Dr. Lawrence Sager and Dr. Rebecca Grier

# Identifying and Measuring the Value of Human Factors to an Acquisition Project


## ABSTRACT

This paper addresses three issues of interest to both human systems integration (HSI) practitioners and to program managers: (1) what is the expected value of human factors engineering (HFE) involvement on an acquisition project? (2) how can that value be measured? and (3) how can that value be optimized? We begin by examining several case studies showing the potential range and magnitude of benefits associated with HFE involvement. Next, we present a simple methodology for determining the cost-effectiveness of HFE intervention within a specific system development context. Finally, we review a sample of HFE methods that have been shown to benefit system development efforts. We conclude with a set of recommendations for putting the above information into practice.


## INTRODUCTION

In traditional systems acquisition, inadequate consideration of human factors engineering (HFE) issues is a familiar problem. That is, HFE issues are not expressly considered, they are considered but their importance is underestimated, or they are considered too late in the design process. In the latter cases, it is not uncommon for project teams to consider HFE issues at only one or two points, rather than throughout, the system development lifecycle, preferring to focus limited resources on technical issues with a higher perceived payoff. Similarly, if a project experiences a budget or schedule overrun, project activities and/or system components with the lowest perceived value are generally the first to be cut.

Such problems could be avoided, in many cases, through better understanding of the value HFE potentially brings to an acquisition project. It is difficult (but not impossible) to estimate the economic impact of usability, but it is extremely important to do so, as the results can be dramatic and significantly impact the course of a project.

While most of the examples in this paper come from the literature on human factors and usability engineering techniques, the underlying principles can be extrapolated to other areas of concern to HSI practitioners (e.g., health, safety, survivability).

## THE BENEFITS OF HFE

In fact, a positive return on investment (ROI) has been documented in a number of studies (e.g., Nielsen, 1993; Bias and Mayhew, 1994; Pressman, 2000). By contrast, systems with bad usability can have significant negative impact in terms of wasted time, unnecessary expense, incorrectly or incompletely performed tasks, user frustration, and safety issues. Moreover, the early inclusion of HFE techniques in the development of a system has proven to reduce development time and costs. Some examples:

### Total Ownership Cost Reduction

Only 20% of software's lifecycle cost is incurred during development. The majority of costs occur after the product is released. Of that work, 80% is due to unmet or unseen user requirements, which could be reduced by following HFE techniques (Karat, 1993).

Similarly, it has been reported by Cunnigham (2001) that typically between 50 and 60% of a military system's total ownership costs is in manning and maintaining the system. Optimizing the crew and considering the maintainer through use of HFE techniques in development will significantly reduce these costs.

### Training Reduction

Systems designed in accordance with usability engineering principles have typically reduced the time needed for training by around 25% (Landauer, 1995). Bias and Mayhew (1994) have many similar examples, such as the company that reduced training on a system from one week to one hour by including HFE experts on the design team. AT&T also reported that having HFE experts on the team for the redesign of an internal system reduced training expenses by $2.5 million a year (Bias and Mayhew, 1994).

### Mission Performance Improvement

Numerous companies have reported that including HFE techniques in the redesign of internal systems has increased the productivity of employees while reducing training time and costs. For example, Dray and Karat (1994) reported that a user interface designed with HFE techniques resulted in a 32% internal rate of return. These costs savings were realized by a 35% reduction in training and a 30% reduction in supervisory time.

It is widely reported that 80% of maritime accidents are ruled due to human error (Fotland, 2004). User-centered design typically cuts errors in user-system interaction from 5% down to 1% (Landauer, 1995). This is significantly great than the 50% reduction in accident rates requested by Secretary Rumsfield in a May 19, 2003 memo.

### Development Cost Reduction

Though these are remarkable figures, they are results seen with patience. HFE techniques can have more immediate benefits to develop time and costs. For example, Bosert (1991) stated that the inclusion of HFE results in a reduction in the product development cycle by 33-50%. Further, Barker (2000) and Nielsen (2003) have reported that the top four reasons given by managers for inaccurate development cost estimates could have been addressed by following HFE best practices. Finally, one high tech company reported that the use of HFE techniques reduced the time spent on one tedious development task by 40% (Bias and Mayhew, 1994).

### Monetary versus Non-Monetary Benefits

Although a compelling case for HFE involvement can be made on economic grounds alone, economic justification may not be the only, or even the most important consideration in justifying HFE activity on a particular endeavor. Managers typically consider multiple criteria in deciding which course of action to follow. Among the factors to be considered might be: safety, user attitudes, reliability, and maintainability. Some of these factors may readily translate into economic objectives (e.g., maintainability can be expressed in terms of annual maintenance costs), but some (e.g., end user morale) may not. (Though in fact end user morale could indirectly affect total ownership cost by reducing costs associated with turnover, this relationship may not be of primary interest to the manager in a particular context.)

When dealing with non-monetary metrics, one loses the ability to speak precisely about “return on investment” (dollars spent vs. dollars earned or saved), but, provided the objectives can be quantified, one can still measure benefits and compare them to costs.

### Overall Impact on a Program

An excellent example of all of the benefits of including HFE is the Air Force’s C-141.

The C-141 was originally designed as a modular cargo plane. That is, various mission packs would be swapped in and out based on the particular mission. According to Hendrick (1996), the inclusion of HFE added less than $500,000 to the development costs of the cargo plan. However, the changes to the design based on the HSI team reduced the cost of producing the plane by nearly $5 million. Some of these changes included eliminating the conversion kits based on a task analysis. Further, these changes are credited with saving $50 for every $1 they cost in total ownership expenses. These savings come from eliminating the man hours required to convert the plane for each mission, reducing the storage needs for unused conversion kits, and reducing training for maintainers in conversion. Additionally, this plane has an exceptional safety record and has accrued untold savings in lost aircraft avoidance. Finally, pilots of the C-141 enjoy flying this plane and think it was smartly designed. As such, there are fewer turnovers, which results in reduced training and improved total system performance.

### The Value of Early Intervention

An oft-quoted statistic is that it is about 40-100 times more expensive to fix problems in the maintenance phase of a program than in the design phase (Boehm, 1981). Slightly restated, if it costs $10 to make a program change during development, it will probably cost $400 to do it after the system is in the field (Pressman, 1992). In fact, Rauterberg and Strohm (1992) found that only 33% of the maintenance effort is spent for debugging, suggesting that 67% is therefore spent on modifying the system due to changes in requirements caused by user dissatisfaction with the system as originally delivered.

Related to the notion of early intervention is the use of low-fidelity prototypes. Virzi (1989) makes a convincing case for the preferability of using low-fidelity prototypes early in the design phase, rather than deferring usability testing until a high-fidelity prototype becomes available. Not only does this allow for early availability of usability and user preference data to drive key decisions, but the low-fidelity prototypes themselves tend to be relatively inexpensive to produce, thereby allowing researchers to quickly explore multiple design approaches.

## A SIMPLE METHOD FOR DETERMINING COST-EFFECTIVENESS

A cost-effectiveness analysis is conceptually no more than a matter of calculating the expected costs of an activity and estimating its potential benefits to determine whether one outweighs the other. In this section, a simplified method is presented for performing such an analysis. The key steps to this method are:

1. Setting goals.
2. Determining when to perform interventions.
3. Identifying candidate alternatives.
4. Estimating costs.
5. Estimating benefits.
6. Comparing alternatives.

Estimating is important because, generally speaking, cost-effectiveness analyses are conducted before an intervention, precisely to decide whether the expense of the intervention is warranted.

### Setting Goals via the Defense Acquisition System Framework

The Defense Acquisition System is the program management process that exists to manage all DoD investments in technologies, programs, and product support. The fundamental principles and procedures that DoD follows are described in DoD Directive 5000.1, *The Defense Acquisition System* (2003) and DoD Instruction 5000.2, *Operation of the Defense*

*Acquisition System* (2003). The Defense Acquisition System calls for the identification of HSI requirements and inclusion of certain HSI activities throughout a program. The HSI domains of particular interest to the Defense Acquisition System include manpower, personnel, training, human factors, safety and occupational health, personnel survivability, and habitability.

DoD Instruction 5000.2 requires every program manager to document program goals prior to program initiation. Program goals consist of an objective value and a threshold value for each parameter. Objective values represent what the user desires and expects. The program manager manages the program to the objective value of each parameter. Thresholds represent the acceptable limits to the parameter values that, in the user's judgment, still provide the needed capability. For performance, a threshold represents either a minimum or maximum acceptable value, while for schedule and cost parameters, thresholds would normally represent maximum allowable values. The failure to attain program thresholds may degrade system performance, delay the program, or make the program too costly. The failure to attain program thresholds, therefore, places the overall affordability of the program and/or the capability provided by the system into question. Thus, it behooves HSI practitioners to identify HSI benefits in terms of objective and threshold values that are meaningful to stakeholders within the context of the overall program goals.

## Guiding Principles in Determining When to Apply HFE Methods

It is also important for HSI practitioners to identify the stages or milestones within the product realization process at which the project team should examine objective and threshold values. Three guiding principles should be considered:

1. Like virtually all modern program management processes, DoD 5000 incorporates basic concepts such as project phases; iterative and recursive processes; documents and deliverables; and technical reviews. Associated with all of these are decision points at which HFE data could be examined and course corrections made as necessary.
2. Optimizing the placement of HFE decision points within the overall program management process can be something of a balancing act: The risk with placing decision points too early in the process is decision-making based on incomplete information. Placing them too late in the process, on the other hand, carries the risk of expending time and money on developing a solution that is eventually shown to be unviable.
3. A related challenge is in defining the optimal number of decision points. It may not simply be a case of "the more, the better," because the activities associated with each decision point themselves take time and consume resources.

Each decision point should be examined with the purpose of addressing one or both of these questions: (1) Can system development proceed as planned? And (2) should it proceed as planned? Question 1 focuses mainly on the technical issues – e.g., do we know enough about the user's visual capabilities to proceed with screen designs? If the answer to Question 1 is "no," then Question 2 is moot. If the answer to Question 1 is "yes," then Question 2 becomes meaningful. Indeed, a "no" response to Question 2 could trigger a reevaluation of system requirements, performance objectives, schedule, and/or budget. In accordance with Defense Acquisition principles, it may be possible simply to rebalance cost, schedule, and performance within a "trade space," provided threshold values are not exceeded. In more extreme cases, issues may need to be escalated and project scope redefined (Pressman, 2000).

## Identifying Candidate Alternatives

A wide variety of HFE methods are available to support analysis, design, and evaluation activities. A brief sample of methods is provided later in this paper. More comprehensive descriptions of methods, tools, and techniques are available in the human factors research literature, as well as via on-line reference sites; e.g., the Human Systems Information Analysis Center (http://iac.dtic.mil/hsiac/index.htm).

## Estimating Costs

There are essentially two ways to estimate the cost of a planned HFE intervention: (1) By drawing analogies to similar situations, either via case studies or personal experience; and/or (2) by applying expert judgment to identify, for example, likelihood and impact of risk events (e.g., difficulties in recruiting subjects, data collection equipment malfunctions, etc.) To assist in this effort, case studies are available both in the human factors research literature and on-line (e.g., www.usability.gov, www.eurocontrol.int).

With some experience, one can learn to estimate quite accurately the cost of performing HFE activities, in terms of number of hours of time, plus additional expenses such as travel and equipment. Estimating the amount of time needed, for example, to recruit subjects or to prepare a report is something learned through practical experience and identification of useful rules of thumb. It is also worth considering for an organization that expects to be involved in HFE activities on an ongoing basis to create their own database of case studies to facilitate estimates of the value of future work.

## Estimating Benefits

The key to a cost-effectiveness analysis, of course, is quantifying the expected benefits so that these can be compared to costs. One must be mindful that different benefits are relevant to different stakeholders. Benefits fall into three general categories:

1. Those that are inherently monetary (e.g. cost to provide training).
2. Those that can be converted into monetary terms (e.g., duration of time to train warfighters can be expressed in dollars by taking into account the labor rates paid to both students and instructors).
3. Those that are quantifiable, but for which the stakeholder may not be primarily interested in money (e.g., warfighter morale or accident rates).

In principle, any type of benefit is adequate to conduct a cost-benefit analysis, provided it is quantifiable. The key advantage to measuring benefits in monetary terms is perhaps a more straightforward comparison of costs and benefits – while one might debate whether a 10% improvement in warfighter morale was "worth" the cost of developing a new system, one is less likely to debate the merits of a system for which the development cost of $1 million yields a savings in training costs of $5 million. Ease of interpretation aside, however, there is no conceptual reason why non-monetary units of measure cannot be used to quantify benefits.

The principal sources of information on which to base assumptions about benefits are published case studies (see URL references above) citing the impact of various HFE treatments on measures of usability, combined with expert judgment as to how readily case studies can extrapolate to a specific project. (This is something learned through experience, and the HSI practitioner who anticipates being faced with the challenge of cost-justifying his work is well-advised to refine this particular form of expertise.) According to Mayhew (1990), most stakeholders will accept an analysis whose claims regarding predicted benefits are based on conservative assumptions supported by this literature.

Even more convincing to stakeholders, and of potentially significant value for subsequent projects, are empirical data from the acquisition program itself that captures actual costs and benefits. Such measures provide the data to validate and/or support the refinement of estimates. For an acquisition program in which the same HFE issues are likely to be revisited many times, a case history based on empirical data can be an especially powerful tool for predicting future benefits.

### Special Considerations

Although beyond the scope of this paper, a couple of special considerations are worthy of mention. The first pertains to the time value of money. Put simply, the value of a given sum of money depends on when it is received. Thus, when performing a cost-effectiveness analysis, one may need to take into account the time lag between when costs are incurred and benefits received. If this amount of time is significant, then benefits may need to be discounted accordingly.

A second consideration applies to the comparison of HFE interventions that offer multiple, non-monetary benefits. How does one, for example, choose the more desirable of the methods shown in Table 1?

**TABLE 1: Cost-Benefit Comparison**

| Method | Cost | Benefits | |
|---|---|---|---|
| | | Increase in morale | Reduction in accidents |
| A | $100K | 20% | 35% |
| B | $150K | 30% | 10% |
| C | $150K | 20% | 10% |

Clearly, Method B is preferable to Method C, since B offers a greater increase in morale for the same cost. Similarly, Method A is preferable to Method C, since A offers a greater reduction in accidents for a lower cost. But what about the relative desirability of A versus B?

One approach to addressing this problem is to apply psychometric scaling methods to transform various measures to a single measure, say, values on a “utility” scale. This would allow for comparisons such as “an X% increase in morale has the same utility as a Y% decrease in accidents.”

## A SAMPLE OF HUMAN FACTORS ENGINEERING METHODS

Many criteria must be considered in choosing a HFE evaluation method. The three most important of these are (1) In what stage of design is the system? (2) How much time is available? and (3) What resources are available? In describing several methods of evaluation, we will consider each of these questions.

### Cognitive Walkthrough

A cognitive walkthrough is an evaluation of task flow. In particular, the evaluation identifies what error messages are needed, whether all information is attainable when required, and whether the tasks flow in an order that users can anticipate. In order to perform a cognitive walkthrough, a wire-frame or flow chart representation of the system is required. As such, it can be completed very early in the development of a system. It also requires very little time, but this depends on the size of the system being developed. Prior to conducting the evaluation, the evaluators should be familiar with the user groups and their goals.

### Heuristic Review

A second technique is the heuristic, or expert, review. In this type of evaluation, a human factors expert, knowledgeable in the guidelines for the system, uses the system to perform common or critical tasks. In so doing, the expert identifies design elements that are contrary to the guidelines or that may cause errors. To perform a heuristic review, a prototype system must exist, although this system can be as simple as screen shots and a wire-frame or flow chart representation of the system. As such, it can be relatively early in the design process. A key advantage of the heuristic review is its short (compared to other methods) turnaround time.

## Usability Testing

A third evaluation method is usability testing, in which a sample of end users is asked to use the system to perform a set of common or critical tasks. A human factors expert observes the user performing these tasks and identifies areas that are difficult, and perhaps identifies alternative designs to reduce the difficulty. There are several different approaches to usability testing, depending on what phase of design the system is in.

Common variations on usability testing include the paper prototype test, the Wizard of Oz test, and the criterion test. The paper prototype test, as the name suggests, has users interact with non-functioning prototypes. As such it can be performed early in the development cycle. In these tests the user is asked to indicate by pointing what they would do to perform the tasks. The tester records these actions and presents the screens in the order indicated by the user's actions. The Wizard of Oz (WOZ) test presents the user with screens on a monitor, but a "man behind the curtain" controls the screens that are presented to the user. As such, two persons are required to run the test. Further, a WOZ test requires that screen designs be fairly set, but that the back end of the system need not be developed. Therefore, it is generally performed later in the lifecycle than the paper prototype test. For a criterion test, the system must be fully functional, and thus the test can only be performed very late in the development cycle when the user can interact with the actual system. Because the system is fully developed, human performance data can be gathered that can reliably be used to compare the total system performance of a new system with the one it is replacing.

## Test Protocol

Another way to view usability testing is in terms of the test protocol, i.e., the procedures used by the tester in collecting the data. The most common protocol in usability testing is the "think aloud test." Participants are asked to vocalize their thoughts while performing the tasks. They are allowed to ask questions of the tester, but the tester does not provide answers. Verbalizations are quite useful in understanding mistakes that are made and getting ideas for what the causes might be and how the interface could be improved to avoid those problems. This is the only protocol that provides reliable data for system comparison. A second protocol is "co-discovery." In this format, two users are asked to perform a task. Their conversation while performing the task provides insight into their approach for performing the task. This is arguably more natural than thinking aloud, but depending on the task may be less representative of the real world and as such may not provide reliable human performance data. The "question asking" protocol is a format where the tester specifically asks the user questions while they are performing the tasks. These questions shed light on how the screens are being interpreted as well as getting users' reactions to parts of a design that the user may otherwise be neglecting. However, as with the co-discovery method, the human performance data is not reliable enough for system comparison. The final technique is the "coaching" protocol, where users are trained to use the system and then are asked to teach other users how to use the system. Again, these data cannot be used for system comparison, but can provide great insight into the learnability of a complex system.

### Sample Size

In designing a usability evaluation, one should be mindful of the finding of Nielsen (1990), that a large proportion of usability problems are revealed by the first few subjects. Thus, it may be more cost-effective to conduct multiple, relatively small, iterative design-test cycles, rather than run a single evaluation with a large number of subjects.

## CONCLUSION

HSI practitioners faced with the challenge of cost-justifying their activities should consider these recommendations:

1. Become thoroughly familiar with the DoD 5000 acquisition process, especially the points at which critical design-affecting decisions are made. These decision points constitute the key opportunities to contribute to the process.

2. Recognize that, all other things being equal, the best opportunities to impact system design tend to occur earlier in the development cycle. On the other hand, early involvement means providing timely deliverables and usability design and test methods based on low-fidelity renditions of the system.

3. Know what goals are important to project stakeholders, and whether these goals should be measured in dollars or some other units. Making the case for HFE by measuring something that doesn't matter to the customer is a wasted exercise.

4. Be sufficiently familiar with a broad range of usability methods, tools, and techniques so that you can identify the most suitable approach for measuring a specific goal within the time, cost, and resource constraints under which the project must operate.

5. Get better at estimating costs and benefits by validating, whenever possible, your predictions. In addition to maintaining familiarity with published case studies, consider creating your own database of cost-benefit data to improve your organization's precision at estimating.

## REFERENCES


Barker, D. T., "Cost Benefits of Usability Engineering," 2000. Retrieved October 9, 2001, from www.interfacearchitecture.net/articles/benefits.htm .

Bias, R.G. and D.J. Mayhew, *Cost Justifying Usability*. Boston, MA.: Academic Press, 1994.

Boehm, B. W., *Software Engineering Economics*. New Jersey: Prentice-Hall, 1981.

Bosert, J. L., "Quality Functional Deployment: A Practitioner's Approach," 1991. In Bias, R. G. & D. J. Mayhew, *Cost-Justifying usability*. Boston, MA: Academic Press, 1994.

*Computer World*, "Norwich Rethinks Customer Service." 24 November 1995 issue.

Cunnigham, P.H., "There Are No Unmanned Systems" In *Human Systems Information Analysis Center Gateway, Volume 12 (2,)* p 1-2, *2001*.

*Department of Defense Directive 5000.1: The Defense Acquisition System*, 2003.

*Department of Defense Instruction 5000.2: Operation of the Defense Acquisition System*, 2003

Dray, S.M. and C.M. Karat, “Human Factors Cost Justification for an Internal Development Project.” In Bias, R.G. and D.J. Mayhew, *Cost Justifying Usability*. Boston, MA.: Academic Press, 1994.

Fotland, H., “Human Error – A Fragile Chain of Contributing Elements” In *Alert: The International Maritime Human Element Bulletin, Issue 3,* April 2004.

Hendrick, H.W., “Good Ergonomics is Good Economics.” Presidential address to the Human Factors and Ergonomics Society. Published in the *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting*, 1996.

Karat, C. “Usability Engineering in Dollars and Cents.” In *IEEE Software, May Issue,* p. 89, 1993.

Landauer, T.K., *The Trouble with Computers*. Cambridge, MA.: MIT Press, 1995.

Mayhew, D.J. “Cost-Justifying Human Factors Support – A Framework.” In *Proceedings of the Human Factors Society 34th Annual Meeting*, 1990.

Nielsen, J. “Evaluating the Thinking Aloud Technique for use by Computer Scientists.” In *Advances in Human Computer Interaction Vol. 3*, H. Hartson and D. Hix (Eds.). Ablex, 1989.

Nielsen, J., *Usability Engineering*. Boston, MA.: Academic Press, 1993.

Pressman, R. *Software Engineering: A Practitioner’s Approach, 5th Edition*. Boston, MA.: McGraw-Hill, 2000.

Rauterberg, M. and O. Strohm, “Work Organization and Software Development.” In *Annual Review of Automatic Programming, Vol. 16(2)*, pp. 121-128, 1992.

Virzi, R.A., “What Can You Learn From a Low-Fidelity Prototype?” In *Proceedings of the Human Factors Society 33rd Annual Meeting*, 1989.

---

**Lawrence Sager, Ph.D.** is a senior human factors engineer with Aptima., Inc., where his current research focuses on the design of tools and methods to improve communication and collaboration within hierarchical teams. Dr. Sager has over 20 years of experience in planning, designing, and evaluating complex human-machine systems and service operations. He has improved the usability of web sites, operations support centers, network services, work environments, and electronic products. Prior to joining Aptima, Dr. Sager worked at AT&T, where he held the positions of human factors consultant, systems engineer, and project manager. Dr. Sager also worked at VocalTec Communications as the director of project management, and at Vonage Holdings as the leader of the systems engineering and project management teams. Dr. Sager holds a Ph.D. in Cognitive Psychology from the Johns Hopkins University and a Master’s Certificate in Project Management from George Washington University.

**Rebecca Grier, Ph.D.** is a human systems engineer with Aptima, Inc. Primarily she conducts analyses to inform the design of systems for both commercial and military domains. Prior to joining Aptima, Dr. Grier was an engineering psychologist in the Human Systems Integration (HSI) Group at the Naval Surface Warfare Center, Dahlgren, VA. In this position she developed courses to educate the Navy’s acquisition workforce about HSI and provided technical support to acquisition programs including writing HSI plans and conducting system evaluations. Rebecca has also worked in the Human Factors Engineering Group at SBC Laboratories, Inc., where she evaluated and designed installation wizards, manuals, dashboards,

CPE, IVRs, web pages, and voice UIs. Dr. Grier received her Ph.D. and M.A. in Human Factors/Experimental Psychology from the University of Cincinnati, and a B.S. Honors in Psychology from Loyola University, Chicago.