APPLICATION OF THE ANALYTIC HIERARCHY PROCESS

TO COMPLEX SYSTEM DESIGN EVALUATION

Michael L. Talbert
Osman Balci
Richard E. Nance

Department of Computer Science
and
Systems Research Center

Virginia Polytechnic Institute and State University
Blacksburg, VA 24061-0106

ABSTRACT

This paper examines the use of the Analytic Hierarchy Process (AHP) to weight indicators used in the evaluation of complex system designs which involve software, hardware, and humanware. Since such a comprehensive design can easily include hundreds of system quality indicators, evaluators need a technique to ensure the identification and emphasis of salient indicators in the determination of the quality of the design. The AHP is a popular technique for determining relative worth among a set of elements. In the present work, we introduce AHP with a simple example, then illustrate the application of the AHP to design evaluation using a subset of indicators from the human component of a system. We note in some detail issues which require added attention when applying the AHP to this domain. The issues include indicator selection, dealing with large numbers of indicators, incorporating group judgments, and conflict resolution. We found AHP to be an effective tool for use in assigning weights of criticality in indicator-based design evaluation, and propose elements of an environment in which the use of AHP is easily incorporated.

1. INTRODUCTION

A complex Navy system is composed of three major components: software, hardware, and humanware, each of which must work in synergy through complex interactions and interfaces (see Figure 1). These components are intertwined with real-time mission-critical characteristics, posing significant technical challenges for system designers, developers, and maintainers independent of warfare domain.

A critically important phase in the engineering of complex systems methodology is system design measurement and evaluation. The commitment to spend millions of dollars for building a system must be justified by thorough and in-depth evaluation of the proposed design considering the array of operational environments in which the system might be deployed.

Measurement at these scales is an important activity of interest in many disciplines. However a standard terminology does not exist. Similar terms are widely applied in various disciplines to convey the same notion. An extensive survey of the literature in the fields of hardware, software, and human factors revealed the following suite of synonymous terms: "metric," "measure," "index," "scale," "factor," "figure of merit," and "indicator." In this paper, we use the term *indicator* and define it as an indirect measure of a qualitative concept. We define *metric* as measurable component in the expression for an indicator, the value of which can be computed by a formula or measured directly.

The common goal in disciplines applying these and other related terms is, of course, to try to accurately measure a concept which can be either quantitative or qualitative. Measurement of quantitative concepts (e.g., response time, throughput, utilization) can be done directly, whereas qualitative concepts (e.g., design utility, maintainability, complexity) must be measured indirectly.

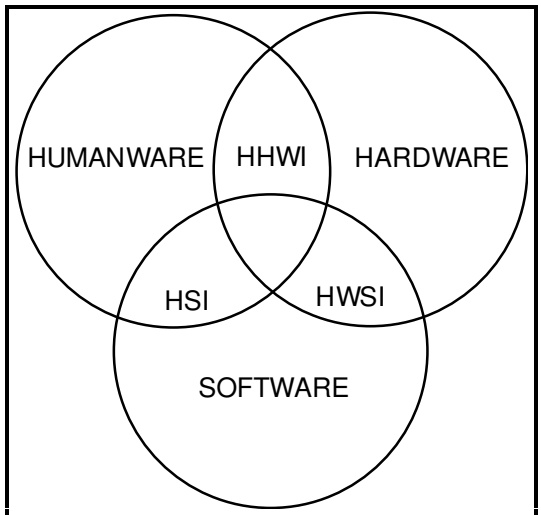

Figure 1. Components of a Complex System.

The most common approaches to assigning values to metrics are through direct measurement or system simulation. We can directly measure response time to a user request, but are better served by simulation in the measure of, for example, resource utilization. More subjective methods are applied in indirect measurement schemes. Balci *et al.* [1993] propose the use of the expert opinions of system design engineers and application domain specialists. These indicators can be expressed as either a single value or a range of values, referred to in the literature as "crisp" or "fuzzy" measures, respectively [Kaufmann and Gupta 1985].

A classic example structure for depicting indirect measurement of indicators is a hierarchy (Figure 2). In a hierarchical approach, concepts are measured by a set of composite indicators. These indicators themselves may be measured by composite indicators, until at the bottom of the hierarchy, we have lowest level indicators, or "leaves," in the sense of a graph. These lowest level indicators must be assigned crisp or fuzzy values.

The composite score for the top level concept is intuitively a function of the leaf scores, aggregated all the way up the hierarchy. The aggregation process involves two phases: (1) assigning or computing *weights of criticality* for all indicators from the leaves to the root, and (2) computing the overall score for the top-level concept (the design element) from the scores assigned to the leaves.

In assigning weights to individual or aggregate indicators, we attach a level of significance, representative of the contribution of the indicators to the quality of the design. The weight of an indicator is our assurance that the most important aspects of the design, in our opinion, are judged the most critically, and the less important ones less so. In so doing, we strive to ensure that the value of the design is measured not as merely the sum of the values of the components; it is more the case that the evaluation of quality of the entire design reflects the degree to which all the components are "pulling their weight."

The assigning of the weights then becomes a critical step in evaluating the design. We propose the Analytic Hierarchy Process (AHP) as a means to assign weights to indicators

The purpose of this paper is to present how the AHP can be used for the evaluation of complex system designs by providing a mechanism for deriving weights for system indicators. Section 2 presents an overview of the AHP and illustrates its use by a simple example. Section 3 provides a more detailed example, applying the process to a suite of system-level indicators germane to the human element of a complex system. The section also addresses some special attention needs. Finally, section 4 provides a brief summary, states the conclusions of our research, and introduces directions for future research. These include (1) the application of fuzzy mathematics as the mechanism for score assignment and aggregation, since some values are derived from subjective judgment, and (2) the incorporation of knowledge-bases and rule sets as a means of ensuring consistency and validity in the weighting and scoring processes.

2. THE ANALYTIC HIERARCHY PROCESS: AN OVERVIEW

The Analytic Hierarchy Process [Saaty 1980] comes from the field of decision theory and has been widely applied in fields ranging from finance [Srinivasan and Bolster 1990], to software engineering [Toshtzar 1988], to combat force evaluation [Lee and Ahn 1991]. It is often used in decision making when choosing among alternatives based upon their relative worth. Its use is keyed on the methodical application of expert judgments to determine the relative contribution of factors to a decision. Saaty's structure for analyzing a decision is a hierarchy and his mathematical vehicle for computation is a positive reciprocal square matrix. Details of the mathematics are beyond the scope of this paper, and the reader is referred to Saaty's original work. AHP has been widely studied and the literature is rife with its adaptations, extensions, criticisms, and applications (see e.g., [Murphy 1993; Masuda 1990; Harker 1987a,b; Saaty 1987; DeTurck 1987; Weiss 1987; Zahedi 1986]).

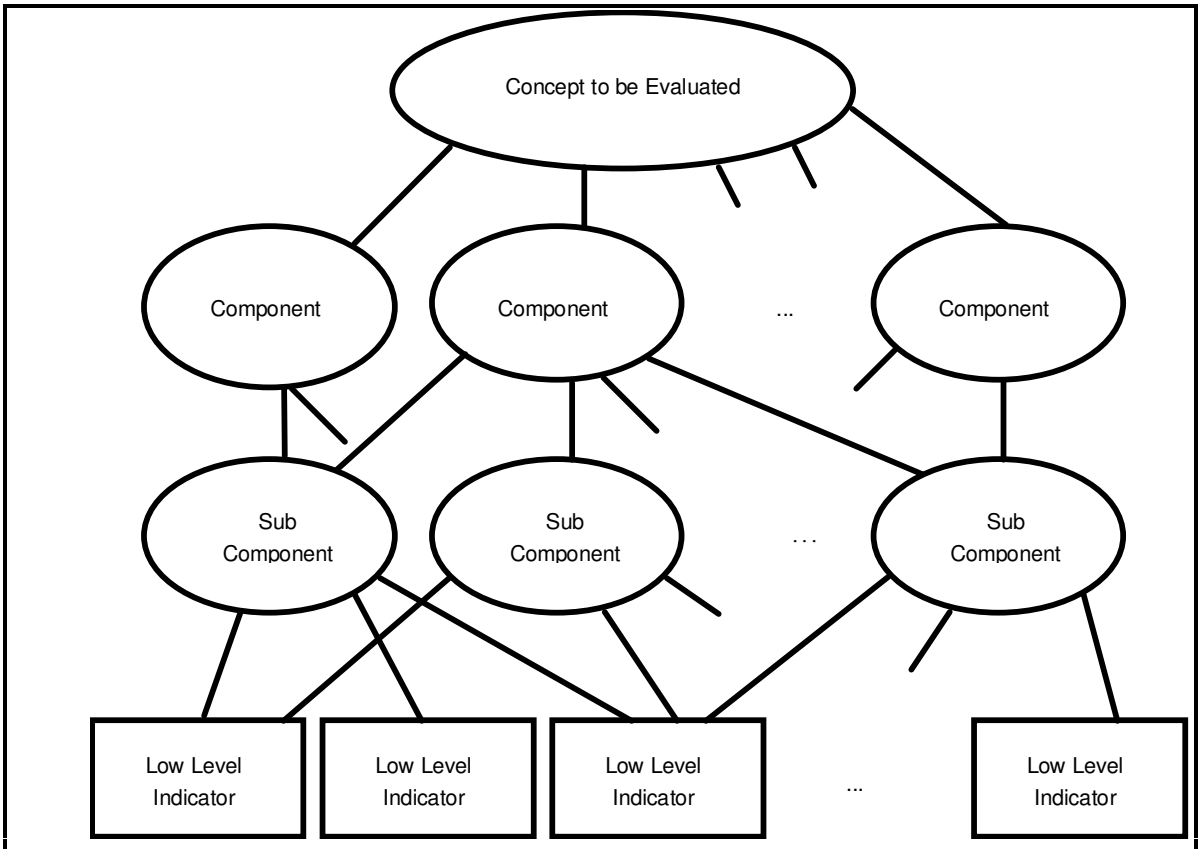

Figure 2. Typical Evaluation Hierarchy.

2.1 How AHP Works

The basic phases to applying the AHP are as follows:

1. Define the decision issue and bound its scope.
2. Structure a hierarchy of the elements contributing to a decision.
3. In a bottom-up fashion, construct a pairwise comparison matrix of the relevant contribution or impact of each element on its governing criterion in the next higher level. Saaty [1980] has developed a pairwise comparison scale for use in computing relative worth of decision elements (see Figure 3). Extensions on this basic scheme exist, including a 1-20 scale and scales which allow real numbers [Saaty 1993; Zahedi 1986].
4. Using the eigenvector method described by Saaty [1980], compute the relative contribution of each element and test the individual matrices for consistency (reflects soundness of judgment, understanding of the interdependencies of criteria, etc.).
5. Working downward through the hierarchy, use hierarchical composition to combine the weight vectors and arrive at global and local relative contributions (priorities) of each element.
6. Perform overall consistency check on the completed hierarchy of weights. Reassign relative weights contributing to a consistency ratio out of tolerance (see section 2.2).

We omit details of the mathematics and direct the reader to the references listed above, especially [Saaty 1980] for a thorough discussion. The key to using the AHP is building the hierarchy of elements and computing their relative weights. To introduce the process, we provide a simple example taken from [Saaty 1990].

2.2 A Simple Example of AHP

The decision to be made is: Which car among Chevrolet (C), Thunderbird (T), and Lincoln (L) to purchase based on comfort?

The decision maker makes the value judgments that the Lincoln is twice as comfortable as the Thunderbird, the Thunderbird is twice as comfortable as the Chevrolet, and demonstrating perfect consistency, the Lincoln is judged as four times as comfortable as the Chevrolet. A brief mention of the meaning of *consistency* is appropriate here (see Figure 4). The judgments as given above are perfectly consistent because their ratios demonstrate transitive equality. Examining the judgments more closely reveals that C = (1/2)T and T = (1/2)L, so by transitivity, C should equal (1/4)L, and it does. As we shall see below, a departure from perfect consistency is generally the rule rather than the exception. Note in Figure 4 the diagonal elements of "1" indicating that each car is judged to be equally as comfortable as itself, and that symmetric entries are computed as inverses.

To compute overall priorities, the matrix is normalized as follows:

1. Obtain the column totals (7, 3.5, 1.75).
2. Divide each entry by its column total.
3. Take the average value for each row.

This produces the vector of priorities for (C, T, L) as (0.14, 0.29, 0.57). An interpretation of this is that the Lincoln culled 57 percent of the comfort vote, the Thunderbird, 28 percent, and the Chevrolet, 14 percent. Note that the computations were quite simple due to the perfect consistency of the matrix. In a less consistent matrix (Figure 5), note the difference in priorities.

The column totals of (7, 5.5, 1.5) lead to a priority vector (row averages) of (0.13, 0.21, 0.66). Note that in the presence of inconsistency, the higher scores increased and the lower score decreased. Saaty [1980] provides an index of consistency (CI) determined as follows:

1. Multiply each entry by its priority (Figure 6).

DEGREE OF IMPORTANCE	DEFINITION	EXPLANATION
1	Equal importance	Two elements contribute equally to the property
3	Moderate importance	Experience and judgment slightly favor one element over another
5	Essential or strong importance	Experience and judgment strongly favor one element over another
7	Very strong importance	An element is strongly favored; its dominance is demonstrated in practice
9	Extreme Importance	The evidence favoring one element over another is of the highest possible order of affirmation
2,4,6,8	Intermediate values between two judgments	Compromise is needed between two adjacent judgments
Reciprocals		When element i compared to j is assigned one of the above numbers, then element j compared to i is assigned its reciprocal.

Figure 3. AHP Scale of Judgments.

2. Obtain a vector of new row sums = (0.41, 0.64, 2.02).
3. Divide the vector of row sums by the vector of priorities yielding: (0.41, 0.64, 2.02) / (0.13, 0.21, 0.66) = (3.15, 3.05, 3.06).
4. Compute the average of these values (approximately 3.09). It is easily verifiable that the corresponding value for a perfectly consistent matrix is 3.0. In fact, it is a convenient property of any such special *n*-by-*n* matrix that the CI (in actuality the maximal eigenvalue) will always be the value *n*.
5. Subtract the value for the consistent matrix from that for the inconsistent matrix and dividing by *n*-1 yields (3.09 - 3)/2 = 0.045.
6. Consult a table of "standard" consistency indicies to find for a matrix where *n*=3, that CI = 0.58. These values represent CI's for matrices with random entries from the range 1/9, 1/8, ... , 8, 9. For reference, we reproduce these values in Table 1 [Saaty 90]. The reader is directed to [Murphy 1993] for a discussion of the limiting nature of the CI.
7. Divide the approximation CI's by the value obtained in Step 6, and obtain the consistency ratio (CR) 0.045/0.58 = 0.08.

Saaty [1980] posits that a CR of less than 0.10 indicates satisfactory consistency. See also [DeTurck 1987] for a discussion of this value.

In concluding this section, it is important to point out that the foregoing examples were contrived and embarrassingly trivial. Intuitively, as the problem size grows to the scale of complex systems, perfect consistency is predictably impossible in practice, due in large part to the nature of human judgment and numerous dependencies among indicators. Section 3 presents an example of a larger scale, which incorporates the steps of the procedure.

3. USE OF AHP TO ASSIGN WEIGHTS TO INDICATORS

An important link between the AHP and our indicator-based method of system design evaluation is the hierarchical structure of the problem. However, instead of a hierarchy of decision factors, we have a hierarchy of indicators. Applying the same techniques as in AHP, we obtain a method for determining relative weights of indicators in the assessment of the design.

COMFORT		C	T	L
	C	1	1/2	1/4
	T	2	1	1/2
	L	4	2	1

Figure 4. Perfectly Consistent Matrix.

COMFORT		C	T	L
	C	1	1/2	1/4
	T	2	1	1/4
	L	4	4	1

Figure 5. Less Consistent Matrix.

COMFORT		C	T	L
	C	0.13	0.11	0.17
	T	0.26	0.21	0.17
	L	0.52	0.84	0.66

Figure 6. Normalized Judgment Matrix.

3.1 Example Application of AHP to Indicator Weighting

Following the method of Saaty, we will state the concept to be evaluated, construct the hierarchy, then compute weights and consistency indices. Additionally, we provide an illustrative example.

1. The concept to be evaluated is: Relative weights of the contribution of human-factors elements to design evaluation.
2. Construction of indicator hierarchy: The hierarchy in Table 2 is an abridgment of numerous indicators culled from the literature. It is representative of human-related indicators pertinent at the highest levels of an integrated system.

Table 1. Consistency Indices for *n*-by-*n* Matrices [Saaty 1990].

n	1	2	3	4	5	6	7	8	9	10
CI	0.0	0.0	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

3. Beginning with the "leaf" indicators, we use the method of pairwise comparisons to construct the matrices of relative weight judgments. Using the judgment of the authors and the AHP tool *Expert Choice* [Saaty 1993], the we present the values in Figure 7 to depict the relative judgments of the highest level indicators. Similarly, we construct the matrices for the next level of indicators (Figures 8-10).
4. Compute the eigenvectors of the individual matrices to obtain the relative weights and consistency indices. Reconsider judgments which lead to intolerable consistency. For the matrices presented above, we compute the respective weights and consistency ratios by using the *Expert Choice* software product and present results in Figure 11 (a-d).
5. Working downward through the hierarchy, synthesize the individual relative weights to obtain local and global priorities of each indicator. These are included in Table 3.
6. Perform an overall consistency check of the entire hierarchy and make adjustments in relative judgments to compensate for poor consistency. This aggregate index is computed from the bottom up by multiplying the consistency index associated with a criterion by its priority, and summing across sibling criteria, etc. Given the consistency ratios associated with our complete tree, we use Saaty's techniques in *Expert Choice* and compute an overall consistency of 0.04.

We omit the matrices for the final level of indices; however, we use their values in the computation of their global weights, as well as the consistency of the entire hierarchy.

HUMAN FACTORS		Human	HSI	HHWI
	Human	1	2	2
	HSI	1/2	1	1
	HHWI	1/2	1	1

Figure 7. Top-Level Judgment Matrix.

Human		Perf	Dep	Cost	Sec
	Perf	1	3	5	1
	Dep	1/3	1	5	1
	Cost	1/5	1/5	1	1/4
	Sec	1	1	4	1

Figure 8. Human Indicator Judgment Matrix.

Table 2. High-level Hierarchy of Indicators.

LEVEL 0	LEVEL 1	LEVEL 2	LEVEL 3
Human Factors			
	Human		
		Performance	
			Skill
			Ability
			Understandability
			Aiding
		Dependability	
			Vigilance
			Availability
			Attitude
			Composure
			Teamwork
		Cost	
			Salary
			Benefits
			Bonuses
		Security	
			Clearance
			Training
			Integrity
	Human-Software Interface		
		Usability	
			Function
			Anthropometrics
			Interface
			Compatibility
		Dependability	
			Stability
			Accuracy
			Availability
			Redundancy
		Configurability	
			Facility
			Flexibility
			Time Required
		User Preference	
			Aesthetics
			User Friendly
			Work Relief
			Work Added
		Cost	
			Design
			Develop
			Test
			Maintain
		Security	
			Physical Access
			Encryption
			Password
	Human-Hardware Interface		
		Usability	
			Function
			Anthropometrics
			Interface
			Compatibility
		Dependability	
			Stability
			Accuracy
			Availability
			Redundancy
		Configurability	
			Facility
			Flexibility
			Time Required
		User Preference	
			Aesthetics
			User Friendly
			Work Relief
			Work Added
		Cost	
			Design
			Develop
			Test
			Maintain
		Security	
			Physical Access
			Encryption
			Password

HSI		Usab	Dep	Conf	Pref	Cost	Sec
	Usab	1	3	7	6	3	1/2
	Dep	1/3	1	4	4	3	1/4
	Conf	1/7	1/4	1	1	1/2	1/6
	Pref	1/6	1/4	1	1	1/2	1/5
	Cost	1/3	1/3	2	2	1	1/5
	Sec	2	3	6	5	5	1

Figure 9. HSI Indicator Judgment Matrix.

HHWI		Usab	Dep	Conf	Pref	Cost	Sec
	Usab	1	3	5	2	2	1/2
	Dep	1/3	1	4	5	1/2	1/4
	Conf	1/5	1/4	1	2	1/3	1/6
	Pref	1/2	1/5	1/2	1	1/2	1/7
	Cost	1/2	2	3	2	1	1/5
	Sec	2	4	6	7	5	1

Figure 10. HHWI Indicator Judgment Matrix.

3.2 Additional Issues

As demonstrated here, the application of the basic AHP methodology is straightforward, especially with the aid of a software tool such as *Expert Choice* used here. However, there are issues inherent in the nature of complex systems which beg additional attention. Not the least of these issues is the selection and structuring of the indicators. Also, the size of an indicator hierarchy for a system with the complexity of, e.g., a naval weapon system, is expected to be very large. Additionally, there are likely to be interdependencies among system indicators. Finally, the issues of group decisions and conflict resolution are certainly to be considered. These issues have been addressed in the literature [Harker 1987a,b; Saaty 1990, 1987, 1980] and are within the capabilities of the AHP and its extensions, albeit not without added cost and complexity. We address these issues below.

3.2.1 Choice and Structure of Indicators

The selection of the indicators and the construction of the hierarchy is not necessarily a simple task. It can require in-depth knowledge, at increasing degrees of granularity, of the system being designed. Such knowledge is not found in a single individual. It is likely to be found in an interdisciplinary workgroup of domain-specific experts. Depending on the system, the group could likely be comprised of software engineers, hardware engineers, and human factors specialists, as well as warfare experts.

The task of the workgroup would be to select the appropriate indicators for the particular level of abstraction at which they are evaluating the design. This could take place separately at the highest levels, intermediate levels, and so on. Saaty [1990] suggests that the AHP can even be used in the decision to keep or discard some of a collection of indicators. Individual preferences from group members would be averaged and used to weight

(a) Human Factors

Human	HSI	HHWI		CR
0.5	0.25	0.25		0.000

(b) Human

Perf	Dep	Cost	Sec		CR
0.415	0.235	0.064	0.286		0.060

(c) HSI

Usab	Dep	Conf	Pref	Cost	Sec		CR
0.283	0152	0.044	0.047	0.078	0.395		0.036

(d) HHWI

Usab	Dep	Conf	Pref	Cost	Sec		CR
0.222	0.129	0.053	0.052	0.129	0.415		0.083

Figure 11. Indicator Weights and Consistency Ratios.

Table 3. Computed Priorities of All Indicators.

	Local Priority	Local/Global		Local Priority	Global Priority
Human	0.500				
	Performance	0.4147			
		0.2070	Skill	0.5480	0.1136
			Ability	0.2800	0.0581
			Understandability	0.0860	0.0178
			Aiding	0.0860	0.0178
	Dependability	0.2352			
		0.1176	Vigilance	0.3300	0.0388
			Availability	0.3960	0.0466
			Attitude	0.0460	0.0054
			Composure	0.1320	0.0155
			Teamwork	0.0970	0.0114
	Cost	0.0638			
		0.0319	Salary	0.7090	0.0226
			Benefits	0.1790	0.0057
			Bonuses	0.1130	0.0036
	Security	0.2863			
		0.1432	Clearance	0.2380	0.0341
			Training	0.6250	0.0895
			Integrity	0.1360	0.0195
HSI	0.250				
	Usability	0.2833			
		0.0708	Function	0.5640	0.0400
			Anthropometrics	0.1920	0.016
			Interaction	0.1790	0.0127
			Compatibility	0.0650	0.0046
	Dependability	0.1518			
		0.0376	Stability	0.1220	0.0046
			Accuracy	0.3670	0.0139
			Availability	0.3960	0.0150
			Redundancy	0.1140	0.0043
	Configurability	0.0442			
		0.0110	Facility	0.4000	0.0044
			Flexibility	0.2000	0.0022
			Time Required	0.4000	0.0044
	User Prefererence	0.0474			
		0.0118	Aesthetics	0.0490	0.0006
			User Friendliness	0.2050	0.0024
			Work Relief	0.3980	0.0047
			Work Added	0.3480	0.0041
	Cost	0.0779			
		0.0195	Design	0.0880	0.0017
			Develop	0.2830	0.0055
			Test	0.2620	0.0051
			Maintain	0.3670	0.0072
	Security	0.3954			
		0.0988	Physical Access	0.0570	0.0056
			Encryption	0.5970	0.0591
			Password	0.3460	0.0342
HHWI	0.250				
	Usability	0.2215			
		0.0554	Function	0.6090	0.0337
			Anthropometrics	0.1730	0.0096
			Interaction	0.0790	0.0044
			Compatibility	0.1390	0.0077
	Dependability	0.1293			
		0.0323	Stability	0.2500	0.0081
			Accuracy	0.2500	0.0081
			Availability	0.2500	0.0081
			Redundancy	0.2500	0.0081
	Configurability	0.0530			
		0.0133	Facility	0.6000	0.0080
			Flexibility	0.2000	0.0027
			Time Required	0.2000	0.0027
	User Preference	0.0520			
		0.0123	Aesthetics	0.0550	0.0007
			User Friendliness	0.0840	0.0011
			Work Relief	0.4230	0.0055
			Work Added	0.4380	0.0057
	Cost	0.1295			
		0.0324	Design	0.0880	0.0028
			Develop	0.1300	0.0042
			Test	0.2110	0.0068
			Maintain	0.5710	0.0186
	Security	0.4146			
		0.1037	Physical Access	0.3330	0.0345
			Encryption	0.3330	0.0345
			Password	0.3330	0.0345

competing indicators. Those finishing in some agreed upon percentile would be deemed "keepers." The others could be discarded or used otherwise, e.g., as tie-breakers or at another level of granularity.

Group judgments are discussed in [Saaty 1990; 1980] and are supported by *Expert Choice*.

3.2.2 Large Numbers of Indicators

Some simple arithmetic gives insight into the cost of using AHP when large numbers and deeply nested indicators are involved. The number of comparisons among n sibling indicators is $(n)(n\text{-}1)/2$.

If we let n be the average number of children per parent node and d be a depth of the hierarchy (root at d=1), the number of pairwise comparisons required is $(n^d\text{-}n)/2$. Table 4 below reveals the fast growing nature of this value. In the table, the number of children per parent increases downward and depth in the tree increases to the right. To obtain the values in the table, we assume a complete tree of n nodes per parent at each depth. For example, a complete hierarchy with a depth of five (5) and an average of five (5) child indicators per parent indicator would require at worst case 1560 pairwise comparisons.

Fortunately, the worst case need not always prevail. In many example hierarchies provided in [Saaty 1990, 1980] and elsewhere in the literature, it is likely that the hierarchy is not necessarily full, reducing the number of required comparisons. Regardless, when the number of comparisons to be made becomes larger than would appear manageable, alternative means for completing massive comparison matrices or matrices with missing judgments have been derived [Harker 1987a]. Additional alternatives are illustrated in [Saaty 1990], including dividing judgments among a panel of decision makers and then combining them to obtain a complete evaluation matrix.

3.2.3 Interdependencies Among Indicators

Saaty [1990, 1987] discusses two types of interdependencies among hierarchical components -- additive and synergistic. In the *additive* case, contributions of interdependent components are simply summed. In the *synergistic* case, "the impact of the interaction of the elements is greater than the sum of the impacts of the elements, with due consideration given to their overlaps" [Saaty 1990, p. 90]. Accounting for such interdependence begins with creating new entities in the hierarchy for each interaction.

In the case of a design involving humanware, hardware, and software, such interactions abound. We have represented these interactions in Figure 1 by new entities for the human-software interface (HSI), human-hardware interface (HHWI), and hardware-software interface (HWSI). We employ the AHP techniques to judge the degree to which each interaction element contributes to the whole. An additional entity can be created for the hyper-interaction of these six entities. This region of interaction is of particular interest at the highest level view of system design evaluation. For our purposes, obtaining its measure of predicted success is an indirect indicator of the quality of the entire design.

3.2.4 Group Judgments and Conflict Resolution

The use of working groups in employing the AHP serves manifold purposes. Not the least of these is the opportunity to avail decision makers of alternative perspectives and possibilities, which may shed new light on the decision making process. Another key benefit is the opportunity to divide the work, reducing any attendant frustration associated with the application of AHP to a very large problem -- especially if the process is unfamiliar to the decision-making body. Still another benefit is the opportunity to mitigate the impact of extreme opinions based on biases, agendas, power struggles, etc. [Saaty 1990].

When conflicts do arise regarding the assignment of a pairwise comparison, and the consequences are significant, the AHP itself can be applied to a resolution. The simplest method is to employ an averaging algorithm to obtain a mean of

Table 4. Number of Pairwise Comparisons.

Nodes/Depth	2	3	4	5	6	7
3	3	12	39	120	363	1092
4	6	30	126	510	2046	8190
5	10	60	310	1560	7810	39060
6	15	105	645	3885	23325	139965
7	21	168	1197	8400	58821	411768
8	28	252	2044	16380	131068	1048572
9	36	360	3276	29520	265716	2391480
10	45	495	4995	49995	499995	4999995

the competing judgments. However, with most averaging techniques, the true judgments are no longer represented in the outcome.

Another method is the application of the AHP to the determination of the relative worth of each judge's input. This technique should be applied with prudence, and is most effective when all parties are objective and agreeable to the process. The usual influences of power, politics, and position must be mitigated as much as possible for the technique to have valid results. Here is the basic process.

1. Set as a goal the assignment of an acceptable judgment to the element(s) in question.
2. Construct a hierarchy (as abbreviated as possible) of the areas of expertise which are required to arrive at such a judgment.
3. Using crisp metrics as a basis for objectivity, (e.g., "5" versus "several"), make pairwise comparisons between the judges as they apply to each of the criterion in the hierarchy from Step 2.
4. Complete the matrices, compute the weights, and determine the consistency in the usual way.
5. When the matrices achieve consistency, the result should reveal the relative worth of each judge's input. Again, with prudence as a guide, this ordering can be applied to choose the value to assign to the judgment.

4. CONCLUDING REMARKS AND FUTURE RESEARCH

We have defined the problem under study as the need for an appropriate scheme to provide relative worth to indicators used in the evaluation of a complex system design. We propose the use of the Analytic Hierarchy Process as a suitable numerical technique and demonstrated its use by examples. We note and address the issues of problem size, problem structure, and conflict resolution when applying group judgments.

Considering the examples and discussion of additional issues presented here, we can conclude that the Analytical Hierarchy Process is a suitable vehicle for assigning weights to indicators as we have done. Though application of the process is not without what can be considerable effort, (especially for the evaluation of systems with large numbers of indicators), the incorporation of software aids such as *Expert Choice* eases the burden on the evaluators.

Regarding future research, we direct our attention to two primary fronts. First, in the vein of Paek, *et al*. [1992], we propose the application of fuzzy mathematics in the assignment and aggregation of indicator scores. This technique allows for the score assignment of either a range of quantitative values, or a qualitative value which can be translated into a fuzzy numerical value. The computation of a final "score" for the design will then be the fuzzy mathematical combination of the indicator scores, weighted by their relative worth as determined by application of the AHP. The ultimate judgment of the quality of the design can then be determined by the degree to which the aggregate score satisfies the design's governing requirements specification. Future research will focus on greater detail of the application of this technique and the derivation of the range of values applicable to a requirements specification.

Secondly, we propose the incorporation of a knowledge-based environment to assist the evaluators in both assigning weight judgments and to achieving greater consistency in the overall assignments. While *Expert Choice* eases the construction of judgment matrices and the computation of the weights and consistency values, there remains room for incorporation of knowledge from the experience of system designers and domain experts to more precisely attach relative judgments. We submit that an appropriate rule base can provide several types of assistance toward achieving this end. Rules can warn evaluators of judgments which differ significantly from those of other evaluators. Additionally a suitable rule base could be used to identify judgments which can lead to inconsistency. Finally, we foresee the use of a knowledge base to serve as an advisor to design evaluators, by providing access to inputs and comments appropriate to the type of system under evaluation, based on previous systems as well as the inputs of previous evaluators.

REFERENCES

Balci, O., D. DeVaux and R. Nance (1993), "Measurement and Evaluation of Complex Navy System Designs," In *Proceedings of the Complex Systems Engineering Synthesis and Assessment Technology Workshop (CSESAW '93)*, S.L. Howell and W. Farr (Eds.), Naval Surface Warfare Center Dahlgren Division, Dahlgren, VA, 126-140.

DeTurck, D.M. (1987), "An Approach to Consistency in the Analytic Hierarchy Process," *Mathematical Modelling 9*, 3-5, 345-352.

Harker, P. (1987a), "Alternative Modes of Questioning in the Analytic Hierarchy Process," *Mathematical Modelling 9*, 3-5, 353-360.

Harker, P. (1987b), "Incomplete Pairwise Comparisons in the Analytic Hierarchy Process," *Mathematical Modelling 9*, 11, 837-848.

Kaufmann, A. and M. Gupta (1985), *Introduction to Fuzzy Arithmetic: Theory and Applications*, Van Nostrand Reinhold Co., New York.

Lee, Y.W. and B.H. Ahn (1991), "Static Valuation of Combat Force Potential by the Analytic Hierarchy Process," *IEEE Transactions on Engineering Management 38*, 3, 237-244.

Masuda, T. (1990), "Hierarchical Sensitivity Analysis of Priority Used in Analytic Hierarchy Process," *International Journal of Systems Science 21*, 2, 415-427.

Murphy, C.K. (1993), "Limits on the Analytic Hierarchy Process from its Consistency Index," *European Journal of Operational Research 65,* 1, 138-139.

Paek, J.H., Y.W. Lee and T.R. Napier (1992), "Selection of Design/Build Proposal Using Fuzzy-Logic System," *Journal of Construction Engineering and Management 118*, 2, 303-317.

Saaty, J. (1993), *Expert Choice User Manual*, Expert Choice, Inc., McLean, VA.

Saaty, T.L. (1980), *The Analytic Hierarchy Process*, McGraw-Hill, Inc.

Saaty, T.L. (1987), "How to Handle Dependence with the Analytic Hierarchy Process," *Mathematical Modelling 9*, 3-5, 369-376.

Saaty, T.L. (1990), *Decision Making for Leaders*, RWS Publications, Pittsburgh, PA.

Srinivasan, V. and P.J. Bolster (1990), "Industrial Bond Rating Model Based on the Analytic Hierarchy Process," *European Journal of Operational Research 48*, 1, 105-119.

Toshtzar, M. (1988), "Multi-criteria Decision Making Approach to Computer Software Evaluation: Application of the Analytic Hierarchy Process," *Mathematical and Computer Modelling 11*, 276-281.

Weiss, E.N. (1987), "Using the Analytical Hierarchy Process in a Dynamic Environment," *Mathematical Modelling 9*, 3-5, 211-216.

Zahedi, F. (1986), "The Analytical Hierarchy Process: A Survey of the Method and its Applications," *Interfaces 16*, 4, 96-108.

AUTHOR INFORMATION

Michael Talbert (Capt, USAF) is a doctoral candidate researching a framework for complex system design evaluation in the Dept. of Computer Science at Virginia Tech. *(talbertm@csgrad.cs.vt.edu)*

Dr. Osman Balci is an associate professor of computer science at Virginia Tech, and principle investigator of the Navy-funded research project "Measurement and Evaluation of Complex Navy System Designs." *(balci@vt.edu)*

Dr. Richard E. Nance is professor of computer science and Director of the Systems Research Center at Virginia Tech. *(nance@vtopus.cs.vt.edu)*