

From “Can They?” to “Will They?”:

Extending Usability to Accommodate Acceptance Predictions

Andrew Dillon
Indiana University
HCI Research Group
SLIS
Bloomington, IN 47405
(812) 855-3249(937)
adillon@indiana.edu

and

Michael Morris
Air Force Institute of Technology
Graduate School of Logistics and Acquisition Management
Wright-Patterson AFB, OH 45433
(937) 426-5290
mmorris@.afit.af.mil

This item is not the definitive copy. Please use the following citation when referencing this material: Dillon, A. and Morris, M. (1998) From "can they?" to "will they?": extending usability evaluation to address acceptance. AIS Conference Paper, Baltimore, August 1998.

Introduction: usability engineering

Within the human-computer interaction (HCI) community, there exists a long and rich research paradigm on “usability engineering (UE).” Within the usability engineering tradition, usability is operationally defined as the effectiveness, efficiency and satisfaction with which specified users can perform particular tasks in a given environment (see e.g., Shackel 1991, Nielsen, 1993). *Effectiveness* answers: can users perform their tasks? *Efficiency* means: what resources do users expend to achieve a given outcome (e.g., time, effort)? Finally, *satisfaction* measures assess how well users like the application. From this perspective, usability is contextually defined in operational terms that designers can see as targets to meet, for example:

"Users should be able to perform specified tasks with new tool after W minutes training, with X% effectiveness, at least Y% efficiency, and Z% greater satisfaction than with old interface"

where $W < \text{infinity}$, and $0 < [X, Y, Z] < 100$

The strengths of the usability engineering approach include:

1. The use of operationalised measures that are negotiated in context,
 2. The direct coupling of usability to tasks the tool must support,
 3. The capability of negotiated targets to fit into an iterative design process,
- and
4. The decoupling of the usability construct from interface features

Each of these strengths gives the approach value to the software industry where design practices require targets to be met and where the success of a new tool is determined contextually rather than in any absolute manner. Thus, the usability engineering paradigm has enjoyed a wide range of support from industry.

Nonetheless, there are associated weaknesses of this approach. Some of these weaknesses include:

1. Usability criteria are dynamic, not fixed,
2. Usability is thus contextually determined so what works in one context may not work in another and design practices must continually ground themselves in
work practices
3. Determining usability criteria requires considerable analytic skill,
4. Generalization beyond context is difficult,
5. Criteria do not determine re-design advice

While the approach advocated by usability engineers of deriving appropriate targets for design and testing to meet is useful, it is clear that usability does *not* fully determine actual system use (see Dillon and Morris 1996). Thus, it is possible that designers may produce a well engineered artifact that meets set criteria, but still fails to gain the acceptance of discretionary users. In other words, usability is a necessary but *insufficient* determinant of use.

So what does predict use?

Research indicates that use is determined by combined attributes of the individual, the situation or the technology. For example, Rogers (1995) posits 5 characteristics of adopted technology that determine its use including relative advantage, compatibility, ease of use, trialability, and observability of outcome. Individual attributes of acceptors have also been identified in Rogers' innovation diffusion (ID) research with "early adopters" being typically individuals who are wealthy, highly educated, and risk-accepting. While ID offers a rich description it seems to afford little predictive power that can be coupled with usability engineering to produce an applicable design framework (although clearly the ease of use attribute of ID is related to usability). Therefore, it seems that the usability engineering approach is at least likely to positively influence *one* of the major determinants of acceptance while the remaining constructs of ID are rather loosely defined. It is not clear that UE can be extended to embrace other components of ID theory. For example, the relationship between the effectiveness and efficiency measures of usability engineering may relate more to relative advantage or observability of outcome than they do to ease of use in ID terms.

A more powerful tool that has been proposed is Technology Acceptance Model (Davis et al., 1989) which explicitly seeks to model the dynamics of acceptance in individual users by tapping their reactions to the system early in the design process. Davis' research shows that TAM offers an R^2 (variance explained) in range of .5 for many common office automation applications--an impressive result when one considers that good predictive tests for human behaviors usually account for .25 (Dillon and Watson, 1996). Davis argues that usefulness is the most important predictor of use, explaining significantly more variance than ease of use ratings by users. However, research on TAM is typically based on a single time period when users are exposed to a ready-made system. Morris (1996) tested Netscape for discretionary use with 263 users, tested every 4 weeks. Users kept self-report usage logs and all data were analyzed for perceptual measures from TAM plus quality and amount of use variables. This study showed that the general TAM relationships hold but over time "quality of use" (whether a user was effective or efficient with a system in usability engineering terms) is a significant mediating variable which influences subsequent ratings of usefulness and ease of use. Ease of use was directly

linked to intention to use in this study suggesting a continuous cycle of attitude influencing behavior which then further shapes attitude, and so forth.

In design terms, it is not clear how well TAM predicts usage when prototypes are evaluated (the time when most scope for re-design exists). Furthermore, the TAM measures are self-ratings of usefulness and ease of use, and it is clear from studies of users over time that user ratings of an interface can change considerably with repeated exposure (Dillon, 1987) and may shift independently of the usability of the interface.

Difficulties with existing approaches

Combining the usability engineering approach with TAM is not as simple as joining the two perspectives together. Usability Engineering measures behavior within controlled region and establishes only that users can perform to a certain level with the system under test. This is useful but does not allow us to predict actual use in the real world. TAM predicts behavior for developed systems once users have the opportunity to test those systems. However, this approach offers little insight for design where feedback from users is needed to shape the emerging technology early in the process. It is possible that by the time designers are able to gather meaningful data on TAM to see whether users will or will not accept a system, it is too costly to make major alterations.

Obviously, both approaches have utility but they do not cleanly complement each other. The operational definitions of effectiveness, efficiency and satisfaction in UE are *not* equivalent to TAM's 'ease of use' construct. Indeed it is possible that measuring usability in the UE manner might produce findings that are contradicted by TAM since part of UE's definition of usability is more likely measured by usefulness in TAM. UE measures behavior of users with the system, while TAM measures affect—and unfortunately, the relationship between the two is complicated. What seems to be missing from the current literature in this area is a unified model of use that supports both the process of design early on and clarifies the relationship between usability and acceptability.

Introducing the P³ model

We propose that use of a given information technology is driven by the following:

Utility (a.k.a. functionality, capability): This refers to the technical capability of the tool to actually support tasks that the user wishes to perform. This can be established in most cases objectively by audit or inspection of the specification or working version.

Usability (a.k.a. operability, user performance with the tool): This is the classic behavioral measure employed HCI studies. Usability refers to the extent to which users can exploit the utility of the system. Thus, systems with equivalent utility may result in different levels of usability depending on how the design is implemented.

User Attitude (i.e., perceptions, affect): While two systems may have identical utility and both prove usable, users may express a preference for a system based on personal judgment, previous experience, aesthetics, cost etc. Therefore, the final driver of use must be the user's attitude to the technology.

To render these drivers meaningful for systems implementation we propose they be conceptualized in the following terms as the P³ framework, i.e.,

Power: An objective measure of the applications capability/functionality, i.e., "what the machine can do".

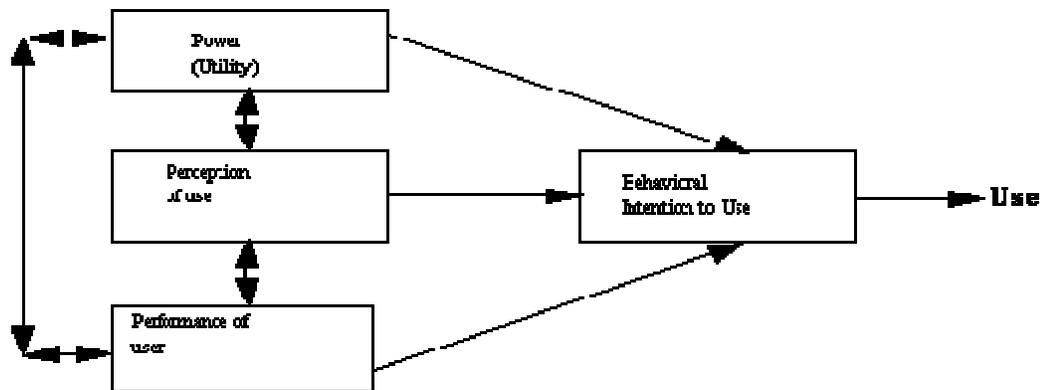
Performance: Behavioral measures of usability as in traditional usability engineering.

Perception: Perceptual measures from users re: usability, utility, etc., (as in TAM).

By conceptualizing systems evaluation in these terms one can offer inputs to design appropriate to the stage of the process in which one is working. Thus, power needs to be considered first, since the inability to even support task performance renders the application limited in use for the intended audience. Assuming task support is provided by the tool we next address the capability of the user to exploit it (here operationalized by traditional usability measures). Finally, given a usable and utilitarian tool, the ultimate question is whether people will choose to use it or not. This is the domain of perception, where TAM most naturally fits.

From a theoretical perspective the P³ framework suggests that the usability engineering constructs of *effectiveness* and *efficiency* of use, and the TAM constructs of *perceived ease of use* and *perceived usefulness* are really different entities that should not be compared directly. The *capability* to use and the *perception* of ease of use are necessarily different constructs and it is plausible that they may even contradict each other in certain situations (e.g., where users performance and perceptions are not equivalent).

We therefore believe that further research might usefully address the extent to which performance data from user trials maps onto later perceptual ratings in TAM. Conversely, comparing TAM data with user performance scores on developed systems will show the degree of overlap between perception and performance for a given context. Where various measures are taken, we may begin to identify the relative weights of each P component (Power, Performance, and Perception) in explaining user behavior with new technology.



The P3 Model: Use is determined by Power, Perception and Performance

References

Davis, F.; R. Bagozzi, and P. Warshaw (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models, *Management Science* 35(8): 982-1003.

Dillon, A. (1987). Knowledge acquisition and conceptual models: a cognitive analysis of the interface. In: D. Diaper and R. Winder (eds.) *People and Computers III*. Cambridge: Cambridge University Press.

Dillon, A., and Watson, C. (1996). User analysis in HCI—the historical lessons from individual differences research. *International Journal of Human-Computer Studies* 45(6), 619-638.

Dillon, A. and M. Morris (1996) User acceptance of information technology: theories and models. In: M. Williams (ed.), *ARIST*, Vol. 31, (M. Williams, ed.), Medford, NJ: Information Today.

Morris, M. (1996) *A Longitudinal Examination of Information Technology Acceptance: The Influence of System Experience on User Perceptions and Behavior*, Ph.D. Dissertation, Indiana University.

Morris, M. and A. Dillon (1997). How user perceptions influence software use, IEEE Software (14)4: 58-65.

Nielsen, J. (1993). Usability Engineering. Boston: Academic Press.

Shackel, B. (1991). Usability—Context, Framework, Definition, Design, and Evaluation. In B. Shackel and S. Richardson (eds.). *Human Factors for Informatics Usability*. Cambridge: Cambridge University Press.