *CHI '83 Human Factors in Computing Systems*, Boston, December 12-15, 1983, ACM, New York, NY, pp. 72-77.
- Bohmstedt, G.W. "Reliability and Validity Assessment in Attitude Measurement," in *Attitude Measurement*, G.F. Summers (ed.), Rand McNally, Chicago, IL, 1970, pp. 80-99.
- Bowen, W. "The Puny Payoff from Office Computers," *Fortune*, May 26, 1986, pp. 20-24.
- Branscomb, L.M. and Thomas, J.C. "Ease of Use: A System Design Challenge," *IBM Systems Journal* (23), 1984, pp. 224-235.
- Campbell, D.T. and Fiske, D.W. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," *Psychological Bulletin* (56:9), March 1959, pp. 81-105.
- Campbell, D.T., Siegman, C.R. and Rees, M.B. "Direction-of-Wording Effects in the Relationships Between Scales," *Psychological Bulletin* (68:5), November 1967, pp. 293-303.
- Card, S.K., Moran, T.P. and Newell, A. *The Psychology of Human-Computer Interaction*, Erlbaum, Hillsdale, NJ, 1984.
- Carroll, J.M. and Carrithers, C. "Training Wheels in a User Interface," *Communications of the ACM* (27:8), August 1984, pp. 800-806.
- Carroll, J.M. and McKendree, J. "Interface Design Issues for Advice-Giving Expert Systems," *Communications of the ACM* (30:1), January 1987, pp. 14-31.
- Carroll, J.M., Mack, R.L., Lewis, C.H., Grishkowsky, N.L. and Robertson, S.R. "Exploring Exploring a Word Processor," *Human-Computer Interaction* (1), 1985, pp. 283-307.
- Carroll, J.M. and Thomas, J.C. "Fun," *SIGCHI Bulletin* (19:3), January 1988, pp. 21-24.
- Cats-Baril, W.L. and Huber, G.P. "Decision Support Systems for Ill-Structured Problems: An Empirical Study," *Decision Sciences* (18:3), Summer 1987, pp. 352-372.
- Cheney, P.H., Mann, R.I. and Amoroso, D.L. "Organizational Factors Affecting the Success of End-User Computing," *Journal of Management Information Systems* (3:1), Summer 1986, pp. 65-80.
- Chin, J.P., Diehl, V.A. and Norman, K.L. "Development of an Instrument for Measuring User Satisfaction of the Human-Computer Interface," *CHI'88 Human Factors in Computing Systems*, Washington, D.C., May 15-19, 1988, ACM, New York, NY, pp. 213-218.
- Cohen, J. and Cohen, P. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, Erlbaum, Hillsdale, NJ, 1975.
- Curley, K.F. "Are There any Real Benefits from Office Automation?" *Business Horizons* (4), July-August 1984, pp. 37-42.
- Davis, F.D. "A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results," doctoral dissertation, MIT Sloan School of Management, Cambridge, MA, 1986.
- Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," *Management Science* (35:8), August 1989, pp. 982-1003.
- Davis, J.A. *The Logic of Causal Order*, Sage, Beverly Hills, CA, 1985.
- Deci, E.L. *Intrinsic Motivation*, Plenum, New

- York, NY, 1975.
- DeSanctis, G. "Expectancy Theory as an Explanation of Voluntary Use of a Decision Support System," *Psychological Reports* (52), 1983, pp. 247-260.
- Dickson, G.W., DeSanctis, G. and McBride, D.J. "Understanding the Effectiveness of Computer Graphics for Decision Support: A Cumulative Experimental Approach," *Communications of the ACM* (29:1), January 1986, pp. 40-47.
- Edelmann, F. "Managers, Computer Systems, and Productivity," *MIS Quarterly* (5:3), September 1981, pp. 1-19.
- Fishbein, M. and Ajzen, I. "*Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*," Addison-Wesley, Reading, MA 1975.
- Franz, C.R. and Robey, D. "Organizational Context, User Involvement, and the Usefulness of Information Systems," *Decision Sciences* (17:3), Summer 1986, pp. 329-356.
- Gallupe, R.B., DeSanctis, G. and Dickson, G.W. "Computer-Based Support for Group Problem Finding: An Empirical Investigation," *MIS Quarterly* (12:2), June 1988, pp. 277-296.
- Ginzberg, M.J. "Early Diagnosis of MIS Implementation Failure: Promising Results and Unanswered Questions," *Management Science* (27:4), April 1981, pp. 459-478.
- Good, M., Spine, T.M., Whiteside, J. and George P. "User-Derived Impact Analysis as a Tool for Usability Engineering," *CHI'86 Human Factors in Computing Systems*, Boston, April 13-17, 1986, ACM, New York, New York pp. 241-246.
- Goodwin, N.C. "Functionality and Usability," *Communications of the ACM* (30:3), March 1987, pp. 229-233.
- Goslar, M.D. "Capability Criteria for Marketing Decision Support Systems," *Journal of Management Information Systems* (3:1), Summer 1986, pp. 81-95.
- Gould, J., Conti, J. and Hovanyecz, T. "Composing letters with a Simulated Listening Typewriter," *Communications of the ACM* (26:4), April 1983, pp. 295-308.
- Gould, J.D. and Lewis C. "Designing for Usability: Key Principles and What Designers Think," *Communications of the ACM* (28:3), March 1985, pp. 300-311.
- Greenberg, K. "Executives Rate Their PCs," *PC World*, September 1984, pp. 286-292.
- Hauser, J.R. and Simmie, P. "Profit Maximizing Perceptual Positions: An Integrated Theory for the Selection of Product Features and Price," *Management Science* (27:1), January 1981, pp. 33-56.
- Hill, T., Smith, N.D., and Mann, M.F. "Role of Efficacy Expectations in Predicting the Decision to Use Advanced Technologies: The Case of Computers," *Journal of Applied Psychology*, (72:2), May 1987, pp. 307-313.
- Ives, B., Olson, M.H. and Baroudi, J.J. "The measurement of User Information Satisfaction," *Communications of the ACM* (26:10), October 1983, pp. 785-793.
- Jarvenpaa, S.L. "The Effect of Task Demands and Graphical Format on Information Processing Strategies," *Management Science* (35:3), March 1989, pp. 285-303.
- Johansen, R. & Baker E., "User Needs Workshops: A New Approach to Anticipating User Needs for Advanced Office Systems," *Office Technology and People* (2), 1984, pp. 103-119.
- Johnson, E.J. and Payne, J.W. "Effort and Accuracy in Choice," *Management Science* (31:4), April 1985, pp. 395-414.
- Johnston, J. *Econometric Methods*, McGraw-Hill, New York, NY, 1972.
- Klein, G. and Beck, P.O. "A Decision Aid for Selecting Among Information Systems Alternatives," *MIS Quarterly* (11:2), June 1987, pp. 177-186.
- Kleinmuntz, D.N. and Schkade, D.A. "The Cognitive Implications of Information Displays in Computer-Supported Decision-Making," University of Texas at Austin, Graduate School of Business, Department of Management Working Paper 87/88-4-8, 1988.
- Kottemann, J.E. and Remus, W.E. "Evidence and Principles of Functional and Dysfunctional DSS," *OMEGA* (15:2), March 1987, pp. 135-143.
- Larcker, D.F. and Lessig, V.P. "Perceived Usefulness of Information: A Psychometric Examination," *Decision Sciences* (11:1), January 1980, pp. 121-134.
- Lucas, H.C. "Performance and the Use of an Information System," *Management Science* (21:8), April 1975, pp. 908-919.
- Malone, T.W. "Toward a Theory of Intrinsically Motivating Instruction," *Cognitive Science* (4), 1981, pp. 333-369.
- Mansfield, E.R. and Helms, B.P. "Detecting Multicollinearity," *The American Statistician* (36:3), August 1982, pp. 158-160.
- Mantei, M.M. and Teorey, T.J. "Cost/Benefit Analysis for Incorporating Human Factors in the Software Lifecycle," *Communications of the ACM* (31:4), April 1988, pp. 428-439.

- Markus, M.L. and Bjorn-Anderson, N. "Power Over Users: It's Exercise by System Professionals," *Communications of the ACM* (30:6), June 1987, pp. 498-504.
- McIntyre, S. "An Experimental Study of the Impact of Judgement-Based Marketing Models," *Management Science* (28:1), January 1982, pp. 17-23.
- Nelson, R.R. and Cheney, P.H. "Training End Users: An Exploratory Study," *MIS Quarterly* (11:4), December 1987, pp. 547-559.
- Nickerson, R.S. "Why Interactive Computer Systems Are Sometimes Not Used by People Who Might Benefit from Them," *International Journal of Man-Machine Studies* (15), 1981, pp. 469-483.
- Norman, D.A. "Design Principles for Human-Computer Interfaces," *CHI '83 Human Factors in Computing Systems*, Boston, December 12-15, 1983, ACM, New York, NY, pp. 1-10.
- Nunnally, J. *Psychometric Theory*, McGraw-Hill, New York, NY, 1978.
- Panko, R.R. *End-User Computing: Management, Applications, and Technology*, Wiley, New York, NY, 1988.
- Payne, J. W. "Contingent Decision Behavior," *Psychological Bulletin*, (92:2), 1982, pp. 382-402.
- Pfeffer, J. *Organizations and Organization Theory*, Pitman, Boston, MA, 1982.
- Radner, R. and Rothschild, M. "On the Allocation of Effort," *Journal of Economic Theory* (10), 1975, pp. 358-376.
- Roberts, T.L. and Moran, T.P. "The Evaluation of Text Editors: Methodology and Empirical Results," *Communications of the ACM* (26:4), April 1983, pp. 265-283.
- Robey, D. "User Attitudes and Management Information System Use," *Academy of Management Journal* (22:3), September 1979, pp. 527-538.
- Robey, D. and Farrow, D. "User Involvement in Information System Development: A Conflict Model and Empirical Test," *Management Science* (28:1), January 1982, pp. 73-85.
- Rogers, E.M. and Shoemaker, F.F. *Communication of Innovations: A Cross-Cultural Approach*, Free Press, New York, NY, 1971.
- Rushinek, A. and Rushinek, S.F. "What Makes Users Happy?" *Communications of the ACM* (29:7), July 1986, pp. 594-598.
- Saracevic, T. "Relevance: A Review of and a Framework for the Thinking on the Notion in Information Science," *Journal of the American Society for Information Science*, November-December 1975, pp. 321-343.
- Schein, E.H. *Organizational Psychology*, third edition, Prentice-Hall, Englewood Cliffs, NJ, 1980.
- Schewe, C.D. "The Management Information System User: An Exploratory Behavioral Analysis," *Academy of Management Journal* (19:4), December 1976, pp. 577-590.
- Schultz, R.L. and Slevin, D.P. "Implementation and Organizational Validity: An Empirical Investigation," in *Implementing Operations Research/Management Science*, R.L. Schultz and D.P. Slevin (eds.), American Elsevier, New York, NY, 1975, pp. 153-182.
- Sharda, R., Barr, S.H., and McDonnell, J.C. "Decision Support System Effectiveness: A Review and Empirical Test," *Management Science* (34:2), February 1988, pp. 139-159.
- Sheppard, B.H., Hartwick, J. and Warshaw, P.R. "The Theory of Reasoned Action: A Meta-Analysis of Past Research with Recommendations for Modifications and Future Research," *Journal of Consumer Research* (15:3), December 1988, pp. 325-343.
- Sherif, M. and Sherif, C.W. "The Own Categories Approach in Attitude Research," in *Readings in Attitude Theory and Measurement*, M. Fishbein (ed.), Wiley, New York, NY, 1967, pp. 190-198.
- Shneiderman, B. *Designing the User Interface*, Addison-Wesley, Reading, MA, 1987.
- Silk, A.J. "Response Set and Measurement of Self-Designated Opinion Leadership," *Public Opinion Quarterly* (35), 1971, pp. 383-397.
- Srinivasan, A. "Alternative Measures of System Effectiveness: Associations and Implications," *MIS Quarterly* (9:3), September 1985, pp. 243-253.
- Swanson, E.B. "Management Information Systems: Appreciation and Involvement," *Management Science* (21:2), October 1974, pp. 178-188.
- Swanson, E.B. "Measuring User Attitudes in MIS Research: A Review," *OMEGA* (10:2), March 1982, pp. 157-165.
- Swanson, E.B. "Information Channel Disposition and Use," *Decision Sciences* (18:1), Winter 1987, pp. 131-145.
- Swanson, E.B. *Information System Implementation: Bridging the Gap Between Design and Utilization*, Irwin, Homewood, IL, 1988.
- Tornatzky, L.G. and Klein, K.J. "Innovation Characteristics and Innovation Adoption-Implementation: A Meta-Analysis of Findings," *IEEE Transactions on Engineering Management* (EM-29:1), February 1982, pp. 28-45.
- Triandis, H.C. *Interpersonal Behavior*, Brooks/

- Cole, Monterey, CA, 1977.
- Vertinsky, I., Barth, R.T. and Mitchell, V.F. "A Study of OR/MS Implementation as a Social Change Process," in *Implementing Operations Research/Management Science*, R.L. Schultz and D.P. Slevin (eds.), American Elsevier, New York, NY, 1975, pp. 253-272.
- Vroom, V.H. *Work and Motivation*, Wiley, New York, NY, 1964.
- Ward, J.R. and Blesser, B. "Interactive Recognition of Handprinter Characters for Computer Input," *IEEE Computer Graphics and Applications*, September 1985, pp. 24-37.
- Warshaw, P.R. and Davis, F.D. "Disentangling Behavioral Intention and Behavioral Expectation," *Journal of Experimental Social Psychology* (21), May 1985, pp. 213-228.
- Weiner, B. "Attribution, Emotion, and Action," in *Handbook of Motivation and Cognition*, R.M. Sorrentino and E.T. Higgins (eds.), Guilford, New York, NY, 1986, pp. 281-312.
- Whiteside, J., Jones, S., Levy, P.S. and Wixon, D. "User Performance With Command, Menu, and Iconic Interfaces," *CHI '85 Proceedings*, San Francisco, April 14-18, 1985, ACM, New York, NY, pp. 185-191.
- Wright, P. "Consumer Choice Strategies: Simplifying vs. Optimizing," *Journal of Marketing Research* (14:1), February 1975, pp. 429-433.
- Young, T.R. "The Lonely Micro," *Datamation* (30:4), April 1984, pp. 100-114.

About the Author

Fred D. Davis is assistant professor at the University of Michigan School of Business Administration. His doctoral research at the Sloan School of Management, MIT, dealt with predicting and explaining user acceptance of computer technology. His current research interests include computer support for decision making, motivational determinants of computer acceptance, intentions and expectations in human behavior, and biased attributions of the performance impacts of information technology.

Appendix

Final Measurement Scales for Perceived Usefulness and Perceived Ease of Use

Perceived Usefulness

Using CHART-MASTER in my job would enable me to accomplish tasks more quickly.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	

Using CHART-MASTER would improve my job performance.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	

Using CHART-MASTER in my job would increase my productivity.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	

Using CHART-MASTER would enhance my effectiveness on the job.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	

Using CHART-MASTER would make it easier to do my job.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	

I would find CHART-MASTER useful in my job.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	

Perceived Ease of Use

Learning to operate CHART-MASTER would be easy for me.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	

I would find it easy to get CHART-MASTER to do what I want it to do.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	

My interaction with CHART-MASTER would be clear and understandable.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	

I would find CHART-MASTER to be flexible to interact with.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	

It would be easy for me to become skillful at using CHART-MASTER.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	

I would find CHART-MASTER easy to use.

likely	_____	_____	_____	_____	_____	_____	_____	unlikely
	extremely	quite	slightly	neither	slightly	quite	extremely	