Advancing the Use of an Analytical Hierarchy Process and Improved Random Indexes for Making Prioritized Decisions in Systems

Connett, Brian; OHalloran, Bryan M.; Pollman, Anthony G.

IEEE

Connett, Brian, Bryan M. OHalloran, and Anthony G. Pollman. "Advancing the Use of an Analytical Hierarchy Process and Improved Random Indexes for Making Prioritized Decisions in Systems." IEEE Transactions on Engineering Management (2019).
http://hdl.handle.net/10945/65003

Advancing the Use of an Analytical Hierarchy Process and Improved Random Indexes for Making Prioritized Decisions in Systems

Brian Connett, *Member, IEEE*, Bryan M. O'Halloran, and Anthony G. Pollman

***Abstract*—In the early stages of the systems engineering process, an important focus is to create an understanding of the stakeholder needs. This is primarily done to prepare the system specification that forms the basis for the system's design. By extension, example steps in this process include surveying stakeholders to better capture their intent, deriving and documenting requirements, and then using those requirements for subsequent activities, such as developing a functional baseline and candidate design alternatives. During this process, it is important to consider the full system lifecycle. As such, one major objective of a systems engineer is to translate the stakeholder's needs into functional and nonfunctional requirements (NFRs). Despite this important role, early system designs are often faulty because important NFRs are poorly prioritized or not prioritized at all. While the prioritization of all requirements can be useful, this work focuses specifically on NFRs. It has been identified that the inability to identify the most useful NFRs can lead to system failure. Furthermore, the lack of NFR prioritization is considered one of the most expensive and difficult errors to correct, as well as one of the ten most significant risks in engineering. Systems need more emphasis on the relationships between the system's elements, rather than on the individual elements or the whole system. Relationships among elements in a system can illustrate more than just the behavior of each element. The illustration can include the purpose for the system and the implications of changing how the NFRs associated with those elements are prioritized. This emphasis requires quantifiable tools and rigor to inform the decision makers. This research's objective is to contribute to quantifiable decision-making methods and prioritization of NFRs in three ways: the development of a process to determine unique random index; the use of a continuous ranking scale; and the development of a universal decision-making heuristic to accompany prioritization of NFRs.**

***Index Terms*—Analytical hierarchy process (AHP), consistency index (CI), consistency ratio (CR), engineering management, nonfunctional requirements (NFRs), prioritization, quality function, quality requirements, random index (RI), ratio, systems engineering (SE).**

Manuscript received February 18, 2019; revised July 17, 2019, September 25, 2019, and November 22, 2019; accepted November 23, 2019. Review of this manuscript was arranged by Department Editor T. Daim. *(Corresponding author: Brian Connett.)*

B. Connett is with the Department of Weapons, Robotics and Control Engineering, United States Naval Academy, Annapolis, MD 21012 USA (e-mail: connett@usna.edu).

B. M. O'Halloran and A. G. Pollman are with the Naval Postgraduate School, Monterey, CA 93943 USA (e-mail: bmohallo@nps.edu; agpollma@nps.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TEM.2019.2958785

I. Introduction

THE management of technical functions, such as research, development, and engineering in industry, government, university, and other settings, according to this journal [1], is used to emphasize the studies carried on within an organization to help in decision-making. The impact of the study presented here is through the interpretation of subject matter experts (SMEs) relative to each other and those needs of the customer. These interpretations are derived from desired behavior understanding of the system.

Adapting the principles presented here is important to the engineering management process from which the decision maker can provide informed prioritization of nonfunctional requirements (NFRs) and subsequently dictating the allocation of resources for system implementation. Ultimately, by applying the advanced analytical hierarchy process (AHP) heuristics to engineering management processes, this effort provides an updated tool set and management framework through which managerial understanding is significantly increased.

The use of AHP as a system design, engineering management, and decision-making tool is evident across several industries. AHP was introduced by T. L. Saaty. In his most referenced work, "What is the analytic hierarchy process?" [2], Saaty discusses how we make choices of "what tasks to do or not to do, when to do them, and whether to do them at all."

Saaty introduces AHP with examples of how those tasks show up in industry and academia, such as buying a home, choosing a school, buying business equipment, allocating funds within a government department, and voting on council issues. These are complex problems of choice, he says, that involve making logical decisions sometimes difficult for the human mind because of the number of factors and the effects in play simultaneously. The crux of the problem with making decisions, Saaty continues, is that humans make haphazard judgments or use models based on unverifiable assumptions. This leads to conclusions that may not be useful.

Developing hierarchical relationships can represent the organization needed for decision-making, but it is not sufficient. The judgments and measurements from stakeholders must be included. AHP is one of the answer's to this problem. As it was when introduced, AHP remains an appropriate scientific experiment for paired comparisons with relative values that are the same as the underlying physical laws dictating the system being measured. Even when underlying characteristics of the system are not quantifiable, AHP serves as a unique tool to harness subjective, or abstract comparisons using relevant facts about the relationships in the system. These relevant facts, in particular, are often derived from the instinct of the SMEs most familiar with the system. As a decision-making tool, AHP makes numerical value assignment to abstract concepts possible. In turn, the numerical assignment of ranking and prioritization can then be applied to a larger system of interest (SoI) in a way that makes measurable sense. The planning, prioritization, and resource allocation [3] that accompanies decision-making in complex systems [4] can be applied to business, economics, technology, energy, health, social, and transportation environments [5], [6]. Akin to decision, prediction, projection, and forecasting, AHP has proven to be a valuable tool in financial markets, political contests, games, contests, and sporting events [7].

The preceding examples are seemingly self-serving as they are presented by the original author of AHP. However, the applications do not end there. AHP follow-on research presents evidence of AHP value for decision-making. Market applications are some of the earliest examples of AHP usefulness, as presented by Wind and Saaty [8]. Albayrakoglu [9] presented a strategic approach to use AHP as a justification of new manufacturing technologies. The justification of advanced manufacturing technology is a complex problem that requires a number of tangible and intangible factors to be considered, especially at higher levels of integration. It is necessary to use multiple-criterion techniques, such as AHP to justify advanced manufacturing technology, since the problem involves a number of diverse and sometimes off-related attributes.

An adaptation of the AHP in [10] by Gawlik, Głuszak, and Małkowska focused on the elicitation of rental housing preferences. To illustrate the application of AHP to the problem of selection, Janic and Reggiani [11] examined a new aviation hub for a hypothetical airline assumed to operate within a liberalized air transport market. The application of AHP in this hypothetical is intended to lead to a preliminary judgment on its utility as additional decision-making tools for practical use. In [12], Bayazit relied on AHP make decisions regarding flexible manufacturing systems. He found that AHP was most useful for engineering management by determining the relative importance and influence of the most critical factors. In the postal industry, Chan *et al.* [13] applied AHP to benchmark logistics performance. This work led to a new benchmarking process for continuous improvement against market leaders. In outsourcing work to external agencies, AHP has also been used by Longaray *et al.* [14] to evaluate the quality of services. Their work focused intently on engineering management through the development of a decision support system in the retail industry. Using AHP, they concluded it was possible to develop an assessment model capable of measuring the performance of quality indicators and evaluating outsourced service products. AHP has also been shown to effectively integrate with other management and decision-making process. In [15], Singh *et al.* developed a strategy using AHP for selecting sustainable manufacturing. In particular, AHP allowed Singh to demonstrate that dynamic external influences can still be addressed in the engineering management aspect. AHP focused on the improvement in performances of economic, environmental, and social aspects of an organization. Moradi *et al.* [16] used AHP to evaluate the performance of the digital game industry. They premise that despite the high market performance of the game industry, there remains many examples of business failure. To design the evaluation system, the strategies and visions, and corresponding attributes and measures of the game companies are extracted and evaluated using AHP. The implementation of AHP decision-making tool provide a measure of efficiency for performance evaluation. Finally, and perhaps the current foremost authority on AHP is E. Forman. In both personal conversation (E. Forman, e-mail to author, June 2019) and through his work [17], Forman has conveyed the value of using AHP in engineering management and requirements prioritization. He uses many examples in his often-cited research and publicly available tool at expert choice.

Buede wrote that the basis of the system engineering process is the original set of requirements [18]. The INCOSE handbook relates that defining engineering requirements is one of the first steps in systems development [19]. However, system designs are often faulty because important NFRs are vague, not properly prioritized, or unaddressed. During the elicitation phase of the system engineering life cycle, NFRs are often overlooked, due in part to the "lack of understanding of NFR and the lack of effective NFR elicitation, modeling, and documentation methods" [20]. Due diligence must be paid by all stakeholders, ensuring an effective decision early in the design process, to inform the analysis effort that follows.

There are several processes to develop NFRs for a system. With specific relevance to this research, one approach is to select the requirements from a given set of candidate NFRs based on customer needs. For example, this selection process is often accomplished by dissecting systems specifications from similar systems as well as handbooks and standards. When a selection process is followed, the system is more likely to achieve preferences of the SME, more likely to retain the technical constraints, and more likely to maximize the overall business value [21].

While exploring possible design solutions, systems engineering (SE) tasks should focus on understanding stakeholder needs holistically and simultaneously. Developing possible design solutions gives life to the primary objective of a system engineer through all phases of the system engineering lifecycle [22]. Recall that the primary objective of a system engineer is to translate customer's needs into requirements, and then into a functioning system [23]. Achieving this objective can be enhanced, in part, with the prioritization of NFRs. Lubars *et al.* suggested that effective and systematic methods for prioritizing NFRs [24] are absent from SE and they link this to the labor and difficulty involved in prioritizing NFRs. Aurum and Wohlin [25] contended that there is not a process and the way to handle NFRs differs greatly among decision makers. As such, a new process that leads to prioritized NFRs is needed. No formal ranking and priority is in place to ensure that there is consensus among the SME rankings of the NFRs. A consensus exists when logical continuity or consistency exists. Consistency exists in the conformity in the way SMEs compare NFRs. This conformity is governed by the logic proposed by the decision makers related to the system. Consistency can be explained through a short illustrative narrative regarding the transitive property. If a decision maker prefers some system, System B, to another system, System A, then it can be said that System B has greater value than System A. Next, if the same decision maker prefers System A, to another system, System C, then it is said that System A has greater value than System C. Since System B is preferred to System A, and System A is preferred to System C, then logically we expect that System B is preferred to System C. This is representative of the transitive property and represents a consistency of ranking [26]. In contrast, the inconsistency is considered the total or partial absence of consistency. This research hypothesizes that SE projects often fail due to the scope creep and failure to meet NFRs. Therefore, if the NFRs can be prioritized and aligned as a consensus of SME NFR rankings, then the risk of project failure can be reduced [28], [29].

SE tasks during a standard design process are carried out using methods that link the technical processes [19], the needs of stakeholders, and the process of defining NFRs; the definition of system NFRs and the process of defining the architecture are identified in ISO15288 (a SE standard covering processes and lifecycle stages) and the INCOSE handbook. The way to execute these methods begins with the traditional SE elicitation, but the contribution of this work is revealed by the use of techniques found in many multicriteria decision-making methods (MCDMs) for building empirical consensus of ranking NFRs between SMEs. Several approaches have been attempted in prior investigations with limited scales of ranking and measurement. Despite some constraints of social order, such as the difficulty of reflecting the preference of a group derived from individual choices, consensus can be developed using rational and interval ranking methods. The system engineer can use the results of MCDM to ensure transparency of rankings and subsequent prioritization through a process known as negotiation or bargaining. The transparency ensures that decision makers make informed decisions with the use of clear and effective communication of information [36].

The AHP has suffered from poor acceptance in SE due to a lack of rigor. However, AHP remains in use in both research and education, demonstrating its core value. As shown throughout this work, the lack of rigor is apparent, especially in the calculation and statistical approach used to determine the primary measurement heuristic, the random index (RI). In practice, various ranking scales are used by decision makers but the choice of RI values is never changed, which leads to an inaccurate measurement of the agreement regarding the ways NFRs are ranked and prioritized. The RI is meant to be a reliable point of reference so that some valuable meaning can be gleaned from the ratio of consistency observed as an AHP output. However, the lack of rigor leads to variation of this AHP output.

There is no singularly correct requirements process in practice, and a cursory review of industry practices reveals that the functional requirements prioritization practices varies [25]. There are few effective or systematic requirements prioritization methods in practice to handle the complex prioritizations necessary [24]. The result of this challenge is that practitioners prioritize through various informal methods that fall short of consistently providing reliable information. Ranking and prioritization remain an elusive practice for NFRs. A variety of methods have been proposed for prioritization, including the numerical assignment technique [31], the MoSCoW technique [32], the priority groups technique [33], the bubble sort technique [34], the binary search tree technique [34], and the cumulative voting technique [35]. In general, these methods do not allow for the evaluation of the SMEs' rankings of the NFRs. The techniques are limited in the number of pairwise comparisons that can be analyzed, and they do not provide a method of transforming subjective SME evaluations to objective measurements.

This research focuses on AHP because of its appropriateness for handling intangibles, such as NFRs, the measurable prioritization outputs, its ability to test the consistency of a stakeholders' preferences, and the opportunity to provide a structure for implementation. This work specifically improves the aspect of AHP used for measuring the agreement of ranking NFRs.

This work is organized in the following way. First, Saaty's original RI is investigated and discussed along with the fundamental ranking scale he offered to the stakeholders. Next, a process is used to validate that a correct approach has been developed for quantifying an RI. The process used to validate the approach includes a discussion to determine the modeling resources necessary. The results of new RI calculation method are compared to the original Saaty RI to show that one RI size can fit all multisets. However, if the minimum and maximum values of those sets are expanded and allowed to be used by the SME, then the RI must be reevaluated to match SME allowable choices. To strengthen these findings, the ranking scales remain bounded at the maximum and minimum but are expanded, granularity of the fundamental ranking scale, RI values are included and larger pairwise comparison matrices are calculated. Note that it is possible to have an unbounded ranking scale, but the decision maker loses control of any standardization when the SMEs set out to rank the NFRs.

This work has led to several contributions including the following.

1) Development of a standard process that delineates the steps necessary to calculate RI and identifies the appropriate mathematical approximation.
2) Identification of the need to include a multiset measurement scale that always requires the inclusion of a "1" with a multiplicity of 2.
3) Determination that the RI is not dependent upon the set of the ranking scale, rather it is strictly dependent upon the size of the pairwise comparison matrix and the maximum and minimum boundary values.
4) Development of an RI scale that is more appropriate for the growing size of current systems. This includes a replication for RIs up to a matrix size, $n = 150$.

A. Research Objective

AHP [26] as a decision-making tool can be used to prioritize system NFRs and thereby improve the decision-making processes. Some traditional parameters of this decision-making tool have been in continual use since the introduction of AHP. The problem with these parameters is that they are not suitable for working with all systems. For example, an ordinal ranking parameter limits the granularity with which designers can rank system elements and relationships. Similarly, an accepted low value of an inconsistency parameter disregards the rapid and natural growth of systems.

The research objective is to address the issue of insufficiently ranking of NFRs for systems by developing an expanded ranking process. A ratio scale for ranking of NFRs is introduced using a continuous ratio scale, thus addressing the current granularity limitations of the traditional AHP. This higher level of sensitivity ranking allows elements to be ranked against each other with any choice of magnitude. Finally, a heuristic tool using the linear and equal interval ranking scale is presented that informs the decision-making process with a more granular level of information.

B. Research Method and Rationale

Literature survey, modeling and simulation, and grounded theory are the methodologies used to support this research. The literature review is used to locate and summarize relevant studies about the topics of: choice theory, modeling theory, utility theory, and decision-making. The literature survey includes research studies, conceptual articles, and opinion articles to provide this research's framework for considering the topic of SE and requirements engineering. Modeling and simulation is the quantitative approach used to develop the necessary data from which prioritization heuristics are developed. Grounded theory is part of the qualitative approach used to derive the general theory that a collection of negotiation methods for NFRs' prioritization is essential to developing a good system.

This study first employs exploratory sequential combined methods [27]. This qualitative research is meant to identify the prioritization order of NFRs that best meets the stakeholder's needs. After collection, the data are analyzed, interpreted, and validated. The nature of this type of SE decomposition is suitable for an embedded combined method [27], an advanced process that requires the repetitive use of quantitative and qualitative data embedded in the overall design.

Combined methods allow researchers and stakeholders to understand increasingly complicated relationships in a broader context [39]. Independently, each qualitative and quantitative method of study has received significant criticism from academia. The criticism of quantitative methods is that they fail to include subjective assertions from stakeholders [37]. Conversely, qualitative methods are criticized for failing to include objective analysis, scalability, and generalizability [38]. The broader context includes the subjective and objective analyses relevant to the motivation presented. Perhaps the most compelling rationale for using combined methods are the following from [46].

1) *Complementarity*: "Using data obtained by one method to illustrate results from another." For example, the data used in AHP are used to illustrate the effectiveness of using SMEs to rank NFRs.
2) *Development*: "Using results from one method to develop or inform the users of the other method." For example, the SMEs' subjective rankings are used to establish priority rankings among NFRs mathematically.
3) *Initiation*: "Using results from different methods specifically to look for areas of incongruence to generate new insights." Testing the subjective rankings from the SMEs against utility theory ensures rational decision-making.
4) *Expansion*: "Setting out to examine different aspects of a research question, where each aspect warrants different methods." The focus of this study is within the system design phase of the SE framework. From the requirements to detailed phase, the method ensures that the relative ranking of an NFR passes to subsequent phases.
5) *Triangulation*: "Using data obtained by both methods to corroborate findings." The corroboration contributes to the modeling and strengthens the findings by combining NFRs to develop aggregated and sufficient prioritization of those NFRs.

II. Background

A. What is the Analytical Hierarchy Process?

AHP is one of the MCDMs that was originally developed by T. L. Saaty. R. W. Saaty summarizes as follows.

The Analytic Hierarchy Process is a general theory of measurement. It is used to derive ratio scales from both discrete and continuous paired comparisons. These comparisons may be taken from actual measurements or from a fundamental scale which reflects the relative strength of preferences and feelings. AHP has a special concern with departure from consistency, its measurement and on dependence within and between the groups of elements of its structure. It has found its widest applications in multicriteria decision-making, planning and resource allocation and in conflict resolution. In its general form AHP is a nonlinear framework for carrying out both deductive and inductive thinking without use of the syllogism by taking several factors into consideration simultaneously and allowing for dependence and for feedback, and making numerical tradeoffs to arrive at a synthesis or conclusion. T. L. Saaty developed AHP in 1971–1975 while at the Wharton School (University of Pennsylvania, Philadelphia, PA, USA) [30].

B. Terms of AHP

In the understanding of AHP and associated parameters, the literature presents several ways of naming characteristics and variables. Ranking scales are used to describe the set of values provided to the SMEs so that pairwise comparisons can be made between NFRs. The measurement scale is the set of values provided for the calculations of the RI. A consistency index (CI) is the value calculated from a pairwise comparison matrix using the chosen mathematical method (power method, geometric method mean, or eigenvector method). An RI is determined by populating many pairwise comparison matrices, accumulating those values and then calculating the average value. Some authors use the term mean CI, or mean random CI. To measure the level of inconsistency in a comparison matrix, a consistency ratio (CR) is used to represent the ratio of the RI and CI. The value calculated is the percentage of inconsistency that exists in the rankings.

Consistency is a term used to illustrate the existence of ranking conformity in the way SMEs compare NFRs. This conformity is governed by the logic proposed by the decision makers related to the system. Consistency can be explained through an illustrative narrative regarding the transitive property. To be consistent, for example, the ranking of three NFRs should be such that if NFR A is weakly or slightly favored to NFR B, and NFR B is moderately-plus favored to NFR C, then by the axiom of transitivity, NFR A is very, very strongly favored to C. While intransitivities are unlikely in a small example with few comparisons, the likelihood is greater with larger and more complex systems. Furthermore, inconsistency can be created even when the axiom of transitivity is met (i.e., $A > B > C$). This is due to the fact that an intensity of importance is assigned between comparisons. For example, if A's intensity of importance to B is 2, and B's intensity of importance to C is 4, then a numerically consistent comparison would expect to have A's intensity of importance to C as 8. However, if the SME indicates that A is intensely important to C by any other value then an inconsistency exists. This is representative of the transitive property and a consistency of decision [26]. Mathematically, Saaty found that in a reciprocal matrix of totally consistent comparison matrix, the largest eigenvalue is equal to the size (nxn) of the comparison matrix, or $\lambda_{\max} = n$. He determined that a measure of consistency of comparison matrix can be shown as (1), and called the CI to determine the deviation from total consistency. Saaty proposed using this CI by comparing it to an appropriate CI that he calls the RI. He randomly generated a reciprocal matrix 500 times from the ranking scale $\frac{1}{9}, \frac{1}{7}, \frac{1}{5}, \frac{1}{3}, 1, 3, 5, 7$, and 9, measured the CI of each randomly generated matrix, summed the 500 matrices, and used the average as the bench mark for each matrix size up size $n = 10$. With both the CI of the comparison matrix being considered and the RI of that same size matrix, Saaty proposed that the amount of consistency be measured using (2). He noted that if the ratio is less than or equal to 0.10, then the inconsistency in the comparisons is acceptable.

$$\mathrm{CI} = \frac{\lambda_{\max} - n}{n - 1} \tag{1}$$

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \tag{2}$$

To determine the relative importance of the NFRs being compared, the use of the principal eigenvalue and the normalized right-eigenvector is warranted. This right-eigenvector weighting is from the local weights calculated with respect to all other NFRs associated with the same parent node. From this, a priority vector is produced. The priority vector is a numerical ranking of the NFRs that indicates an order of priority among them reflecting intensity or priority as indicated by the ratios of the numerical values [26]. The priority vector shows relative weights among the things that we compare.

C. *How Others Have Contributed to AHP*

There are many researchers who provide amplification of the original AHP [43], [44], [51], [55], [56]. Those researchers address classic SE challenges. Specifically, the challenge of making decisions that involve intangibles that need to be traded off. The tradeoff, though, must be measured against the understood objectives of the stakeholders. Pairwise comparisons, rankings, and prioritization of NFRs by SMEs are processes that are used to measures those tradeoffs. The decision makers rely on those SMEs to provide a good ranking of the elements, so that a true priority representation can be provided to inform further decision-making.

The SMEs provide a complete collection of rankings that illustrate how much more one NFR dominates another NFR concerning any chosen system attribute. The inherent flaw of subjective human decision-making injects inconsistency into the comparisons. AHP provides a method to assess the inconsistency, using established mathematical rigorous concepts. A survey finds that follow-on practitioners of AHP are built on two distinct characteristics of AHP. The first is that the ranking scale is a linear ordinal scale and bounded to nine distinct values $\frac{1}{9}, \frac{1}{7}, \frac{1}{5}, \frac{1}{3}, 1, 3, 5, 7$, and 9. The second characteristic is that the level of comparison inconsistency was generally acceptable when found to be 10% or less. These two traits are still found in subsequent research, although much debate surrounds the value of strict compliance to these two characteristics.

Harker and Vargas defend some of the major criticism of AHP stating that the strict compliance to original ranking scale and the low inconsistency level creates a firm "theoretical foundation and is a viable, usable decision-making tool" [40]. With the unique set of axioms of AHP in mind, designers should not view the process as a subset of the traditional design methods. The common theme is that humans are inconsistent and that their inconsistency should be considered in a formal manner, rather than one that is *ad hoc*. They propose that future research leading to the development of efficient stand-alone implementations of AHP is a worthwhile venture.

Systems need more emphasis on the relationships between the system's elements, rather than on the individual elements, or the whole system [41]. In [42], Lootsma begins by addressing concerns found in systems design aggregation. He investigates the original process that disregards cross-ranking of subcriterion within a hierarchy of evaluation. Lootsma then summarizes his findings by addressing ways in which design methodologies can be enhanced when the processes of both top–down and bottom–up design are combined.

Donegan and Dodd [43] also argue the need to prioritize system elements. Like other AHP practitioners, Donegan and Dodd focus on a feature of AHP, which is used to check the consistency of the SME rankings, and subsequent prioritization of the elements ranked. They adopt Saaty's comparison of the CI of its matrix with the chosen RI from a matrix of that same order. Donegan and Dodd seek to capitalize on Saaty's

TABLE I
SAATY'S RI (SOURCE: [58])

n	1	2	3	4	5	6	7	8	9	10
RI	0.00	0.00	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49

self-acknowledged lack of statistically significant estimates. They discuss two statistical tools that future researchers might consider to build a reliable RI. The first tool necessary is the need to determine a good sample size so that an appropriate random value can be generated and analyzed. The second tool generates multiple samples for each system size so that the RI for that system-size equates to the mean of all of the samples [43], [56]. Exhaustive tests by other practitioners, using Saaty's ranking set, demonstrated very close to the exact RI for matrices of order 3 and 4. Before proceeding to larger systems Donegan and Dodd indicate that the "estimates by Saaty are systematically overlarge" [43]. Donegan and Dodd conclude that it "might make sense to standardize the raw data" [43] of the set. The final survey includes an examination by Salo and Hämäläinen into the measurement of preferences [51] in AHP. The work describes the effects caused by bounded and discrete ratio scales and that the process can be modified to produce improved value measurements. Still, opportunities exist to contribute to these works. In this research, a few areas are explored: the use of a near-zero scale, the use of an infinite real number ranking scale, and the introduction of a decision-making heuristic.

III. UNDERSTANDING SAATY'S RI

Saaty's original RI calculation is not well understood, and still receives significant academic attention to better understand its role in prioritization. Some challenges found in reviewing Saaty's work, which are specific to creating a process that quantifies the RI, include: determining the mathematical approximation method used, determining the membership used in the original ranking scale (i.e., the scale of ranking values from which the SME will choose), and measurement scale (i.e., the scale of ranking values from which the decision maker will use to develop a unique RI), and measuring the accuracy of the traditional process beyond the matrix size, $n = 10$. The purpose of this section is to understand the process to develop the original Saaty RI in AHP. An RI is calculated using the same scale (measurement scale) that the stakeholder will use (ranking scale) to populate the pairwise matrix. Saaty used the scale $\{\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ in the traditional RI estimation. He then generated a random matrix selecting from the ranking scale, of which each entry has a equal probability of being selected. For each size matrix, the corresponding CI is calculated. This process was completed 500 times for each size matrix, and then the mean of all those consistency indices is the RI for the matrix size. Refer to Table I. The next section replicates the original values provided by Saaty, given that these RI values have remained unchanged since their inception. A pairwise matrix is developed by a single person and represents the relative rankings between NFRs. A specific cell within the matrix represents the relative ranking between two specific NFRs.

Intensity of Importance	*Definition*	*Explanation*
1	Equal Importance	Two activities contribute equally to the objective
2	Weak or slight	
3	Moderate importance	Experience and judgement slightly favour one activity over another
4	Moderate plus	
5	Strong importance	Experience and judgement strongly favour one activity over another
6	Strong plus	
7	Very strong or demonstrated importance	An activity is favoured very strongly over another; its dominance demonstrated in practice
8	Very, very strong	
9	Extreme importance	The evidence favouring one activity over another is of the highest possible order of affirmation
Reciprocals of above	If activity *i* has one of the above non-zero numbers assigned to it when compared with activity *j*, then *j* has the reciprocal value when compared with *i*	A reasonable assumption
1.1–1.9	If the activities are very close	May be difficult to assign the best value but when compared with other contrasting activities the size of the small numbers would not be too noticeable, yet they can still indicate the relative importance of the activities.

Fig. 1. Saaty's fundamental ranking scale of absolute numbers. Source [30].

A. Use of the Fundamental Ranking Scale of Absolute Numbers

As mentioned, various ranking scales are used by decision makers leading to unreliable measures of consistency in AHP. To set the stage for how users apply a preference to a pair of NFRs, understanding Saaty's fundamental ranking scale of absolute numbers is necessary. Once understood, this work introduces an extension of the fundamental ranking scale that can be applied universally, to include the use of a multiset continuous ranking scale and a varied maximum and minimum ranking scale. While normally reserved for background and related work, the inclusion of this section here should highlight its critical contribution to the understanding and findings of this research.

The calculation of the RI was based on a ranking scale from Saaty's notion that humans are relatively comfortable with the use of a ranking system that has an absolute reference frame. The absolute reference in a ranking scale arises from some social standard that indicates a convenience and familiarity in ranking (i.e., good, better, and best); in contrast, a relative reference frame is used to explain something that is the best among a pair of NFRs (i.e., preferred or not preferred). For the purpose of creating an RI heuristic, Saaty based his calculations on the static fundamental ranking scale of absolute numbers.

Because Saaty states that he uses "a randomly generated reciprocal matrix using the scale l/9, l/8, ..., 1, ..., 9" [30], we must investigate how to employ this ranking scale. First, a nominal scale is used to apply linguistic definitions (e.g., equal importance, weak importance, very strong importance, and extreme importance). Second, an intensity of importance is associated with each of the linguistic definitions (e.g., $1 =$ equal importance, $5 =$ strong importance, and $9 =$ extreme importance). Third, the set of reciprocal values of the aforementioned intensities of importance are identified. If an NFR has a nonzero number assigned when compared to another NFR, then the second NFR has the reciprocal value when compared to the first. Fig. 1 summarizes Saaty's fundamental ranking scale.

IV. Developing a Process to Determine the RI

This section highlights a major contribution of this research. This effort addresses the need to model and/or replicate the original RI values to validate a process that produces all possible RIs. To do that effectively, the following process is provided that replicates the results of Saaty's original experiment [58]. The evolution of an RI, presented in this work, is intended to improve on the currently available RI. More specifically, this part of the research develops a detailed process to determine the RI for any ranking scale and matrix size n. If the proposed expansion of the fundamental ranking scale is to be used, then it follows that a measurement scale should be commensurate with that same expansion. That is to say, each unique ranking scale has its own unique RI. However, that has not been the approach followed in the traditional AHP, and it is evident in the literature review that the basic RI developed by Saaty is still used for AHP evaluations. Because this investigation and research reveal the process of determining an RI through exploratory data analysis, the following steps are provided in pseudocode as a contribution to those decision makers developing their own RI for use in AHP.

The process is delineated in Section IV-B with appropriate steps determined to adequately replicate Saaty's RI measurements. This process is a generic approach to determining the measurements. However, it is found through this research that specific conditions should be enacted if a universally applicable RI is desired. In particular, it is known that the iterations used to calculate the original RI was 500, and follow-on authors used repeated bootstrapping methods. However, using statistical parameters, the precise number of iterations can be identified, as discussed in Section IV-A. Before choosing the measurement scale for this process, it is important to understand how the original ranking scale was presented as discussed in Section V-A. Then, the understanding needs to be turned to how the process can be used with a continuous scale. The scale includes predefined minimum and maximum values of the ranking scales presented to the SMEs, as discussed in Section V-C. This interpretation of the ranking scale can now be presented in a way that provides granular options to the SME, as discussed in Section V-B. Finally, the process calls for an iteration through matrix sizes to foster scaling of the RI, as discussed in Section V-D.

A. Determining the Modeling Iterations Necessary to Identify the RI

This section identifies the number of simulation runs to determine RI. In the context of this work, a simulation run is defined as determining a single CI value. Therefore, this section equivalently determines the number of CI. Since the RI is the average CI, the idea is to understand how many CIs are needed to achieve an adequate RI. For clarity, the CI referenced here is based on the random selection of values from the ranking scale. This is in comparison to the traditional use of CI that is discussed, where CI is calculated based on the rankings provided by SMEs during the ranking process.

In light of this, there are two reasons for why this section is needed. First, the RI values published by Saaty are going to be compared to the values produced by the process in Section IV-B. In order to do this, given that Saaty's value are presented as scalars, the RI produced in this work must include a confidence interval. Otherwise, the two scalar RI values would have no basis for comparison. The second reason for this section is to offer insight into "when is enough" for practical users of AHP. Less emphasis is placed on this reason due to the abundance of computing resources available in most instances where RI would be quantified. That said, AHP is broadly used across many application areas, and therefore, this point is worth noting.

To determine the number of runs necessary using a Monte Carlo simulation, one must calculate a confidence interval of a particular width given a desired confidence level. A confidence interval is a range of values used to estimate the true value of a population parameter [54]. A confidence level is the probability $1 - \alpha$ that the confidence interval contains the population parameter during the estimation process that is repeated [54]. In constructing the intervals, most literature lists the critical assumption that the random variables must be independent and identically distributed.

To determine an RI for use in the method developed in this work, a model ratio or interval data must be used. The RI being developed is a unit less measurement and can be described on a target confidence interval width containing positive and real values, v: $\{v \in \mathbb{R} \mid v > 0\}$. Equation (3) is calculation for the confidence interval, where $\bar{x}$ is the sample mean, $z_{\frac{\alpha}{2}}$ is the value of the standard normal (alpha level's z-score for a two-tailed test), s is the sample standard deviation, and "# runs" is the number of runs necessary for the simulation. In the value of the standard normal, α is the probability of rejecting the null hypothesis when the null hypothesis is true, or simply it is the probability of making a wrong decision.

$$\text{Confidence Interval}: \bar{x} \pm \frac{z_{\frac{\alpha}{2}} s}{\sqrt{\text{\# runs}}} \tag{3}$$

Because the comparison of the importance of the NFRs is a representative model of human decision-making, the assumption can be made that the standard deviation will scale with the mean [53]. That is, models with matrix size $n = 131$ have larger standard deviation than models with matrix size $n = 3$. This scaling, (4) [54], is called the coefficient of variation (CV) and can be assumed to be constant for the purpose of adequately determining the number of runs of the model ($\#runs$).

$$\text{CV} = \frac{s}{\bar{x}} \tag{4}$$

These two assumptions can be combined mathematically with the standard equation for a confidence interval. According to Byrne [53], the process is as follows. If the interval width v is taken to be a multiplier of mean $v\bar{x}$, then the width of the confidence interval can be computed using (5), and then, solving for "# runs," the number of the model runs necessary, (6), and replacing $\frac{s}{\bar{x}}$ with (4) to determine (7).

$$v\bar{x} = z_{\frac{\alpha}{2}} \frac{s}{\sqrt{\text{\# runs}}} \tag{5}$$

$$\text{\# runs} = \left(\frac{z_{\frac{\alpha}{2}}}{v} \frac{s}{\bar{x}}\right)^2 \tag{6}$$

$$\text{\# runs} = \left(\frac{z_{\frac{\alpha}{2}}}{v} \text{CV}\right)^2 \tag{7}$$

TABLE II
OF MODEL RUNS NECESSARY AS A FUNCTION OF STATISTICAL PARAMETERS TARGET CONFIDENCE INTERVAL AND THE CV (0.1–1.6)

		Coefficient of Variation															
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.1	1.2	1.3	1.4	1.5	1.6
	0.005	1537	6147	13830	24586	38416	55319	75295	98345	124468	153664	185933	221276	259692	301181	345744	393380
	0.01	384	1537	3457	6147	9604	13830	18824	24586	31117	38416	46483	55319	64923	75295	86436	98345
	0.015	171	683	1537	2732	4268	6147	8366	10927	13830	17074	20659	24586	28855	33465	38416	43709
	0.02	96	384	864	1537	2401	3457	4706	6147	7779	9604	11621	13830	16231	18824	21609	24586
	0.025	61	246	553	983	1537	2213	3012	3934	4979	6147	7437	8851	10388	12047	13830	15735
	0.03	43	171	384	683	1067	1537	2092	2732	3457	4268	5165	6147	7214	8366	9604	10927
	0.035	31	125	282	502	784	1129	1537	2007	2540	3136	3795	4516	5300	6147	7056	8028
	0.04	24	96	216	384	600	864	1176	1537	1945	2401	2905	3457	4058	4706	5402	6147
	0.045	19	76	171	304	474	683	930	1214	1537	1897	2295	2732	3206	3718	4268	4857
	0.05	15	61	138	246	384	553	753	983	1245	1537	1859	2213	2597	3012	3457	3934
	0.055	13	51	114	203	317	457	622	813	1029	1270	1537	1829	2146	2489	2857	3251
	0.06	11	43	96	171	267	384	523	683	864	1067	1291	1537	1803	2092	2401	2732
	0.065	9	36	82	145	227	327	446	582	736	909	1100	1309	1537	1782	2046	2328
Target Confidence Interval (v)	0.07	8	31	71	125	196	282	384	502	635	784	949	1129	1325	1537	1764	2007
	0.075	7	27	61	109	171	246	335	437	553	683	826	983	1154	1339	1537	1748
	0.08	6	24	54	96	150	216	294	384	486	600	726	864	1014	1176	1351	1537
	0.085	5	21	48	85	133	191	261	340	431	532	643	766	899	1042	1196	1361
	0.09	5	19	43	76	119	171	232	304	384	474	574	683	802	930	1067	1214
	0.095	4	17	38	68	106	153	209	272	345	426	515	613	719	834	958	1090
	0.1	4	15	35	61	96	138	188	246	311	384	465	553	649	753	864	983
	0.105	3	14	31	56	87	125	171	223	282	348	422	502	589	683	784	892
	0.11	3	13	29	51	79	114	156	203	257	317	384	457	537	622	714	813
	0.115	3	12	26	46	73	105	142	186	235	290	351	418	491	569	654	744
	0.12	3	11	24	43	67	96	131	171	216	267	323	384	451	523	600	683
	0.125	2	10	22	39	61	89	120	157	199	246	297	354	416	482	553	629
	0.13	2	9	20	36	57	82	111	145	184	227	275	327	384	446	511	582
	0.135	2	8	19	34	53	76	103	135	171	211	255	304	356	413	474	540

The problem, however, is that there is no obvious way to know a model's CV [54]. Instead the decision makers should create a CV that is appropriate for the amount of variance acceptable for simulation [60]. Table II indicates the number of runs necessary to meets these statistical parameters.

B. Steps of a Process Designed to Develop RI

The following pseudocode was developed to replicate Saaty's work and provided to verify a repeatable process designed to produce an RI based on the appropriate mathematical formulations and appropriate sample iterations. This process loops through a randomly populated matrix a predetermined amount of times, with each iteration producing a CI measurement based on the measurement scale provided and with the chosen estimation process. The CIs are accumulated and then averaged at the conclusion of the predetermined iterations, and then averaged to provide an RI for that matrix size. This process is repeated until all matrix sizes are satisfied.

Inputs to this process are as follows:

1) measurement scale and membership;
2) number of runs necessary to match statistical parameters indicated in Table II;
3) number of matrices desired;

The enumerated process are as follows.

1) Initialize the ranking scale set.
2) Initiate iteration first loop using the maximum number of samples appropriate for each matrix size.
3) Initiate iteration second loop using the maximum matrix size (i.e., $n \times n$) desired.
4) Populate each matrix element on upper triangle with random values and immediately assign reciprocal value to the lower triangle.
5) Populate matrix diagonal with ones.
6) End second loop.
7) Calculate the maximum lambda value.
8) Calculate the CI. CI $= \frac{\lambda_{\max} - n}{n-1}$.
9) End first loop.
10) Calculate average CI. This is the RI for that matrix size.

C. Calculating the RI With an Established Confidence Interval and CV

This section is used to systematically compare RI values published by Saaty (i.e., those from Table I) to values produced using the process presented in Section IV-B. As such, this is a validation for the process presented in Section IV-B. This validation is done for all RI values of comparison matrix sizes 3 to 10, and for the ranking scale $\{\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

Despite not knowing the matrix entries used by Saaty in the original 500 runs to determine an RI, the aforementioned discussion on determining the number of simulation runs can be used as an approach to compare RI values. As an example, to establish a 99% target confidence interval with an average CV of 0.8, the chart indicates that 98 345 simulations will provide an adequate RI for comparison. Similarly, the same amount of simulations are needed if one desires to establish a 98% target confidence interval with an average CV of 1.6. The values used to develop the RIs, sampled from the ranking scale of $\{\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, are illustrated in Fig. 2. These values resulted in the RI value of 0.5237 in the confidence interval of 0.5174 to 0.5300, which includes Saaty's 0.52 RI value for matrix size $n = 3$.

Using this same process with 98 345 simulations, the additional RIs, specifically those associated with matrix sizes $n = 4$ to $n = 10$, are determined with associated confidence intervals and are presented in Table III. Each of the RIs are within the constructed confidence interval, indicating a correct process. The RI column are those values calculated by the process described to replicate Saaty's work. The confidence interval includes the constructed 99% interval for all 98 345 values from that simulation. The last column, Saaty RI, is the original value of RI calculated by Saaty in 1980.

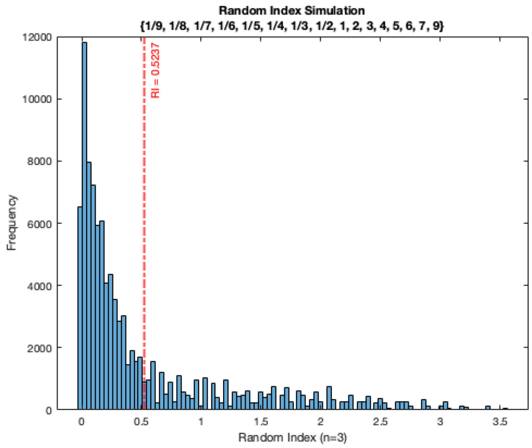

Fig. 2. RI simulation results for matrix size $n = 3$, 98 345 simulations from fundamental ranking scale of absolute measure. RI = 0.5237.

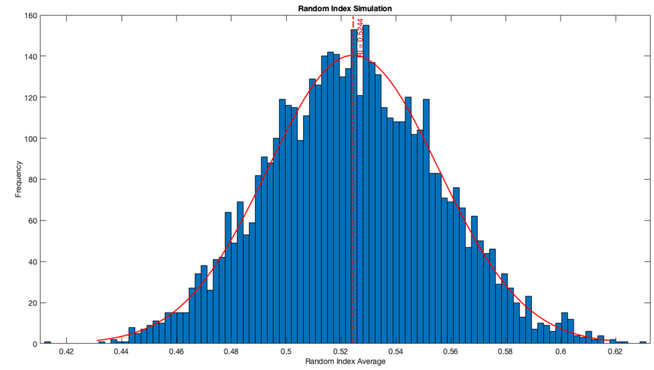

Fig. 3. RI simulation results for matrix size $n = 3$, 5000 simulations for 500 CIs on ranking scale $\{\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. RI = 0.5244.

TABLE III
CALCULATED CI ON FUNDAMENTAL RANKING SCALE
$\{\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$
FOR 98 345 SIMULATION RUNS

Matrix Size	Confidence Interval	RI	Saaty RI
3	0.5174 - 0.5300	0.5237	0.52
4	0.8797 - 0.8901	0.8849	0.89
5	1.1049 - 1.1133	1.1091	1.11
6	1.2466 - 1.2533	1.2499	1.25
7	1.3401 - 1.3455	1.3428	1.35
8	1.4010 - 1.4055	1.4033	1.40
9	1.4481 - 1.4521	1.4501	1.45
10	1.4837 - 1.4872	1.4854	1.49

TABLE IV
CALCULATED RI ON FUNDAMENTAL RANKING SCALE
$\{\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ FOR 5000
SIMULATION RUNS WITH 500 CIS AT EACH ITERATION

Matrix Size	Confidence Interval	RI	Saaty RI
3	0.5233 - 0.5255	0.5244	0.52
4	0.8833 - 0.8853	0.8843	0.89
5	1.1084 - 1.1100	1.1092	1.11
6	1.2487 - 1.2501	1.2494	1.25
7	1.3502 - 1.3513	1.3507	1.35
8	1.4036 - 1.4045	1.4041	1.40
9	1.4505 - 1.4513	1.4509	1.45
10	1.4857 - 1.4863	1.4860	1.49

D. *Statistical Comparison of the Process Results*

To provide a side-by-side statistical and value comparisons of RIs using the aforementioned simulation methods, the CI of each matrix size is calculated 500 times as was done by Saaty in his original work. Recall, a stakeholder is consistent if the following are satisfied [26].

1) $a_{ij} \cdot a_{jk} = a_{ik} \quad \forall i, j, k$.
2) $\lambda_{\max} = n$.
3) $\mathrm{CI} = 0$.

This process is then iterated 5000 times to determine the RI convergence. The RIs converge over many simulation runs resulting in the CI. Since Saaty only reports that he took 500 samples once to determine his RI, this process will help to validate his findings without knowing the reciprocal matrix entries he randomly selected.

The average of averages is an acceptable process in mathematics to develop a population parameter. If the number of elements of all groups is the same or when all the group averages are zero then the average of averages is equal to the average of all values. Consider two sets $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_n\}$ and their averages

$\overline{x} = \frac{\sum_{i=1}^n x_i}{n}, \overline{y} = \frac{\sum_{i=1}^m y_i}{m}$.

The average of the averages is

$$\text{average}(\overline{x}, \overline{y}) = \frac{\frac{\sum_{i=1}^n x_i}{n} + \frac{\sum_{i=1}^m y_i}{m}}{2} = \frac{\sum_{i=1}^n x_i}{2n} + \frac{\sum_{i=1}^m y_i}{2\tilde{}m}.$$

Now, consider $Z = \{x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n\}$ and its average

$$\overline{z} = \frac{\sum_{i=1}^n x_i + \sum_{i=1}^m y_i}{n + m}.$$

When the sets are the same size (i.e., $n = m$), then an average of the averages is an appropriate measurement tool. This notion of the average of averages is an acceptable tool to determine the convergence of a value to represent the population being analyzed. Using this approach, the process presented is run for 5000 simulations with 500 CIs calculated at each iteration. This results in the average RI. Fig. 3 illustrates the results of this simulation for a matrix size of 3.

Because Saaty originally calculated the RI value for each of the matrix sizes up to and including $n = 10$, those simulated values are provided in Table IV.

To check if the simulation is correct, the table includes the 99% confidence interval constructed around the simulated RI. It appears that Saaty's lack of significant digits is a result of rounding, although he never indicates as such. Still, the confidence intervals are provided alongside each of the RIs and demonstrate that the process is valid for replicating Saaty's original work. A side-by-side comparison is provided in Table V. This part of the process is provided to lend credibility calculating each RI with 98 345 simulations for each matrix size, and then using the average of those 98345 CIs to determine the RI. The original Saaty process for RI determination has now been confirmed. We understand how he developed the original RI and, therefore, can use this process to address limitations of AHP, such as

TABLE V
SIDE-BY-SIDE COMPARISON OF THE REPLICATED RIS USING TWO METHODS

Matrix Size, n	98,345 CI Averaged	5000 RIs @ 500 Averaged	Saaty's Original RI
3	0.5237	0.5244	0.52
4	0.8849	0.8843	0.89
5	1.1091	1.1092	1.11
6	1.2499	1.2494	1.25
7	1.3428	1.3507	1.35
8	1.4033	1.4041	1.40
9	1.4501	1.4509	1.45
10	1.4854	1.4860	1.49

Note: The first column is a method of determining the iterations necessary (98 345) to provide a high level of confidence with the smallest possible interval width to calculate individual CIs, then take one average for each matrix size. The second column provides the values calculated from 500 iterations averaged for each matrix size, then replicated 5000 times to develop the RI. Both processes serve to validate the process, ranking scale, and probability of matrix entry selection necessary to codify the unique RI process. The third column is the original published RI from Saaty.

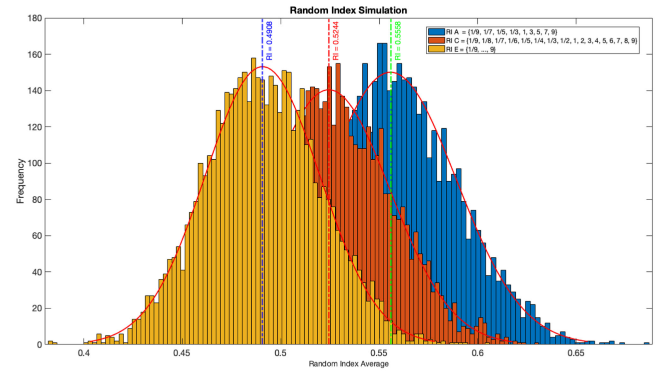

Fig. 4. RI simulation results for matrix size $n = 3$, 5000 simulations for 500 CIs using various ranking scales: $\{\frac{1}{9}, \frac{1}{7}, \frac{1}{5}, \frac{1}{3}, 1, 3, 5, 7, 9\}$, $\{\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5, 6, 7, 9\}$, and $\{\frac{1}{9}, \ldots, 9\}$.

ranking scales, measurement scales, and matrix size RIs for larger comparison matrices.

V. IMPROVING THE PROCESS TO DETERMINE THE RI

With the development of a standardized process regarding the calculation of the RI, some improvements are introduced. The following sections investigate four distinct improvements designed to provide the decision maker more flexibility and insight into the ranking and prioritization process. These improvements include a departure from the original fundamental ranking scale of absolute measures, examination of the ranking scale to include ranking granularity, an interpretation of the fundamental ranking scale of absolute measures, and the scaling of the heuristic to include larger matrices.

A. Departure From the Original Fundamental Ranking Scale of Absolute Measures

Saaty used the 17-element set of $\{\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5, 6, 7, 9\}$ for calculation of his published RI; he called this the fundamental ranking scale of absolute measures. There is evidence in the literature and textbooks that some practitioners of AHP offer ranking scales with fewer members, such as $\{\frac{1}{9}, \frac{1}{7}, \frac{1}{5}, \frac{1}{3}, 1, 3, 5, 7, 9\}$. One conjecture of this research was that changing the number of members from which the stakeholder can choose would require the recalculation of the heuristic to determine consistency of rankings. Using the validated process aforementioned in this section to replicate the original Saaty process, an investigation was conducted into varying the number of pairwise matrix entries from which to choose. Fig. 4 illustrates the three side-by-side calculations of RI with varying element membership (increasing discretization). Those sets that include the fewest entries bounded by values $\frac{1}{z}$ to z will converge to an RI value that is larger than a set bounded by the same values, but with a larger membership. The impact of using only one RI for all ranking scale sets is that the total consistency of the pairwise comparison will be miscalculated. If a decision maker relies on the same Saaty methodology for RI determination, they must also change their RI for consistency measurement. It is important that the decision makers take the time and effort to establish the RI commensurate with the bounds and discretization of the ranking scale provided to the SMEs. However, this approach is neither efficient nor necessary. Section V-C examines how to remedy the inefficiency.

TABLE VI
CALCULATED RIS FOR RI A: $\{\frac{1}{9}, \frac{1}{7}, \frac{1}{5}, \frac{1}{3}, 1, 3, 5, 7, 9\}$; RI C: $\{\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$; AND RI E: $\{\frac{1}{9}, \ldots, 9\}$

Matrix Size	RI A	RI C	RI E
3	0.5503	0.5254	0.4904
4	0.9406	0.8847	0.8255
5	1.1805	1.1087	1.0337
6	1.3306	1.2501	1.1629
7	1.4277	1.3420	1.2487
8	1.4955	1.4054	1.3070
9	1.5424	1.4500	1.3515
10	1.5815	1.4853	1.3837
11	1.6111	1.5136	1.4105
12	1.6346	1.5369	1.4325
13	1.6544	1.5548	1.4489
14	1.6706	1.5703	1.4639
15	1.6847	1.5844	1.4761
16	1.6977	1.5951	1.4868
17	1.7079	1.6050	1.4966
18	1.7178	1.6144	1.5050
19	1.7256	1.6223	1.5120
20	1.7327	1.6290	1.5190
21	1.7395	1.6355	1.5248
22	1.7459	1.6413	1.5302
23	1.7516	1.6464	1.5351
24	1.7560	1.6505	1.5394
25	1.7606	1.6554	1.5437
50	1.8139	1.7061	1.5914
100	1.8388	1.7312	1.6147
110	1.8413	1.7330	1.6163
120	1.8440	1.7345	1.6182
130	1.8460	1.7355	1.6201
140	1.8477	1.7387	1.6206
150	1.8489	1.7384	1.6221

Table VI and Fig. 5 are corresponding table and graph to demonstrate the differences in RI due to varying the ranking scale matrix entries. Specifically, this shows RI for three separate rankings scales: RI A: $\{\frac{1}{9}, \frac{1}{7}, \frac{1}{5}, \frac{1}{3}, 1, 3, 5, 7, 9\}$; RI C: $\{\frac{1}{9}, \frac{1}{8}, \frac{1}{7}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$; and RI E: $\{\frac{1}{9}, \ldots, 9\}$. Using the process presented in Section IV-B, any RI can be developed to inform the prioritization process.

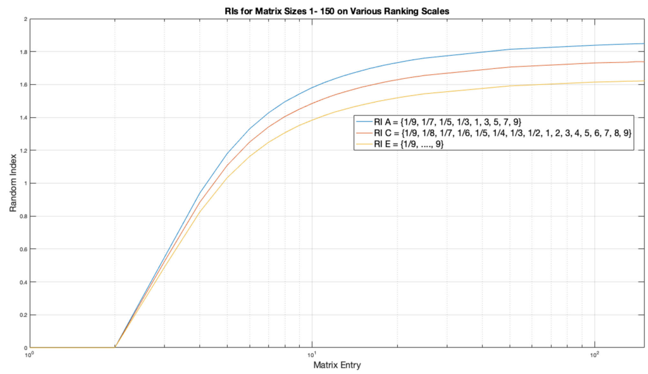

Fig. 5. RIs for matrix size 1–150 on various ranking scales.

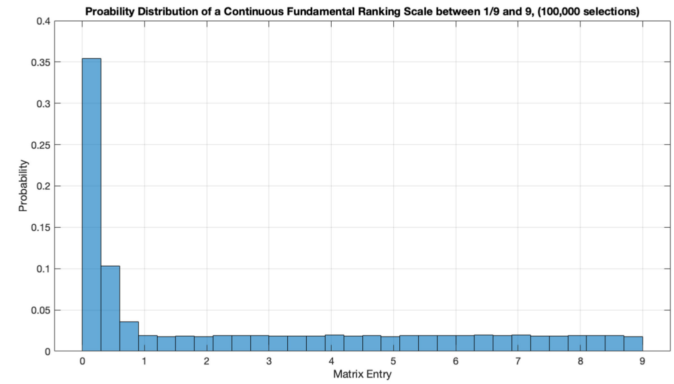

Fig. 6. Probability distribution of continuous ranking scale $\{\frac{1}{9}, \ldots, 9\}$.

B. Examining the RI Heuristic to Include Granularity From Fundamental Ranking Scale

The use of different scales of measurements beyond the fundamental ranking scale of absolute numbers is an important consideration when seeking opportunities to resolve conflicts or contradictions in rankings of NFRs. Suppose a decision maker wants to limit or expand the ranking scale membership provided to the SMEs for use during pairwise comparisons. The current published RIs do not compensate for such a change, although there is recurring evidence in the literature that the same RI scale is used against modified ranking scales to determine the amount of consistency in a comparison matrix.

When using a ranking scale of discrete values,ability of selecting a matrix entry is equal. In comparison though, if a granular ranking scale of continuous values is to be used, it is imperative that the distribution of those values be appropriate. For example, a continuous ranking scale such as $S = \{\frac{1}{z}, \ldots z\}$ should have an equal probability of selection on both sides of the median value, "1," where z is some value on the continuous number line. The goal here is not to define the probability density function (PDF) mathematically. Rather, the point is to ensure that when creating a new RI, the sample values for calculation be selected half of the time from a range from 1 to z and then half of the time be selected by from range of inverted vales $\frac{1}{z}$ to 1. The PDF is used to specify the probability of the random variable falling within a particular range of values, as opposed to taking on any one value. Fig. 6 illustrates this PDF for values $\frac{1}{9}$ to 9. Because the range from $\frac{1}{9}$ to 1 is smaller than the range of 1 to 9, it appears to build up on the left-hand side of 1. However, the left and right sides of the median value represent the same relative probability (50%) of selection. No absolute values for probability of selection can be inferred from this distribution.

C. Discussion on the Literal Interpretation of the Fundamental Ranking Scale of Absolute Measures (See Fig. 1) in the Calculation of Saaty's RI

Some researchers [43], [59] disagree on the utilization of the ranking scale for calculation of the RI. For example, if the set of fundamental rankings exists, $S_1 = \{a, b, c, d, e\}$, then it is also true that a second set of fundamental rankings exists, $S_2 = \{\frac{1}{e}, \frac{1}{d}, \frac{1}{c}, \frac{1}{b}, \frac{1}{a}\}$. The union of these two sets offer a set of discrete calculation values $S_1 \cup S_2 = S_3 = \{a, b, c, d, e, \frac{1}{e}, \frac{1}{d}, \frac{1}{c}, \frac{1}{b}, \frac{1}{a}\}$, and permit multiset representation with higher multiplicity for the value "1." This set combination provides the first insight into how follow-on authors investigating Saaty's RI might have misrepresented the original calculation of the RI. In [43], Donegan and Dodd surmised that "the estimates of Saaty are systematically overlarge" but they cannot provide conclusive evidence supporting this conjecture. Still, their explanation focuses on a perceived missed opportunity to include the value "1" in each of the sets, which provides a bias against the choice of "1" as a matrix entry. The choice to not include the reciprocal of "1" leads to a larger RI when compared to including the reciprocal. They conclude that because all of the possible matrix entry options must be equally likely, then the probability of selecting "1" must inherit a probability that is twice that of each other entry option. The use of "1" and "1/1," treated as distinct members for the reasons given by Donegan and Dodd in [43], changes the probability of selection for only the purpose of calculating the RI. More specifically, the inclusion of "1/1" in the ranking scale is not relevant to the decision maker when populating the pairwise comparison matrix because the reciprocal is equal to the original value and has no inherent linguistic value to the SME.

The interpretation of the way a fundamental ranking scale includes the value of "1" in its list of matrix entries and the statistical effect of that decision is worth noting. The discussion is important to determine if the inclusion of a repeated value will shift or skew the statistical representation of the population being analyzed. Using a discrete set of ordinal, interval, or ratio values will all have a uniform and equal probability of being selected in a random process as applied to emulate human choice.

The inclusion of "1" in the measurement scale to determine the RI is important and will have significant effect on the way that AHP is used to measure consistency of SMEs' rankings. When a bounded ranking scale is used (i.e., $\{\frac{1}{z}, \ldots z\}$), and a balanced set union that includes "$\frac{1}{1}$" and "1" for measurements is used, then there is no need for different RIs regardless of the original ranking scale membership. Fig. 7 shows the RI

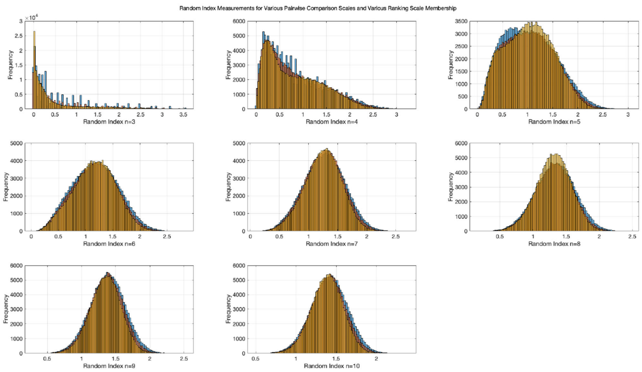

Fig. 7. RI simulation results for matrix size $n = 3$ to $n = 10$ on a ranking scale bounded by $\frac{1}{9}$ and 9 but with increasing membership of 10, 18, and infinite members in each plot.

TABLE VII
REVISED RI FOR INFINITE SCALE $\frac{1}{9}$ TO 9 REGARDLESS OF RANKING SCALE MEMBERSHIP

n	1	2	3	4	5	6	7	8	9	10
RI	0.00	0.00	0.50	0.83	1.04	1.17	1.26	1.32	1.36	1.40

simulation results of three ranking scales $\{\frac{1}{9},\ldots,\frac{1}{1},1,\ldots,9\}$, $\{\frac{1}{9},\frac{1}{8},\frac{1}{7},\frac{1}{6},\frac{1}{5},\frac{1}{4},\frac{1}{3},\frac{1}{2},\frac{1}{1},1,2,3,4,5,6,7,8,9\}$, and $\{\frac{1}{9},\frac{1}{7},\frac{1}{5},\frac{1}{3},\frac{1}{1},1,3,5,7,9\}$.

This is strong evidence that the original RI determined by Saaty used only the membership of his fundamental ranking scale of absolute measures. This poses considerable problems for any decision maker who chooses to use a ranking scale with a different amount of members. Instead, if $S_1 = \{a,b,c,d,e\}$, then it is also true that a second set of fundamental rankings exists, $S_2 = \{\frac{1}{e},\frac{1}{d},\frac{1}{c},\frac{1}{b},\frac{1}{a}\}$. The union of these two sets offer a set of discrete calculation values $S_1 \cup S_2 = S_3 = \{a,b,c,d,e,\frac{1}{e},\frac{1}{d},\frac{1}{c},\frac{1}{b},\frac{1}{a}\}$). By designing a simulation to include the appropriate union of ranking members and their reciprocals, we will be able to use a singular constant RI in the decision-making processes that only changes with the size of the comparison matrix, refer to Table VII.

D. Scaling the RI Heuristic to Include Larger Matrices

The use of RIs larger than those published ($2 \leq n \leq 10$) in the literature are necessary to inform the decision-making for applications to large and complex systems. An evaluation of matrices up to $n = 150$ is provided, see Table VIII.

The current limitations of the RI are tied to the design of those systems that require the use of tools capable of handling the prioritization of many NFRs but cannot due to the limited size of published heuristics. Military, commercial, and instances of personal-use systems present hundreds or thousands of individual requirements. It is difficult to prioritize such a large collection of requirements, however, this should not relegate a decision maker into groupings of requirements because of a limitation in the design tools available.

TABLE VIII
RI MATRICES ON THE INFINITE RANKING SCALE FOR $n = 1$ TO $n = 150$

n	RI		n	RI
1	0.00		21	1.53
2	0.00		22	1.53
3	0.49		23	1.53
4	0.83		24	1.54
5	1.03		25	1.54
6	1.16		30	1.56
7	1.25		35	1.57
8	1.31		40	1.58
9	1.35		45	1.59
10	1.38		50	1.59
11	1.41		60	1.60
12	1.43		70	1.60
13	1.45		80	1.61
14	1.46		90	1.61
15	1.48		100	1.61
16	1.49		110	1.62
17	1.50		120	1.62
18	1.51		130	1.62
19	1.51		140	1.62
20	1.52		150	1.62

VI. IN PRIORITIZING NFRS, HOW DOES ALL OF THIS HELP US?

The process to quantify RI is a critical calculation for AHP in that it helps to determine the total consistency of ranking of NFRs in a pairwise matrix. To be consistent, for example, the ranking of three NFRs should be such that if NFR A is weakly or slightly favored to NFR B, and NFR B is moderately plus favored to NFR C, then by the axiom of transitivity NFR A is very, very strongly favored to C. While intransitivities are unlikely in a small example with few comparisons, the likelihood is greater with a larger number of comparisons. Furthermore, inconsistency can be created even when the axiom of transitivity is met (i.e., $A \succ B \succ C$). This is due to the fact that a numerical preference is assigned between comparisons. For example, if A's intensity of importance to B is 2, and B's intensity of importance to C is 4, then a numerically consistent comparison would expect to have A's intensity of importance to C as 8. However, if the SME indicates that A is intensely important to C by any other value then an inconsistency exists. However, the method and heuristic with which to measure inconsistencies using AHP are still debated.

Saaty's original RI was investigated and discussed along with the fundamental ranking scale he offered to the stakeholder. To identify and validate that the correct process for determining the RI is developed, a process was presented along with a discussion to determine the modeling resources necessary. These results were compared to the original Saaty RI to show that one RI size can fit all membership multisets that are bounded by the same minimum and maximum boundary values. However, if those minimum and maximum values are expanded and allowed to

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE IX
AGD—CONTINUOUS RANKING SCALE $\frac{1}{9}, \ldots, 9$

	NFR 1	NFR 2	NFR 3	NFR 4	NFR 5	NFR 6	NFR 7
NFR 1	1.00	0.23	0.18	0.17	0.15	0.89	0.13
NFR 2	4.38	1.00	2.63	0.82	0.76	6.85	0.16
NFR 3	5.71	0.38	1.00	0.35	0.13	2.12	0.19
NFR 4	6.06	1.22	2.86	1.00	0.53	7.89	0.11
NFR 5	6.62	1.31	7.69	1.87	1.00	4.61	0.83
NFR 6	1.13	0.15	0.47	0.13	0.22	1.00	0.13
NFR 7	7.76	6.08	5.29	8.73	1.21	7.96	1.00

TABLE X
AGD—AHP 1 $\{\frac{1}{9}, \frac{1}{8}, \ldots, 8, 9\}$

	NFR 1	NFR 2	NFR 3	NFR 4	NFR 5	NFR 6	NFR 7
NFR 1	1.00	0.25	0.17	0.17	0.14	1.00	0.13
NFR 2	4.00	1.00	3.00	1.00	1.00	7.00	0.17
NFR 3	6.00	0.33	1.00	0.33	0.13	2.00	0.20
NFR 4	6.00	1.00	3.00	1.00	0.50	7.00	0.11
NFR 5	7.00	1.00	8.00	2.00	1.00	5.00	1.00
NFR 6	1.00	0.14	0.50	0.14	0.20	1.00	0.13
NFR 7	8.00	6.00	5.00	9.00	1.00	8.00	1.00

be used by the SME, then the RI must be reevaluated to maintain commensurate measure of SME rankings. To strengthen these findings, the bounds of the ranking scales were expanded, granularity of the fundamental ranking scale RI values were included, and larger pairwise comparison matrices were calculated.

This work has led to several contributions, which include the following:

1) development of a standard process that delineates the steps necessary to calculate RI and identifies the appropriate mathematical approximation;
2) identification of the need to include a multiset measurement scale that always requires the inclusion of a "1" with a multiplicity of 2;
3) determination that the RI is not dependent upon the membership of the ranking scale, rather it is strictly dependent upon the size of the pairwise comparison matrix and the maximum and minimum boundary values;
4) development of an RI scale that is more appropriate for the growing size of current systems. This includes a replication for RIs up to a matrix size, $n = 150$.

VII. APPLICATION IN EXISTING/CURRENT METHOD: ARTIFICIALLY GENERATED DATA (AGD)

A case study is generated from artificial data. The artificial data are created using the continuous set $\frac{1}{9}, \ldots, 9$. For comparison, this same data are then modified to fit the traditional ranking scales, AHP 1 : $\frac{1}{9}, \frac{1}{8}, \ldots, 8, 9$ and AHP 2: $\frac{1}{9}, \frac{1}{7}, \ldots, 7, 9$ used in the relevant research. The comparisons that will be made from these different scales will demonstrate that the measure of consistency and prioritization ranking can be apparent.

A. Preparing AGD for AHP

The precedence of using an AGD set has been established by prior authors. In particular, Kamishima and Akaho [61] generated 100 sample sets for the use in ranking comparisons. For each sample set, they ran sample algorithms five times using different clustering methods, and then the best cluster of data was selected. Similarly, artificial data are generated here to create a sample pairwise comparison matrix of size $n = 7$ with a CR that falls below 10%, this is shown in Table IX. The generation of data for this case study was achieved through a similar process that is used in Section IV-B. However, the difference is that the loop was terminated once a dataset was found to satisfy the 10% threshold for CR compared to the proposed universal RI.

B. Apply SDNPM to the AGD

Again, the axioms of completeness and continuity are the first checks of rationality included in Step 1 of the SDNPM. The third axiom, transitivity, is combined with the AHP check for magnitude transitivity using axiomatic design. The AHP measurements associated with Table IX are $\lambda_{\max} = 7.756$ and $\mathrm{CI} = 0.126$ (1) leading to $\mathrm{CR} = 0.099$ (2). If the threshold of consistency is satisfied with the calculated CR, then the process ends and a prioritized list of NFRs is produced. If the decision maker does not accept this measured level of consistency, then negotiation can be applied.

In negotiation, the axiom of transitivity is used as a reference tool to facilitate scrutiny of each of the pairs of comparisons when the threshold of inconsistency is exceeded. An example shows the NFR 5 intensity of importance to NFR 3 is 7.69 and the NFR 3 intensity of importance to NFR 6 is 2.12; however, the NFR 5 intensity of importance to NFR 6 is only 4.61. The expected transitive value for the intensity of importance of NFRs 5 to 6 is 16.30. This is a specific example of where a glaring inconsistency exists in the pairwise rankings, and a point of information is now available to the decision maker to initiate negotiation among SMEs.

For a comparative analysis regarding the choice of ranking scales, the artificial data from Table IX are manipulated next to create a sample pairwise comparison matrix of size $n = 7$. This uses the traditional ranking scale (AHP 1) provided by Saaty and is presented in Table X. The measurements associated with this table are $\lambda_{\max} = 7.870$ and $\mathrm{CI} = 0.145$ leading to $\mathrm{CR} = 0.107$. The first observation from these calculations is that the inconsistency of the rankings has increased simply by restricting the ranks from which the SME can choose. Even still, a second observation is reached if the limited ranking scale is permitted and compared to the proposed universal RI. When $\mathrm{RI} = 1.26$, which reflects the multiset for $\frac{1}{9}, \ldots, 9$, then $\mathrm{CR} = 0.115$. The traditional RI overestimates the total amount of consistency in the comparison rankings.

Continuing the comparative analysis, the artificial data from Table IX are manipulated once more here to create a sample pairwise comparison matrix of size $n = 7$ but using the second traditional ranking scale (AHP 2) provided by Saaty, this is shown in Table XI. The measurements associated with this table are $\lambda_{\max} = 7.942$ and $\mathrm{CI} = 0.157$ leading to $\mathrm{CR} = 0.116$. Again, viewed differently and using the proposed universal RI $\mathrm{RI} = 1.26$ that reflects the multiset for $\frac{1}{9}, \ldots, 9$ then $\mathrm{CR} = 0.125$. The conclusion is that the traditional RI overestimates the total amount of consistency in the comparison rankings.

TABLE XI
AGD—AHP 2 $\{\frac{1}{9}, \frac{1}{7}, \frac{1}{5}, \ldots, 5, 7, 9\}$

	NFR 1	NFR 2	NFR 3	NFR 4	NFR 5	NFR 6	NFR 7
NFR 1	1.00	0.20	0.14	0.14	0.14	1.00	0.14
NFR 2	5.00	1.00	3.00	1.00	1.00	7.00	0.14
NFR 3	7.00	0.33	1.00	0.33	0.14	2.00	0.20
NFR 4	7.00	1.00	3.00	1.00	0.33	8.00	0.11
NFR 5	7.00	1.00	7.00	3.00	1.00	5.00	1.00
NFR 6	1.00	0.14	0.50	0.13	0.20	1.00	0.14
NFR 7	7.00	7.00	5.00	9.00	1.00	7.00	1.00

TABLE XII
PRIORITIZATION OF AGD CASE STUDY

	Continuous	AHP 1	AHP 2
NFR 7	0.386	0.376	0.369
NFR 5	0.225	0.232	0.235
NFR 4	0.136	0.131	0.129
NFR 2	0.128	0.139	0.142
NFR 3	0.066	0.065	0.068
NFR 6	0.031	0.030	0.030
NFR 1	0.027	0.027	0.026

Table XII summarizes the prioritization of the AGD case study. Clearly, the prioritization values are distributed differently among the NFRs. Particular attention should be paid to the change in priorities that occurs with NFRs 2 and 4. Limiting the ranking scale granularity and membership has an outward effect on the prioritization of the NFRs.

VIII. CONCLUSION

AHP is a valuable tool to inform the decision-making process. Some traditional parameters have been used since the introduction of the decision-making tool. The parameters, ordinal rankings and a 10% inconsistency allowance, have been identified by the authors in related research and require extensive consideration and evaluation. The authors of this research have taken the steps to address deliberate ranking of NFRs using a modified ranking process and a new measurement heuristic. These two contributions allow elements to be ranked against each other with a sensitivity-level of magnitude comparison appropriate to the needs of the system designer.

REFERENCES

[1] IEEE Transactions on Engineering Management, "ABOUT IEEE TEM," Accessed: Dec. 5, 2019. [Online]. Available: https://www.ieee-tems.org/ieee-transactions-on-engineering-management/

[2] T. L. Saaty, "What is the analytic hierarchy process?" in *Mathematical Models for Decision Support*. Berlin, Germany: Springer, 1988, pp. 109–121.

[3] T. L. Saaty, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. New York, NY, USA: McGraw-Hill, 1980.

[4] T. L. Saaty, *Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World*. Pittsburgh, PA, USA: RWS Publications, 1990.

[5] T. L. Saaty and L. G. Vargas, *The Logic of Priorities: Applications of Business, Energy, Health and Transportation*. Berlin, Germany: Springer, 2013.

[6] T. L. Saaty and L. G. Vargas, *Decision Making in Economic, Social and Technological Environments*. Pittsburgh, PA, USA: RWS Publications, 2006.

[7] T. L. Saaty and L. G. Vargas, *Prediction, Projection, and Forecasting (Applications of the Analytic Hierarchy Process in Economics, Finance, Politics, Games, and Sports)*. Norwell, MA, USA: Kluwer, 1991.

[8] Y. Wind and T. L. Saaty, "Market applications of the analytic hierarchy process," *Manage. Sci.*, vol. 26, no. 7, pp. 641–658, 1980.

[9] M. M. Albayrakoglu, "Justification of new manufacturing technology: A strategic approach using the analytic hierarchy process," *Prod. Inventory Manage. J.*, vol. 37, no. 1, pp. 71–76, 1996.

[10] R. Gawlik, M. Głuszak, and A. Małkowska, "The measurement of housing preferences in the analytic hierarchy process," *Folia Oeconomica Stetinensia*, vol. 17, no. 1, pp. 31–43, 2017.

[11] M. Janic and A. Reggiani, "An application of the multiple criteria decision making (MCDM) analysis to the selection of a new hub airport," *Eur. J. Transp. Infrastruct. Res.*, vol. 2, no. 2, pp. 113–141, 2002.

[12] O. Bayazit, "Use of AHP in decision-making for flexible manufacturing systems," *J. Manuf. Technol. Manage.*, vol. 16, no. 7, pp. 808–819, 2005.

[13] F. T. S. Chan, H. K. Chan, H. C. W. Lau, and W. L. I. P. Ralph, "An AHP approach in benchmarking logistics performance of the postal industry," *Benchmarking, Int. J.*, vol. 13, no. 6, pp. 636–661, 2006.

[14] A. A. Longaray, J. de Deus Rodrigues Gois, and P. R. da Silva Munhoz, "Proposal for using AHP method to evaluate the quality of services provided by outsourced companies," *Proc. Comput. Sci.*, vol. 55, pp. 715–724, 2015.

[15] S. Singh, E. U. Olugu, S. N. Musa, A. B. Mahat, and K. Y. Wong, "Strategy selection for sustainable manufacturing with integrated AHP-VIKOR method under interval-valued fuzzy environment," *Int. J. Adv. Manuf. Technol.*, vol. 84, no. 1–4, pp. 547–563, 2016.

[16] N. Moradi, H. Malekmohammad, and S. Jamalzadeh, "A model for performance evaluation of digital game industry using integrated AHP and BSC," *J. Appl. Res. Ind. Eng.*, vol. 5, no. 2, pp. 97–109, 2018.

[17] E. N. D. Forman, "The math of AHP—Computing priorities from pairwise comparisons." Accessed: Dec. 5, 2019. [Online]. Available: http://professorforman.com/

[18] D. M. Buede, "Developing originating requirements: Defining the design decisions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, no. 2, pp. 596–609, Apr. 1997.

[19] A. Incose, "A world in motion: Systems engineering vision 2025," in *Proc. Int. Council Syst. Eng.*, San Diego, CA, USA, 2014, pp. 1–30.

[20] A. Mahmoud and G. Williams, "Detecting, classifying, and tracing non-functional software requirements," *Requirements Eng.*, vol. 21, no. 3, pp. 357–381, 2016.

[21] G. Ruhe, A. Eberlein, and D. Pfahl, "Quantitative WinWin: A new method for decision support in requirements negotiation," in *Proc. 14th Int. Conf. Softw. Eng. Knowl. Eng.*, Jul. 2002, pp. 159–166.

[22] C. Haskins, K. Forsberg, M. Krueger, D. Walden, and D. Hamelin, "Systems engineering handbook," in *Proc. INCOSE*, Jun. 2006, p. 34.

[23] B. S. Blanchard, W. J. Fabrycky, and W. J. Fabrycky, *Systems Engineering and Analysis*, vol. 4. Englewood Cliffs, NJ, USA: Prentice-Hall, 1990.

[24] M. Lubars, C. Potts, and C. Richter, "A review of the state of the practice in requirements modeling," in *Proc. IEEE Int. Symp. Requirements Eng.*, Jan. 1993, pp. 2–14.

[25] A. Aurum and C. Wohlin, "The fundamental nature of requirements engineering activities as a decision-making process," *Inf. Softw. Technol.*, vol. 45, no. 14, pp. 945–954, 2003.

[26] T. L. Saaty, "Decision making with the analytic hierarchy process," *Int. J. Services Sci.*, vol. 1, no. 1, pp. 83–98, 2008.

[27] J. W. Creswell, *Research Design: Qualitative and Quantitative Approach*. Thousand Oaks, CA, USA: Sage, 1996.

[28] M. Dabbagh and S. P. Lee, "An approach for integrating the prioritization of functional and nonfunctional requirements," *Sci. World J.*, vol. 2014, 2014, Art. no. 737626.

[29] R. K. Chopra, V. Gupta, and D. S. Chauhan, "Experimentation on accuracy of non functional requirement prioritization approaches for different complexity projects," *Perspectives Sci.*, vol. 8, pp. 79–82, 2016.

[30] R. W. Saaty, "The analytic hierarchy process—What it is and how it is used," *Math. Model.*, vol. 9, nos. 3–5, pp. 161–176, 1987.

[31] I. Sommerville and P. Sawyer, *Requirements Engineering: A Good Practice Guide*. Hoboken, NJ, USA: Wiley, 1997.

[32] D. Tudor and G. A. Walter, "Using an agile approach in a large, traditional organization," in *Proc. AGILE*, 2006, pp. 1–7.

[33] J. Karlsson, C. Wohlin, and B. Regnell, "An evaluation of methods for prioritizing software requirements," *Inf. Softw. Technol*, vol. 39, no. 14–15, pp. 939–947, 1998.

[34] A. V. Aho, J. E. Hopcroft, and J. D. Ullman, *Data Structures and Algorithms*. London, U.K.: Pearson Education, 1983.

[35] P. B. Andrews and A. Amschler, "Engineering and managing software requirements," *Requirements Prioritization*, pp. 69–94, 2005.
[36] M. Hosseini, A. Shahri, K. Phalp, and R. Ali, "Engineering transparency requirements: A modelling and analysis framework," *Inf. Syst.*, vol. 74, pp. 3–22, 2018.
[37] A. Toomela, "Culture of science: Strange history of the methodological thinking in psychology," *Integrative Psychol. Behav. Sci.*, vol. 41, no. 1, pp. 6–20, 2007.
[38] O. Gelo, D. Braakmann, and G. Benetka, "Quantitative and qualitative research: Beyond the debate," *Integrative Psychol. Behav. Sci.*, vol. 42, no. 3, pp. 266–290, 2008.
[39] S. Schulze, "Views on the combination of quantitative and qualitative research approaches," *Progressio*, vol. 25, no. 2, pp. 8–20, 2003.
[40] P. T. Harker and L. G. Vargas, "The theory of ratio scale estimation: Saaty's analytic hierarchy process," *Manage. Sci.*, vol. 33, no. 11, pp. 1383–1403, 1987.
[41] D. O. Norman and M. L. Kuras, "Engineering complex systems," in *Complex Engineered Systems*. Berlin, Germany: Springer, 2006, pp. 206–245.
[42] F. A. Lootsma, "Scale sensitivity in the multiplicative AHP and SMART," *J. Multi-Criteria Decis. Anal.*, vol. 2, no. 2, pp. 87–110, 1993.
[43] H. A. Donegan and F. J. Dodd, "A note on Saaty's random indexes," *Math. Comput. Model.*, vol. 15, no. 10, pp. 135–137, 1991.
[44] F. J. Dodd, H. A. Donegan, and T. B. M. McMaster, "A statistical approach to consistency in AHP," *Math. Comput. Model.*, vol. 18, no. 6, pp. 19–22, 1993.
[45] B. S. Blanchard, *System Engineering Management*. Hoboken, NJ, USA: Wiley, 2004.
[46] S. Tariq and J. Woodman, "Using mixed methods in health research," *JRSM Short Rep.*, vol. 4, no. 6, 2013, Art. no. 2042533313479197.
[47] D. Hanna and M. Dempster, *Psychology Statistics for Dummies*. Hoboken, NJ, USA: Wiley, 2012.
[48] M. Marcus and H. Minc, *Introduction to Linear Algebra*. Chelmsford, MA, USA: Courier Corp., 1988.
[49] J. Cleland-Huang, R. Settimi, X. Zou, and P. Solc, "Automated classification of non-functional requirements," *Requirements Eng.*, vol. 12, no. 2, pp. 103–120, 2007.
[50] P. Laininen and R. P. Hämäläinen, "Analyzing AHP-matrices by regression," *Eur. J. Oper. Res.*, vol. 148, no. 3, pp. 514–524, 2003.
[51] A. A. Salo and R. P. Hämäläinen, "On the measurement of preferences in the analytic hierarchy process," *J. Multi-Criteria Decis. Anal.*, vol. 6, no. 6, pp. 309–319, 1997.
[52] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychol. Rev.*, vol. 63, no. 2, pp. 81–97, 1956.
[53] M. D. Byrne, "How many times should a stochastic model be run? An approach based on confidence intervals," in *Proc. 12th Int. Conf. Cogn. Model.*, Ottawa, Canada, Jul. 2013, pp. 445–450.
[54] M. F. Triola, *Elementary Statistics* (Technology Update). London, U.K.: Pearson Education, 2010.
[55] J. Franek and A. Kresta, "Judgment scales and consistency measure in AHP," *Proc. Econ. Finance*, vol. 12, pp. 164–173, 2014.
[56] J. A. Alonso and M. T. Lamata, "Consistency in the analytic hierarchy process: A new approach," *Int. J. Uncertainty, Fuzziness, Knowl.-Based Syst.*, vol. 14, no. 4, pp. 445–459, 2006.
[57] V. R. Tummala and Y. W. Wan, "On the mean random inconsistency index of analytic hierarchy process (AHP)," *Comput. Ind. Eng.*, vol. 27, nos. 1–4, pp. 401–404, 1994.
[58] T. L. Saaty, *The Analytic Hierarchy Process*. New York, NY, USA: McGraw-Hill, 1980.
[59] D. V. Budescu, R. Zwick, and A. Rapoport, "A comparison of the eigenvalue method and the geometric mean procedure for ratio scaling," *Appl. Psychol. Meas.*, vol. 10, no. 1, pp. 69–78, 1986.
[60] W. Oberle, "Monte Carlo simulations: Number of iterations and accuracy," United States Army Res. Lab., Aberdeen, MD, USA, Tech. Rep. ARL-TN-0684, 2015.
[61] T. Kamishima and S. Akaho, "Efficient clustering for orders," in *Mining Complex Data*. Berlin, Germany: Springer, 2009, pp. 261–279.

Brian Connett (M'18) received the B.S. degree in information systems from Drexel University, Philadelphia, PA, USA, in 1999, and the M.S. degree in systems engineering in 2006, and the M.S. degree in space operations from the Naval Postgraduate School, Monterey, CA, USA, in 2012.

He is currently assigned to the United States Naval Academy, Annapolis, MD, USA, examining concepts of cyber-physical systems and the complexity of decision-making regarding its trustworthiness. As a Career Naval Officer, he has recently been a Junior Faculty Member with the United States Naval Academy, leading midshipmen in the classroom as a controls systems and cyber operations instructor. Prior to this classroom mission, his naval experiences include assignment to the U.S. Naval Air Forces, Navy Information Operations Command, Special Boat Team TWENTY, and USS Lake Erie (CG70).

Prof. Connett personal decorations include the Meritorious Service Medal, Defense Meritorious Medal, Navy and Marine Corps Commendation, and Navy and Marine Corps Achievement Medal.

Bryan M. O'Halloran received the B.Sc. degree in engineering physics and the M.Sc. and Ph.D. degrees in mechanical engineering from Oregon State University, Corvallis, OR, USA, in 2009, 2011, and 2013, respectively.

He is currently an Assistant Professor with the Systems Engineering Department, Naval Postgraduate School, Monterey, CA, USA, and the Academic Associate for the Reliability and Maintainability certificate program (curriculum 242). He was a Senior Reliability and Systems Safety Engineer with Raytheon Missile Systems, Tucson, AZ, USA, and the Lead Reliability and Safety Engineer for hypersonic missile programs. His current research interests include risk, reliability, safety, and failure modeling in the early design of complex systems.

Anthony G. Pollman received the B.S. and M.S. degrees in nuclear engineering from Purdue University, West Lafayette, IN, USA, both in 2004, the Executive MBA in diagnosis and analysis of marine corps organizational culture from the Naval Postgraduate School (NPS), Monterey, CA, USA, in 2015 and the Ph.D. degree in mechanical engineering from the University of Maryland, College Park, MD, USA, in 2011.

He is currently an Assistant Professor with the Systems Engineering Department and the Deputy Director of the Littoral Operations Center, Naval Postgraduate School, Monterey, CA, USA. He teaches courses in system dynamics, sensors, and mathematical modeling. He is the author of more than 30 papers. His research interests include process modeling and simulation, operational energy, autonomous systems, and radiation detectors.

Prof. Pollman academic awards include the Mills Medal for Service that had a conspicuous impact on the Naval Services from the NPS, Clements Award for Excellence in Education from the United States Naval Academy, Meritorious Service Award from the Marine Corps Association, and Annual Essay Award from the U.S. Naval Institute. He is a member of the American Society of Mechanical Engineers, American Nuclear Society, International Council on Systems Engineering, the Marine Corps Association, and the U.S. Naval Institute.