PUNGGAWA
GLOBAL RESEARCH:
JURNAL MULTIDISIPLIN

Punggawa Global Research: Jurnal Multidisiplin
Volume 1, Number 4, 2026
E-ISSN: 3164-0117
Open Access: punggawajournal.com/pgr

A Comparison of Different Measurement Scales Adopted in the Analytic Hierarchy Process (AHP) Based on Evaluation Test Criteria

Mohammad Azadfallah[1]
[1]Researcher, Business Studies and Development Office, Saipayadak (SAIPA After Sales Services Organization), Islamic Republic of Iran
Correspondence: m.azadfallah@yahoo.com

Article Info

Article history:
Received July 20th, 2026
Revised Aug 25th, 2026
Accepted Aug 31th, 2026

Keyword:
Analytic Hierarchy Process (AHP); Decision analysis; Multi-Criteria Decision Making (MCDM); Measurement scale; Pairwise comparison; Ratio scale.

ABSTRACT

This study reviewed and compared seven measurement scales adopted in the Analytic Hierarchy Process (AHP), including Saaty's 1–9 linear scale, power, geometric, logarithmic, root-square, inverse linear, and balanced scales. The scales were evaluated through twelve criteria covering gradation, scale range, value distribution, midpoint error, discreteness, finiteness, discrimination, ratio preservation, parameter changes, priority effects, sensitivity to pairwise-comparison changes, and consistency. Numerical examples were used to examine how alternative scales affected priorities and rankings. The results showed that no single alternative scale consistently dominated the benchmark across all evaluation criteria. Several scales performed better on specific criteria, but their advantages were not uniform. Overall, Saaty's 1–9 scale demonstrated the most balanced performance across the tests and remained the most robust general-purpose scale among the alternatives examined.

INTRODUCTION

One of the most important tasks faced by decision makers in business and government is that of selection. Selection problems are challenging, because they require the balancing of multiple, often conflicting objectives, criteria, or attributes (Olson, 1996). Multi-Criteria problems are therefore extremely important and request an appropriate treatment (Brans & Mareschal, 2005).

The concept of Multi-Criteria Decision Analysis (MCDA, also often called Multi-Criteria Decision Making; MCDM) has been used since life began on surface of the earth. It might interest you to know that even animals use instinctive reasoning or judgment to make their own decisions. If animals can make good use of this concept, what can humans, who are far more intelligent than animals, achieve? Although this concept has existed since the origins of life, it was not documented or recorded to assist in making decisions when the need arose (Uzun, Ozsahin, Agbor & Ozsahin, 2021). Since the 1970's, MCDM research has developed rapidly and has become a hot research topic because many complex practical decision problems involve multiple and conflicting criteria as well as multiple objectives (Saaty & Daji, 2015).

According to Qu, Wan, Yang and Lee (2018), MCDM methods can be defined as structure frameworks that deal with the process of making decisions in the presence of multiple objectives. They are used to find the best opinion from all of the feasible alternatives in the presence of multiple, usually conflicting, decision criteria. There is some empirical evidence that multi-criteria techniques are especially helpful in enlarging insight into the choice-possibilities and to make the various characteristics of those choice-possibilities better debatable (Voogd, 1982). Therefore, over the past few decades a number of MCDM methods have been developed to deal with the measurement of tangible/intangible conflicting criteria and with the measurement of the alternatives of a decision with respect to these criteria (Saaty & Daji, 2015). In this paper, Analytic Hierarchy Process (AHP) that is one of these methods is discussed.

One of the well-known methods for solving discrete problems is the AHP, which is applied to prioritize alternatives and is also used with other techniques in many different contexts (Sequra, Maroto, Belton, Ginestar & Marques, 2019).

AHP is a MCDM method, which was originally developed by Saaty (Saaty, 2000). AHP is a useful decision analysis tool and has been widely used in many areas (Ji & Jiang, 2003), with impressive results (Ishizaka, Balkenborg & Kaplan, 2011). According to Ishizaka and Labib (2009), AHP is a MCDM method that helps the decision maker facing a complex problem with multiple conflicting and subjective criteria. Forman and Selly (2001) believe that, AHP is composed of several previously existing but unassociated concepts and techniques such as hierarchical structuring of complexity, pair wise comparisons, redundant judgments, an eigenvector method for deriving weights, and consistency considerations. Although each of these concepts and techniques were useful in and of themselves, Saaty's synergistic combination of the concepts and techniques (along with some new developments) produced a process whose power is indeed far more than the sum of its parts.

The three terms in the name of the method give us information about the characteristics of the methodology (Grunig & Kuhn, 2005):

- "Analytical" means that the decision goal is broken down into criteria. The option can be compared with respect to both qualitative and quantitative criteria. The criteria can be weighted. The overall assessment of the options is found with linear algebra.
- "Hierarchical" refers to the presentation of criteria, environmental conditions and options. In AHP, these are always arranged on different hierarchical levels.
- "Process" indicates that the solution of a complex decision problem is organized as a systematic sequence of component steps.

According to Saaty (2000), despite the rigorous mathematical foundations of the AHP, it still meets the criteria of simplicity to use by both practitioner and academics steeped in their specialization without complete prior knowledge of the development of the theory. In sum, the AHP has a systematic procedure for better judgments (Saaty, 2001_a). Further, AHP utilizes qualitative description to define a problem and to represent the interactions of its parts. It also makes use of quantitative judgments to assess the strengths of these interactions. The decision maker first identifies his or her main purpose in solving a problem. Criteria are chosen and weighted according to the priority of their importance to the decision maker. The different alternatives are then evaluated in terms of these criteria, and a best one or best mix is chosen. The alternatives are then potential solutions to the problem (Saaty & Vargas, 1991).

In sum, AHP is based on three basic principles: decomposition, comparative judgments, and hierarchic composition or synthesis of priorities. The decomposition principle is applied to structure a complex problem into a hierarchy of clusters, sub-clusters, sub-sub-clusters and so on. The principle of comparative judgments is applied to construct pair wise comparisons of all combinations of elements in a cluster with respect to the prevent of the cluster. These pair wise comparisons are used to derive "local" priorities of the elements in a cluster with respect to their parent. The principle of hierarchic composition or synthesis is applied to multiply the local priorities of elements in a cluster by the "global" priority of the parent element, producing global priorities throughout the hierarchy and then adding the global priorities for the lowest level elements (the alternatives) (Forman & Selly, 2001).

On the other side, measurement theory is a branch of applied mathematics that is useful in measurement and data analysis (Sarie, 1997). Stevens (1946) believe that, measurement is defined as the assignment of numerals to objects or events according to rules. The only rule not allowed would be random assignment, for randomness amounts is effect to a non-rule (Luce, 1997).

In MCDA every performance is associated with a suitable scale to evaluate or rank alternatives to achieve goal. For instance, A is two times better than B. here, linguistic term 'two times better than' needs a scale to measure degree of preference (Mukherjee, 2017). In addition, in the AHP, a decision maker first gives linguistic pair wise comparisons, then obtains numerical pair wise comparisons by selecting certain numerical scale to quantify them, and finally derives a priority vector from the numerical pair wise comparisons (Dong, Xu, Li & Dai, 2008).

A scale is a triplet, encompasses a set of numbers, a set of objects, and mapping of objects to the number. There are different types of scale such as:

1. Nominal scale: A number is assigned to each object. For example, queue in front of railway ticket reservation counter.
2. Ordinal scale: Numbers are assigned to each object to represent their order, increasing or decreasing.
3. Interval scale: For example, $Y=B-30$, where Y is a dependent variable and B is an independent variable.
4. Ratio scale: For example, $K=BL$, where $B>0$, and is a proportional constant.
5. Absolute scale: Number is used directly for pair wise comparison. It is commonly used in AHP (Mukherjee, 2017).

According to Ishizaka and Labib (2009), one of AHP's strengths is the possibility to evaluate quantitative as well as qualitative criteria and alternatives on the same preference scale of nine levels (i.e., integers from one to nine). These can be numerical, verbal or graphical. Theoretically, there is no reason to be restricted to these numbers. Therefore, other scales have been proposed (Ishizaka, Balkenborg & Kaplan, 2011). On the other side, Harker and Vargas (1987) and Ishizaka and Labib (2009) believe that, the choice of a correct scale in the AHP is an open research issue, and a very heated debate, respectively. Therefore, in this paper we compares the performance of seven measurement scales adopted in AHP (Saaty's 1-9, power, and so on), on twelve identified criteria (i.e., number of gradation, range of scale value, and so on), to select the appropriate measurement scale(s).

Accordingly, this study compared seven measurement scales used in AHP against twelve evaluation criteria to determine whether any alternative scale performs at least as well as Saaty's 1–9 scale across the full set of tests. The contribution lies in evaluating scale behavior across several mathematical and decision-analytic properties rather than judging suitability from a single performance characteristic.

In the literature, there are several examples where different measurement scale is used and compared in the choice of the final solution. Here, we will mention some of them. For instance, Harker and Vargas (1987) addressed theoretical issues related to whether or not the AHP is a valid and acceptable method of eliciting and analyzing subjective judgments (particularly, the scale used to measure the intensity preference), and present four new scales developed based on the Saaty's 1-9 scale. Lootsma (1993) explained context-related scaling of human judgment in some of MCDM (i.e., the multiplicative AHP, SMART and ELECTRE) methods. Triantaphyllou, Lootsma, Pardalos and Mann (1994) used two evaluative criteria (1. The real continuous pair wise [*or RCP*] matrix and 2. The closest discrete pair wise [*or CDP*] matrix) to examine a total of *78* scales which can be derived from two widely used scales (i.e., the original Saaty's *1-9* scale and the exponential scale). Poyhonen, Hamalainen and Salo (1997) performed a comparative study in which subjects were requested to quantify verbal ratio statements by adjusting the heights of visually displayed bars. Salo and Hamalainen (1997) proposed new balanced scales to improve the sensitivity of the AHP ratio scales. Leskinen (2001) discussed the properties of alternative measurement scales and the probabilistic modeling of interval judgments in the framework of the regression analysis of pair wise comparisons data. Sato (2001) evaluated Saaty's linear scale and the power scale using two types of data (i.e., 1. Non-biased random sample and 2. Biased actual samples), from the point of view of appropriateness for representing each decision maker's feeling. Ji and Jiang (2003) first reviewed and compared different scales from different aspects, then discussed the transitivity of AHP scales and derive a scale based on the transitivity. Barzilai (2006) summarized the conditions for the construction of measurement scales that enable the application of the operations of algebra, calculus, and statistics to scale values. Dong, Xu, Li and Dai (2008) presented two performance measure algorithms (further the essence of these two performance measure algorithms is to calculate the difference between input and output of AHP) to evaluate the numerical scales and the prioritization methods. Cox (2009) illustrated that the graphs are a great aid in interpreting multidimensional data (particularly, in AHP and ANP [Analytic Network Process] context). Monat (2009) proposed the use of global scales (the remapped scale based on decision maker's experience, aspirations, or imagination) instead of local scale (the natural scale). Ishizaka, Balkenborg and Kaplan (2011) demonstrated the influence of aggregation and measurement scale on ranking a compromise alternative in AHP. Azadfallah (2017_a) analyzed the impacts of various aggregation rules, normalization methods and measurement scales on ranking alternatives in AHP. Azadfallah (2017_b) first reviewed different measurement scales adopted in AHP, then with reduction of different measurement scale ranges (to left position, middle position, etc.) the effects of different measurement scale on priorities and discrimination level of alternatives are investigated. Azadfallah (2017_c) presented a comparative analysis of different measurement scales adopted in AHP, by testing them versus a problem with known composite answers. Then experimentally, the impact of the different measurement scale elements alteration from three aspects (1. The limited scale upper bound, 2. Changing the scale parameters and 3. Changing the system numbers) on priorities are investigated. Starczewski (2017) analyzed the dependence of the quality of the final decisions about priority vectors on the assumed scale that expresses the pair wise comparisons judgments. Goepel (2019) founded a set of parameters allowing a further classification of AHP scales (i.e., the maximum entry value is kept at nine, reduced to lower values than nine and extended to values higher than nine) and impact of different scales on the resulting priorities. Lipovetsky (2021) considered the nonlinear scaling (i.e., transforming the ratio scale into the additive or logarithmic scales) in the AHP approach to different problems. Romanchak (2023) proposed an adjustment to the mathematical model of the AHP using the Stevens direct measurement scale/model. Azadfallah (2025) suggested four new measurement scales (three Fibonacci-based and one Golden Ratio-based) to obtain a more suitable measurement scale in the MCDM (particularly, AHP) context. Lukinskiy, Lukinskiy, and Bazhina (2025) analyzed of nearly 500 expert judgment matrices reveals that the upper bounds of Saaty's 1–9 scale are rarely used. then supporting the adoption of a simplified 1–7 scale to reduce experts' cognitive load and enhance judgment reliability. Royo and Chiclana (2026) applied the Intentional Bounded Rationality Methodology (IBRM) to study the performance of the most used AHP ratio scales in the literature.

In sum, according to Goepel (2019), the discussion about the right scale for the AHP is ongoing for many years. Therefore, our paper aims to shed some light on the select of the appropriate measurement scale(s) in AHP's environment.

RESEARCH METHODS

Research Design

This study used a comparative numerical research design. The analytical units were seven measurement scales used in the Analytic Hierarchy Process: Saaty's 1–9 linear scale, power, geometric, logarithmic, root-square, inverse-linear, and balanced scales. Because the study evaluated mathematical scale behavior rather than human participants, no population or respondent sample was employed. Saaty's 1–9 scale was treated as the benchmark because it remains the most widely used AHP scale and provides the reference against which the alternatives were compared.

The analysis examined each scale through numerical examples, pairwise-comparison matrices, priority vectors, ranking outcomes, and consistency measures. Calculations followed the standard AHP procedures described below. The comparison focused on whether an alternative scale could match or improve upon the benchmark without losing performance on other evaluative dimensions.

Performance Evaluation Criteria

Twelve criteria were used to evaluate the seven scales: number of gradations, range of scale values, uniformity of value distribution, midpoint error, discreteness, finiteness, discrimination among alternatives, ratio preservation, effects of scale-parameter changes, effects of scale type on priorities, sensitivity to changes in pairwise-comparison elements, and consistency. For each criterion, the numerical result for an alternative scale was compared with the corresponding result obtained from Saaty's 1–9 scale.

The resulting values were summarized in tables and interpreted comparatively. Where a criterion produced a direct error, inconsistency, or sensitivity measure, lower or higher values were assessed according to the logic of that criterion. The final assessment considered the pattern across all twelve tests rather than selecting a scale from a single favorable result.

Analytic Hierarchy Process (AHP)

AHP can easily be employed with the following steps (Mukherjee, 2017):

Step1: Define goal or objective.

Step2: Define criteria and sub-criteria to accomplish goal.

Step3: Use Saaty's nine-point preference scale to form pair wise comparison matrix. Let *B* is a *n.n* pair wise comparison matrix.

$$B = \begin{pmatrix} b_{11} & b_{21} \cdots & b_{1n} \\ b_{21} & b_{22} \cdots & b_{2n} \\ \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} \cdots & b_{nn} \end{pmatrix} where b_{ij} = 1, i = j \tag{1}$$

step4: Normalize the matrix with geometric mean as follows:

$$u_i = \left[\sum_{j=1}^{n} b_{ij}\right]^{1/n} / \sum_{i=1}^{n} \left[\sum^{n}{}_{j=1} b_{ij}\right]^{1/n}, \ i = j = 1,\ 2,\ 3,\ \ldots, n \tag{2}$$

step5: If *R* denotes *n-dimensional* column vector describing the sum of

$$R = \left[R_i\right]_{n.1} = Bu^T,\ i = 1,\ 2,\ \ldots,\ n,$$

Where

$$Bu^T = \begin{pmatrix} 1 & b_{21} \cdots & b_{1n} \\ b_{21} & 1 \cdots & b_{2n} \\ \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} \cdots & 1 \end{pmatrix} \begin{bmatrix} u_1 & u_2 & \ldots & u_n \end{bmatrix} = \begin{bmatrix} R1 \\ R2 \\ \vdots \\ Rn \end{bmatrix} \tag{3}$$

Step6. Saaty proposed the use of maximum eigenvalue λ_{max} to check consistency of judgment. The maximum eigenvalue λ_{max} can be determined as follows:

$$\lambda_{max} = \sum_{i=1}^{n} R_{ui} / n, \mathbf{i} = 1,\ 2,\ 3,\ \ldots,\ n \tag{4}$$

Step7. Calculate consistency index *(CI)* as follows:

$$CI = \lambda_{max} - n \ / \ n - 1 \tag{5}$$

Step8. Calculate consistency ratio *(CR)* to check the validity of decision making process.

$$CR = CI \ / \ RI \tag{6}$$

Where *RI* denote average random index. Saaty proposed that *CR* value should be less than or equal to 0.1. Some researchers also mentioned that random index *(RI)* can be represented as follows; *RI= {1.98*(m-2) / m}*, where *m* is the size of the matrix.

Measurement scale

According to Ji and Jiang (2003), AHP is a useful decision analysis tool and has been widely used in many areas. Pair wise comparisons are fundamental in the use of the AHP. The values defined on a scale are used to represent the relative importance between two compared objects in terms of ratio. The scale to be used has a significant effect on the outcome's consistency and accuracy.

Saaty in 2001 (Saaty, 2001_b) demonstrate how the integers *1 to 9* used in the fundamental scale of the AHP to represent pair wise comparison judgments can be derived from stimulus-response theory. This is the problem we wish to address here.

According to Saaty (2000, 2008), historically, in 1846 Weber formulated his law regarding a stimulus *(S)* of measurable magnitudes (Eq. 7).

$$S + \Delta s, \tag{7}$$

Where Δ_S is called the just noticeable difference.

In 1860 Fechner considered a sequence of just noticeable increasing stimulus (the first one by S_0) (Eqs. 8-10).

$$S_1 = S_0 + \Delta_{S0} = S_0 + \left(\left(\Delta_{S0} / S_0\right)*S_0\right) = S_0\left(1+r\right) \tag{8}$$

Similarly

$$S_2 = S_1 + \Delta_{S1} = S_1\left(1+r\right) = S_0\left(1+r\right)^2 = S_0\alpha^2 \tag{9}$$

In general

$$S_n = S_{n-1}\alpha = S_0\alpha^n \, (n = 1, \; 2, \ldots) \tag{10}$$

Thus stimuli of noticeable differences follow sequentially in a geometric progression. Fechner noted that the corresponding sensations should follow each other in an arithmetic sequence at the discrete points at which just noticeable differences occur. But the latter are obtained when we solve for *n* (Eq. 11), have

$$n = \left(Log \; S_n - Log \; S_0\right) \; / \; Log \; \alpha \tag{11}$$

And sensation is a linear function of the logarithm of the stimulus. Thus if *M* denotes the sensation and *S* the stimulus, the psychophysical law of Weber-Fechner is given by (Eq. 12):

$$M = a \; Log \; S + b, \; a \neq 0 \tag{12}$$

We take the ratios M_i / M_1, $i=1, \ldots, n$ of these responses in which the first is the smallest and serves as the unit of comparison (and $b=0$), thus obtaining the integer values *1, 2, …, n* of the fundamental scale of the AHP.

Why the upper limit of the scale is *9* is due to the requirement of homogeneity as a condition for the stability of the eigenvector of priorities that also depends on comparing a few elements at a time (Saaty, 2001_b). In addition, according to Triantaphyllou, Shu, Sanchez and Ray (1998), psychological experiments have also shown that individuals cannot simultaneously compare more than seven objects (plus or minus two).

As discussed earlier, to derive priorities, the verbal comparisons must be converted into numerical ones. In Saaty's AHP the verbal statements are converted into integers from one to nine. Theoretically there is no reason to be restricted to these numbers and verbal gradation (Ishizaka & Labib, 2009). Therefore, other measurement scales have been proposed (Ishizaka, Balkenborg & Kaplan, 2011), as follows.

1. Saaty's 1-9 (or Linear) scale

$$C = a \, . \, x, \; a > 0; \; x = 1, \; 2, \; \ldots, \; 9 \tag{13}$$

2. Power scale

$$C = x^a, \; a > 1; \; x = 1, \; 2, \; \ldots, \; 9 \tag{14}$$

3. Geometric scale

$$C = a^{x-1}, \; a > 1; \; x = 1, \; 2, \; \ldots, \; 9 \tag{15}$$

4. Logarithmic scale

$$C = Log_a \left(x + 1\right), \; a > 1; \; x = 1, \; 2, \; \ldots, \; 9 \tag{16}$$

5. Root Square

$$C = \sqrt[a]{x}, \; a > 1; \; x = 1, \; 2, \; \ldots, \; 9 \tag{17}$$

6. Inverse linear

$$C = 9 \; / \; \left(10 - x\right); \; x = 1, \; 2, \; \ldots, \; 9 \tag{18}$$

7. Balanced

$$C = \frac{w}{(1-w)}; \; w = 0.5, \; 0.55, \; 0.60, \; \ldots, \; 0.9 \tag{19}$$

Table 1 and Figure 1 compare the seven measurement scales.

Table 1 Different Measurement Scales For Comparing Two Alternatives*, **

Scale	**Value**															
Saaty's	0.111	0.125	0.143	0.167	0.200	0.250	0.333	0.500	1	2	3	4	5	6	7	8

Scale	Value																
(or Linear)																	
Power	0.012	0.016	0.020	0.028	0.040	0.063	0.111	0.250	1	4	9	16	25	36	49	6
Geometric	0.004	0.008	0.016	0.031	0.063	0.125	0.250	0.500	1	2	4	8	16	32	64	1
Logarithmic	0.301	0.315	0.333	0.356	0.388	0.431	0.500	0.633	1	1.58	2	2.32	2.58	2.81	3	3.
Root Square	0.333	0.353	0.377	0.408	0.422	0.500	0.578	0.709	1	1.41	1.73	2	2.37	2.45	2.65	2.
Inverse Linear	0.111	0.222	0.333	0.444	0.556	0.667	0.775	0.885	1	1.13	1.29	1.5	1.8	2.25	3	4
Balance	0.111	0.176	0.250	0.333	0.429	0.538	0.667	0.820	1	1.22	1.5	1.86	2.33	3	4	5.

*. Calculated by presented formula (Eqs. 13-19),

**. Notice; a=1 for the Saaty's scale and a=2 for all other scales.

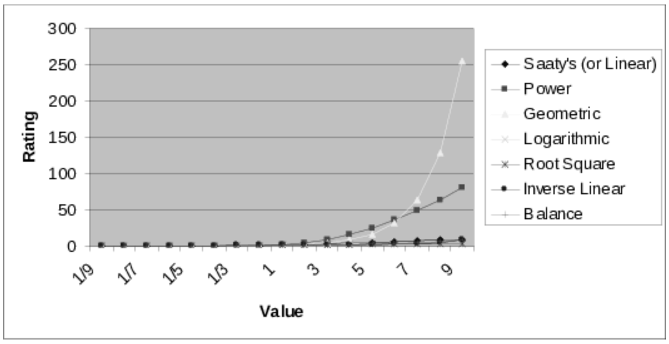

Figure 1 Comparison Of Results

RESULTS AND DISCUSSION

The number of gradation

According to Triantaphyllou, Lootsma, Pardalos and Mann (1994), psychological experiments (from Miller in 1956) have also shown that individuals cannot simultaneously compare more than seven objects (plus or minus two). This is the main reasoning used by Saaty to establish *9* as the upper limit of his scale, *1* as the lower limit and a unit difference between successive scale values.

As can be seen from table *1*, the all of the measurement scales adopt nine grade for right side (i.e., *1-9* based on Saaty's scale) and seventeen (i.e., *1/9-9* based on Saaty's scale) for both sides (right side include: intensity of importance and left side include: reciprocals of intensity of importance). Therefore, the result is the same for all scales as the benchmark (i.e., Saaty's *1-9* scale or linear scale).

The range of scale values

This is the simplest way of describing the spread of data (such as subtract the minimum value from the maximum value, Eq. 20) (Denscombe, 2017). For instance, for Saaty's *1-9* scales have:

$$Range\ (R) = Maximum\ value - Minimum\ value \quad (20)$$

$$= 9 - 0.111 = 8.889$$

In addition, implicitly, the interval of scale values, by (Eq. 21):

$$Minimum\ value \leq a_{ij} \leq Maximum\ value$$

$$0.111 \leq a_{ij} \leq 9$$

A comparison of the test results is given in Table 2 and Figure 2.

Table 2 The Range And Interval Of Values For Different Measurement Scale

Type of scale	Minimum of value	Maximum of value	Range*	The interval of scale values**
Saaty's (or Linear)	0.111	9	8.889	$0.111 \leq a_{ij} \leq 9$
Power	0.012	81	80.988	$0.012 \leq a_{ij} \leq 81$
Geometric	0.004	256	255.996	$0.004 \leq a_{ij} \leq 256$
Logarithmic	0.301	3.32	3.019	$0.301 \leq a_{ij} \leq 3.32$
Root Square	0.333	3	2.667	$0.333 \leq a_{ij} \leq 3$
Inverse Linear	0.111	9	8.889	$0.111 \leq a_{ij} \leq 9$
Balance	0.111	9	8.889	$0.111 \leq a_{ij} \leq 9$

**. By using Eq. (20),*

***. By using Eq. (21).*

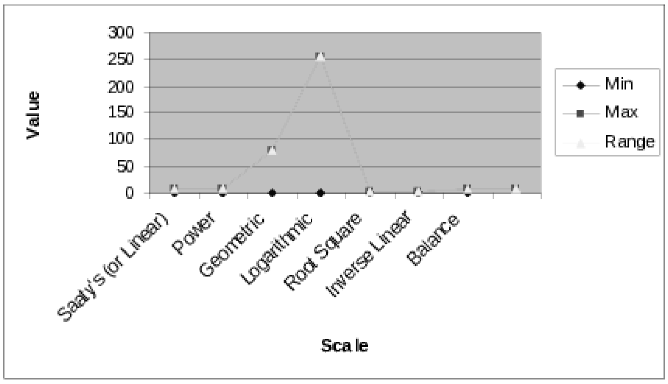

Figure 2 Comparison Of Results

As can be seen from table 2 and Figure 2, the inverse linear and balanced measurement scales exhibit the closeness range and interval of scale values as compared with standard (Saaty's 1-9 scale).

The uniformity of value distribution

In this test uniformity in increasing and decreasing, the value of the scale in different scales has been investigated. For instance, for Saaty's *1-9* scales (table 3) have:

Table 3 The Uniformity Of Value Distribution On Saaty's Scale

Value / Scale	1/9	1/8	1/7	1/6	1/5	1/4	1/3	1/2	1	2	3	4	5	6	7	8	9
Saaty	0.014*	0.018	0.024	0.033	0.050	0.083	0.167	0.500	1	1	1	1	1	1	1	1	1

**. =0.125-0.111 (based on table 1).*

A comparison of the test results is given in Table 4.

Table 4 The Uniformity Of Value Distribution On Different Scale

Scale	Distribution Value														
Saaty's (or Linear)	0.014	0.018	0.024	0.033	0.050	0.083	0.167	0.500	1	1	1	1	1	1	1

Scale	Distribution Value														
Power	0.003	0.005	0.007	0.012	0.023	0.049	0.139	0.750	3	5	7	9	11	13	15
Geometric	0.004	0.008	0.016	0.031	0.063	0.125	0.250	0.500	1	2	4	8	16	32	64
Logarithmic	0.014	0.018	0.023	0.032	0.043	0.069	0.133	0.367	0.580	0.420	0.320	0.260	0.230	0.190	0.170
Root Square	0.020	0.024	0.031	0.014	0.078	0.078	0.131	0.291	0.410	0.320	0.270	0.370	0.080	0.200	0.180
Inverse Linear	0.111	0.111	0.111	0.111	0.111	0.109	0.110	0.115	0.130	0.160	0.210	0.300	0.450	0.750	1.500
Balance	0.065	0.074	0.083	0.096	0.108	0.129	0.153	0.180	0.220	0.280	0.360	0.470	0.670	1.000	1.670

As can be seen from Table 4, the inverse linear, geometric and power scales, holds uniformly of values for left side, both sides and right side, respectively.

The middle –scales error

This test (borrowing the idea from Shinohara, Miyake and Ohsawa, 2001) deal with the error caused by determining the middle of the scale as the value by the decision maker and the proposed value of the scale. For instance, for Saaty's *1-9* scale:

Error = | the midpoint of scale – Closest value provided by scale|

For right side of scale:

$E_{Right} = |4.5 - 5| = 0.5$

For left side of scale:

$E_{Left} = |0.250 – 0.200| = 0.050$

A comparison of the test results is given in Table 5.

Table 5 The Middle-scales Error For Different Scale

Scale type	Error measure	
	left side	right side
Saaty's (or Linear)	0.050	0.500
Power	0.085	15.5
Geometric	0.187	112
Logarithmic	0.071	0.920
Root Square	0.067	0.870
Inverse Linear	0.111	2.7
Balance	0.019	2.17

As can be seen from Table 5, the geometric and balance for the left side and the geometric and linear (or Saaty's scale) for the right side of the scale have the highest and lowest errors, respectively.

The discreteness of scale

According to Ji and Jiang (2003), in a judgment matrix *A*, a perfectly consistent judgment *(i.e., satisfying* $a_{ij}=a_{ik}.\ a_{kj}$*)* cannot be achieved since $a_{ik}.\ a_{kj}$ yields the values that are not those defined on the scale, although they do not break the bound of the scale. As a result, the values have to be rounded up to the adjustment values defined on the scale.

In this section, the values that are not those defined on the different measurement scales considered (table 6 and table 7). For instance, for Saaty's scale:

Table 6 The Discreteness Of Saaty's Scale

Saaty's scale	1	1/2	1/3	1/4	1/5	1/6	1/7	1/8	1/9
1	1	0.50	0.33	0.25	0.20	0.17	0.14	0.13	0.11
2	2	1	0.67	0.50	0.40	0.33	0.29	0.25	0.22
3	3	1.50	1	0.75	0.60	0.50	0.43	0.38	0.33
4	4	2	1.33	1	0.80	0.67	0.57	0.50	0.44
5	5	2.50	1.67	1.25	1	0.83	0.71	0.63	0.56
6	6	3	2	1.50	1.20	1	0.86	0.75	0.67
7	7	3.50	2.33	1.75	1.40	1.17	1	0.88	0.78
8	8	4	2.67	2	1.60	1.33	1.14	1	0.89
9	9	4.50	3	2.25	1.80	1.50	1.29	1.13	1

As can be seen from table 6, *44/81* is not the defined values by the Saaty's measurement scale.

A comparison of the test results is given in table 7.

Table 7 The Discreteness Of Different Measurement Scale

Scale type	Inconsistency caused by the discreteness of scale
Saaty's (or Linear)	44/81
Power	44/81
Geometric	0
Logarithmic	54/81
Root Square	49/81
Inverse Linear	47/81
Balance	53/81

As can be seen from table 7, only a geometric measurement scale has no such inconsistencies.

The finiteness of scale

In continuing, kike above test (test number 5), Ji and Jiang (2003) believe that, a perfectly consistent judgment *(i.e., satisfying* $a_{ij}=a_{ik}.\ a_{kj}$*)* may not be achieved since $a_{ik}.\ a_{kj}$ probably breaks the upper/ lower bound of the scale, and a_{ij}, a_{ik} *and* a_{kj} must take the values within the scale. As a result, the inconsistency is caused as the cases *'7.7=9' and '8.8=9'*.

In this section, the inconsistency caused by the scale finiteness is considered (table 8 and table 9). For instance, for Saaty's scale:

Table 8 The Finiteness Saaty's Scale

Saaty's scale	1	2	3	4	5	6	7	8	9
1	1	2	3	4	5	6	7	8	9
2	2	4	6	8	10	12	14	16	18
3	3	6	9	12	15	18	21	24	27

Saaty's scale	1	2	3	4	5	6	7	8	9
4	4	8	12	16	20	24	28	32	36
5	5	10	15	20	25	30	35	40	45
6	6	12	18	24	30	36	42	48	54
7	7	14	21	28	35	42	49	56	63
8	8	16	24	32	40	48	56	64	72
9	9	18	27	36	45	54	63	72	81

As seen from the above table, 58/81 is not values within the Saaty's measurement scales.

A comparison of the test results is given in table 8.

Table 9 The Finiteness Of Different Measurement Scale

Scale type	Inconsistency caused by the finiteness of scale
Saaty's (or Linear)	58/81
Power	58/81
Geometric	36/81
Logarithmic	64/81
Root Square	64/81
Inverse Linear	60/81
Balance	63/81

As can be seen from table 9, all of the existing measurement scales are subject to inconsistency.

Before continuing, it is worth noting here that, according to Wedley (2001), "*at the 5th international symposium of the AHP in Kobe Japan, Thomas Saaty suggested to the author that the controversy regarding correct synthesis modes for the AHP should be tested with problems with known true values ...*". In addition, Saaty (2000) said; "*in the AHP one needs to be careful with criteria measured on the same absolute scale. Criteria measured in dollars are a common example of this. The priority of each criterion must be equal to the sum of the measurements of it's alternative's divided by the sum of the measurements of the alternative's with respect to all such criteria. Only then can one normalize the measurements of the alternative's, weight them by these priorities and add to obtain the relative weights of the alternative's with respect to all these criteria*". Since, the use of problems with known answers is the aim of this paper.

To illustrate these basic ideas, with assumes that all the criteria are expressed in the same unit, a simple 4.3 example is presented (table 10).

Notice; according to axiom 3 of AHP, the criteria are assumed to be independent of the alternatives (Wedley, 2001). Here, we violate of this properties. Clearly, this is an assumption that is made here in order to study different measurement scales. Hence, this could cause some bias in the result.

Table 10 Decision Matrix (the Problem With Known Answer)

Criteria / Alternative	C_1^*	C_2^*	C_3^*	total	True weights
A_1	1	9	5	15	0.227

Criteria / Alternative	C_1^*	C_2^*	C_3^*	total	True weights
A_2	3	7	7	17	0.258
A_3	7	3	5	15	0.227
A_4	9	3	7	19	0.288
total	20	22	24	66	1

*. Benefit-type criteria.

The discriminating level of alternatives

According to Sato (2001), one of the principal purposes of the AHP is to quantify the degree of importance for each alternative and to discriminate an important alternative from others. Since, in this section, the difference in the intensity of preference between the two top alternatives has been considered. For instance, for Saaty's scale (table 11):

Table 11 The Discriminating Level Of Alternatives For Saaty's Scale

Criteria & weights / Alternative	C_1^* 20/66	C_2^* 22/66	C_3^* 24/66	True weights
A_1	0.050	0.409	0.208	0.227
A_2	0.150	0.318	0.292	0.258
A_3	0.350	0.136	0.208	0.227
A_4	0.450	0.136	0.292	0.288

As can be seen from table 11, the AHP result with Saaty's *1-9* scale (priorities), are the same as compared with standard (the problem with known answer). Nevertheless, we have:

The discriminating level between the two top alternatives = $|A_i - A_j| = |0.288 - 0.258| = 0.030$.

A comparison of the test results is given in table 12.

Table 12 The Discriminating Level Of Alternatives For Different Measurement Scale

Scale type	The discriminating level
Saaty's (or Linear)	0.030
Power	0.074
Geometric	0.063
Logarithmic	0.010
Root Square	0.013
Inverse Linear	0.038
Balance	0.049

As can be seen from table 12, only the Saaty's *1-9* scale result *(0.030)* is the same as compared with standard. However, the power *(0.074)*, geometric *(0.063)*, inverse linear *(0.038)*, and balance *(0.049)* measurement scales could apply to improve the discriminating of level of importance alternatives.

The ratio preservation test

Wedley (2001) said; ratio preservation rather than rank preservation is chosen as the measure of effectiveness, because AHP produces ratio answers. In this section, the ratio to best is considered. The ratio to best is dividing the score of each alternative by the score of the best alternatives (Azadfallah, 2017_a). For instance, for Saaty's scale (table 13):

Table 13 The Ratio To Best For Saaty's Measurement Scale

Method / Alternative	Standard		AHP with Saaty's scale	
	True weights	Ratio to best	Priorities	Ratio to best
A_1	0.227	0.788	0.227	0.788
A_2	0.258	0.896	0.258	0.896
A_3	0.227	0.788	0.227	0.788
A_4	0.288	1	0.288	1

As can be seen from table 13, the AHP with Saaty's scale ratios present the correct ratios, as compared with standard ratios.

A comparison of the test results is given in table 14.

Table 14 The Ratio To Best For Different Measurement Scale

Method / Alternative	standard		Saaty		Power		Geometric		Logarithmic		Root Square		Inverse Linear		Balance	
	TW*	RB**	P***	RB	P	RB	P	RB	P	RB	P	RB	P	RB	P	RB
A_1	0.227	0.788	0.227	0.788	0.245	0.768	0.336	0.842	0.224	0.830	0.231	0.862	0.307	0.890	0.279	0.851
A_2	0.258	0.896	0.258	0.896	0.245	0.768	0.162	0.406	0.260	0.963	0.255	0.951	0.189	0.548	0.215	0.655
A_3	0.227	0.788	0.227	0.788	0.190	0.596	0.103	0.258	0.246	0.911	0.245	0.914	0.158	0.458	0.177	0.540
A_4	0.288	1	0.288	1	0.319	1	0.399	1	0.270	1	0.268	1	0.345	1	0.328	1

*. True weights, **. Ratio to best, ***. Priorities.

Changing the scale parameters

In this section, according to Azadfallah (2017_c), evaluating of the effect of changing the different measurement scales parameter on priorities is considered. Further, the change of the scale parameters from *a=2* (except for Saaty's scale; from *a=1*), to *a=3* for all scales is investigated. For instance, for Saaty's scale (table 15-17):

Table 15 The Change Of Saaty's Scale Parameter Result

Parameter *a*	Value								
1	1	2	3	4	5	6	7	8	9
3	3	6	9	12	15	18	21	24	27

Table 16 New Decision Matrix*

Criteria / Alternative	C_1^*	C_2^*	C_3^*
A_1	3	27	15

Criteria / Alternative	C_1^*	C_2*	C_3*
A_2	9	21	21
A_3	21	9	15
A_4	27	9	21

*. Based on table 10 and table 15.

Table 17 AHP Result For Modified Saaty's Scale (a=3)

Criteria & weights / Alternative	C_1^* 60/198	C_2* 66/198	C_3* 72/198	True weights
A_1	0.050	0.409	0.208	0.227
A_2	0.150	0.318	0.292	0.258
A_3	0.350	0.136	0.208	0.227
A_4	0.450	0.136	0.292	0.288

As can be seen from table 17, the obtained values (or priorities) for parameter *a=3*, are the same as compared with conventional parameters (i.e., *a=1*). Therefore, the change of the Saaty's scale parameters, does not impact on results.

A comparison of the test results is given in table 18.

Table 18 AHP Result For Different Modified Measurement Scales

Scale type	Criteria weights a=1	Criteria weights a=3	Priority a=1	Priority a=3
Saaty	0.303 0.333 0.364	0.303 0.333 0.364	$A_4>A_2>A_1\approx A_3$ 0.288 0.258 0.227 0.227	$A_4>A_2>A_1\approx A_3$ 0.288 0.258 0.227 0.227
Power	0.321 0.339 0.339	0.348 0.356 0.296	$A_4>A_2\approx A_1>A_3$ 0.319 0.245 0.245 0.190	$A_4>A_1>A_2>A_3$ 0.348 0.270 0.225 0.157
Geometric	0.400 0.403 0.197	0.450 0.450 0.100	$A_4>A_1>A_2>A_3$ 0.399 0.336 0.162 0.103	$A_4>A_1>A_2>A_3$ 0.450 0.409 0.090 0.050
Logarithmic	0.303 0.335 0.362	0.303 0.335 0.362	$A_4>A_2>A_3>A_1$ 0.270 0.260 0.246 0.224	$A_4>A_2>A_3>A_1$ 0.270 0.260 0.246 0.224
Root Square	0.304 0.331 0.365	0.313 .335 0.352	$A_4>A_2>A_3>A_1$ 0.268 0.255 0.245 0.231	$A_4>A_2>A_3>A_1$ 0.264 0.256 0.246 0.233
Inverse Linear*	-	-	-	-

Scale type	Criteria weights		Priority	
	a=1	a=3	a=1	a=3
Balance*	-	-	-	-

*. Will not be examined, because different parameters (instead *a* parameter) were used.

The effect of scale type on priorities

According to Ishizaka, Balkenborg and Kaplan (2011); and Azadfallah (2017_a), the type of measurement scale applied could determine the final results (or priorities). Therefore, in this section, we are focusing on the different measurement scale and their effects on priorities in AHP. Consider the following pair wise comparison matrices (For instance, for Saaty's scale) (table 19):

Table 19 The Pair Wise Comparison Matrix

1	1	3	9
	1	3	7
		1	5
			1

For determining priority weights, Saaty and Vargas (1991) proposed to use the following formula (Eq. 22):

$$Lim_{k\to\infty} A^{k}{}_{\mathbf{e}} / \mathbf{e}^{t} A^{k}{}_{\mathbf{e}} = \mathbf{C}_{w} \qquad (22)$$

Where $\boldsymbol{e}$ is the column vector unity, $\boldsymbol{e}^t$ is its transpose, and $\boldsymbol{C}$ is a positive constant. The result is the following:

W^1= (0.410 0.352 0.195 0.043),

W^2= (0.411 0.390 0.157 0.042),

W^3= (0.410 0.389 0.158 0.043),

W^4= (0.410 0.389 0.158 0.043).

As can be seen, the process has convergence in fourth iteration and the calculation got stabled. In other words, w^4 is the final result.

A comparison of the test results is given in Table 20 and Figure 5.

Table 20 The Comparative Results For Different Measurement Scales

Scale type	priority
Saaty	$A_1>A_2>A_3>A_4$ 0.410 0.389 0.158 0.043
Power	$A_1>A_2>A_3>A_4$ 0.481 0.439 0.074 0.006
Geometric	$A_1>A_2>A_3>A_4$ 0.546 0.360 0.090 0.004
Logarithmic	$A_1>A_2>A_3>A_4$ 0.355 0.347 0.200 0.099

Scale type	priority
Root Square	$A_1>A_2>A_3>A_4$ 0.343 0.334 0.215 0.109
Inverse Linear	$A_1>A_2>A_3>A_4$ 0.410 0.292 0.216 0.082
Balance	$A_1>A_2>A_3>A_4$ 0.404 0.320 0.207 0.069

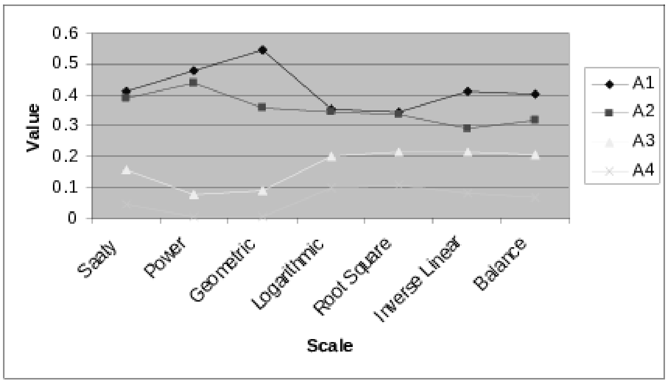

Figure 5 The Comparative Results

The change of pair wise comparison matrices element

In this section, we are looking for the answer to this question that (borrowing the idea from Mirhedayatian & Saen, 2011), *if one of the elements of the pair wise comparison matrix changes, what change in priorities occurs?* Consider the previous pair wise comparison matrices (table 19) (For instance, for Saaty's scale – table 21). So, a_{13} change from *3 to 1*:

Table 21 The Modified Pair Wise Comparison Matrix

1	1	1	9
	1	3	7
		1	5
			1

Similarly, by using Eq. (22), we have:

$W^1 = (0.366\ 0.366\ 0.224\ 0.044)$,

$W^2 = (0.327\ 0.405\ 0.224\ 0.044)$,

$W^3 = (0.336\ 0.399\ 0.220\ 0.045)$,

$W^4 = (0.336\ 0.398\ 0.221\ 0.044)$,

$W^5 = (0.336\ 0.398\ 0.221\ 0.044)$.

As can be seen, the process has been convergence in fifth iteration and the calculation got stabled. In other words, w^5 is the final result. Further, as can be seen from above results, the AHP is sensitive to change in a_{13} from *3 to 1*, for Saaty's measurement scales.

A comparison of the test results is given in Table 22 and Figure 6.

Table 22 The Comparative Results For Different Measurement Scales

Scale type	priority
Saaty	$A_2>A_1>A_3>A_4$ 0.398 0.336 0.221 0.044
Power	$A_2>A_1>A_3>A_4$ 0.531 0.318 0.145 0.006
Geometric	$A_1>A_2>A_3>A_4$ 0.458 0.373 0.165 0.005
Logarithmic	$A_2>A_1>A_3>A_4$ 0.356 0.304 0.242 0.099
Root Square	$A_2>A_1>A_3>A_4$ 0.338 0.302 0.250 0.110
Inverse Linear	$A_1>A_2>A_3>A_4$ 0.392 0.290 0.235 0.083
Balance	$A_1>A_2>A_3>A_4$ 0.375 0.319 0.237 0.070

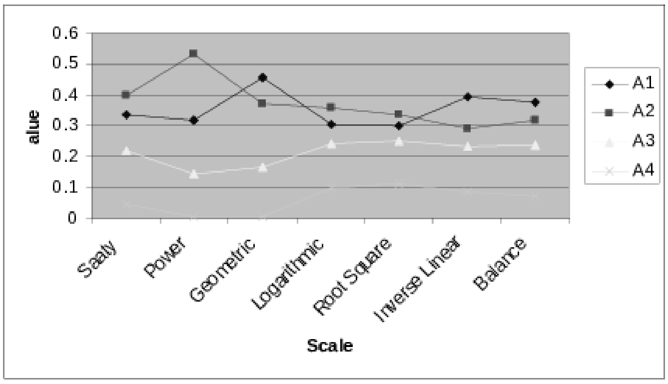

Figure 6 The Comparative Results

In addition, a comparative result for two previous tests (test numbers 10 and 11), is given in Table 23 and Figure 7.

Table 23 The Comparative Results For Different Measurement Scales (after & Before Change)

Scale type	Before	After
Saaty	$A_1>A_2>A_3>A_4$ 0.410 0.389 0.158 0.043	$A_2>A_1>A_3>A_4$ 0.398 0.336 0.221 0.044
Power	$A_1>A_2>A_3>A_4$ 0.481 0.439 0.074 0.006	$A_2>A_1>A_3>A_4$ 0.531 0.318 0.145 0.006
Geometric	$A_1>A_2>A_3>A_4$ 0.546 0.360 0.090 0.004	$A_1>A_2>A_3>A_4$ 0.458 0.373 0.165 0.005
Logarithmic	$A_1>A_2>A_3>A_4$ 0.355 0.347 0.200 0.099	$A_2>A_1>A_3>A_4$ 0.356 0.304 0.242 0.099
Root Square	$A_1>A_2>A_3>A_4$	$A_2>A_1>A_3>A_4$

Scale type	Before	After
	0.343 0.334 0.215 0.109	0.338 0.302 0.250 0.110
Inverse Linear	$A_1>A_2>A_3>A_4$	$A_1>A_2>A_3>A_4$
	0.410 0.292 0.216 0.082	0.392 0.290 0.235 0.083
Balance	$A_1>A_2>A_3>A_4$	$A_1>A_2>A_3>A_4$
	0.404 0.320 0.207 0.069	0.375 0.319 0.237 0.070

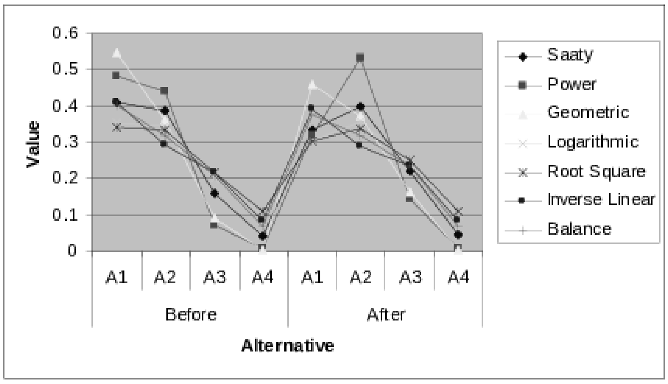

Figure 7 The Comparative Results (after & Before Change)

The consistency measure

Saaty (2000) defined a consistency index (*CI*) as (Eq. 23):

$$CI = (\lambda_{max} - n) \ / \ (n-1), \ \lambda_{max} = maximal \ \ eigenvalue \tag{23}$$

If we divide *CI* by the random consistency number for the same size matrix, we obtain the consistency ratio (*CR*). The value of *CR* should be around *10%* or less to be acceptable (Saaty & Kearns, 1991).

On the other side, according to Siraj (2011), the possibility of having an acceptable pair wise comparison matrix is sensitive to the selected scale of measurement. In addition, Franek and Kresta (2014) calculated the random indices (*RI*) for different measurement scales. For instance, for *n (or matrices size) =4*, we have (table 24):

Table 24 Measures Of RI For Judgment Scales Used In AHP* (particularly, For n=4)

-	Scale						
	Saaty	Power	Geometric	Logarithmic	Root Square	Inverse Linear	Balance
RI	0.882	6.987	9.299	0.241	0.179	0.333	0.440

*. For more details, see, Franek and Kresta, 2014.

In this section, Consider the following pair wise comparison matrices (i.e., *n=4*) (table 25):

Table 25 The Pair Wise Comparison Matrix

1	3	7	9
	1	5	5
		1	1/3
			1

By using *CR* formula (*CR = CI / RI*), we have (for example, for Saaty's scale):

CR = 0.088 / 0.882 = 0.099.

As can be seen from above result, the consistency ratio (CR < 0.1) is acceptable.

A comparison of the test results is given in table 26.

Table 26 The Inconsistency Measure For Different Measurement Scale

Scale type	**The inconsistency measures**
Saaty's (or Linear)	0.099
Power	0.074
Geometric	0.040
Logarithmic	0.141
Root Square	0.123
Inverse Linear	0.251
Balance	0.127

In continuing, table 27 represents a summary of the results of this analysis.

Table 27 A Summary Of The Results

Evaluation criteria Scale	**Test No. 1***	**Test No. 2***	**Test No. 3***	**Test No. 4***	**Test No. 5***	**Test No. 6***	**Test No. 7***	**Test No. 8***	**Test No. 9***	**Test No. 10***	**Test No. 11***
Saaty's (or Linear)**	Right side: 9 Both side:17	8.889, 0.111≤aij≤9	Right side: evenly filled Both side: not evenly filled	Right side: 0.500 Left side: 0.0.50	44/81	58/81	0.030	A1: 0.788 A2: 0.896 A3: 0.788 A4: 1	For a=1: $A_4>A_2>A_1\approx A_3$ 0.288 0.258 0.227 0.227 For a=3: A4>A2>A1≈A3 0.288 0.258 0.227 0.227	A1>A2>A3>A4 0.410 0.389 0.158 0.043	Before: A1>A2>A3>A4 0.410 0.389 0.158 0.0 After: A2>A1>A3>A4 0.398 0.336 0.221 0.0
Power	Right side: 9 Both side:17	80.988, 0.012≤aij≤81	Right side: evenly filled Both side: not evenly filled	Right side: 15.50 Left side: 0.085	44/81	58/81	0.074	A1: 0.768 A2: 0.768 A3: 0.596 A4: 1	For a=1: A4>A2≈A1>A3 0.319 0.245 0.245 0.190 For a=3: A4>A1>A2>A3 0.348 0.270 0.225 0.157	A1>A2>A3>A4 0.481 0.439 0.074 0.006	Before: A1>A2>A3>A4 0.481 0.439 0.074 0.006 A2>A1>A3>A4 0.531 0.318 0.145 0.0
Geometric	Right side: 9 Both side:17	255.996, 0.004≤aij≤256	Right side: evenly filled Both side:	Right side: 112 Left side:	0	36/81	0.063	A1: 0.842 A2: 0.406 A3: 0.258 A4: 1	For a=1: A4>A1>A2>A3 0.399 0.336 0.162 0.103	A1>A2>A3>A4 0.546 0.360 0.090 0.004	Before: A1>A2>A3>A4 0.546 0.360 0.090 0.004 A1>A2>A3>A4

Evaluation criteria Scale	**Test No. 1***	**Test No. 2***	**Test No. 3***	**Test No. 4***	**Test No. 5***	**Test No. 6***	**Test No. 7***	**Test No. 8***	**Test No. 9***	**Test No. 10***	**Test No. 11***
			evenly filled	0.187					For a=3: A4>A1>A2>A3 0.450 0.409 0.090 0.050		0.458 0.373 0.165 0.0
Logarithmic	Right side: 9 Both side:17	3.019, $0.301 \le aij \le 3.32$	Right side: not evenly filled Both side: not evenly filled	Right side: 0.920 Left side: 0.071	54/81	64/81	0.010	A1: 0.830 A2: 0.963 A3: 0.911 A4: 1	For a=1: A4>A2>A3>A1 0.270 0.260 0.246 0.224 For a=3: A4>A2>A3>A1 0.270 0.260 0.246 0.224	A1>A2>A3>A4 0.355 0.347 0.200 0.099	Before: A1>A2>A3>A4 0.355 0.347 0.200 0.099 A2>A1>A3>A4 0.356 0.304 0.242 0.0
Root Square	Right side: 9 Both side:17	2.667, $0.333 \le aij \le 3$	Right side: not evenly filled Both side: not evenly filled	Right side: 0.870 Left side: 0.067	49/81	64/81	0.013	A1: 0.862 A2: 0.951 A3: 0.914 A4: 1	For a=1: A4>A2>A3>A1 0.268 0.255 0.245 0.231 For a=3: A4>A2>A3>A1 0.264 0.256 0.246 0.233	A1>A2>A3>A4 0.343 0.334 0.215 0.109	Before: A1>A2>A3>A4 0.343 0.334 0.215 0.109 A2>A1>A3>A4 0.338 0.302 0.250 0.1

Evaluation criteria Scale	**Test No. 1***	**Test No. 2***	**Test No. 3***	**Test No. 4***	**Test No. 5***	**Test No. 6***	**Test No. 7***	**Test No. 8***	**Test No. 9***	**Test No. 10***	**Test No. 11***
Inverse Linear	Right side: 9 Both side:17	8.889, 0.111≤aij≤9	Right side: not evenly filled Both side: evenly filled	Right side: 2.70 Left side: 0.111	47/81	60/81	0.038	A1: 0.890 A2: 0.548 A3: 0.458 A4: 1	-	A1>A2>A3>A4 0.410 0.292 0.216 0.082	Before: A1>A2>A3>A4 0.410 0.292 0.216 0.082 A1>A2>A3>A4 0.392 0.290 0.235 0.0
Balance	Right side: 9 Both side:17	8.889, 0.111≤aij≤9	Right side: not evenly filled Both side: not evenly filled	Right side: 2.17 Left side: 0.019	53/81	63/81	0.049	A1: 0.851 A2: 0.655 A3: 0.540 A4: 1	-	A1>A2>A3>A4 0.404 0.320 0.207 0.069	Before: A1>A2>A3>A4 0.404 0.320 0.207 0.069 A1>A2>A3>A4 0.375 0.319 0.237 0.0

**. Test number: 1. the number of gradation, 2. the range of scale values, 3. the uniformity of value distribution, 4. the middle –scales error, 5. the discreteness of scale, 6. the finiteness of scale, 7. the discriminating level of alternatives, 8. the ratio preservation test, 9. changing the scale parameters, 10. The effect of scale type on priorities, 11. The change of pair wise comparison matrices element, 12. The consistency measure.*

***. As mentioned earlier, these results (especially, for Saaty's 1-9 scale) have been used as a standard to compare the performance of other scales.*

Discussion

Across the first three criteria, all scales provided nine gradations on the positive side and seventeen when reciprocal values were included. Saaty's 1–9, inverse-linear, and balanced scales also shared the same overall range and interval (8.889 and 0.111 ≤ aij ≤ 9), whereas the geometric scale produced the widest spread (255.996 and 0.004 ≤ aij ≤ 256). Distributional behavior differed more sharply: the geometric scale was uniform on both sides, while Saaty's linear and power scales were monotonic on the positive side and the inverse-linear scale was monotonic on the reciprocal side (Tables 2, 4, and 27).

The midpoint-error, discreteness, finiteness, discrimination, and ratio-preservation tests showed that the relative advantage of a scale depended on the property being assessed. Saaty's 1–9 scale had the smallest midpoint error on the positive side (0.500), while the balanced scale had the smallest reciprocal-side error (0.019); the geometric scale produced the largest midpoint errors. For discreteness, the geometric scale was the only scale without the identified inconsistency, and it also had the lowest finiteness inconsistency (36/81). The power scale produced the strongest discrimination between the two leading alternatives (0.074), whereas the logarithmic scale produced the weakest (0.010). Only Saaty's 1–9 scale preserved the benchmark ratios exactly in the ratio-preservation test (Tables 5, 7, 9, 12, 14, and 27).

For the remaining tests, changing scale parameters produced the same priority pattern for Saaty's linear scale at a = 1 and the logarithmic scale at a = 3. All seven scales preserved the ordering A1 > A2 > A3 > A4 in the scale-type test, although the priority intensities differed substantially. When one element of the pairwise-comparison matrix was changed, Saaty's 1–9, power, root-square, and logarithmic scales were more sensitive than the other scales. Consistency results favored Saaty's 1–9, power, and geometric scales, with consistency ratios of 0.099, 0.074, and 0.040 respectively; the remaining scales exceeded the benchmark by larger margins (Tables 18, 20, 23, 26, and 27).

Taken together, the tests do not identify a single alternative that can replace Saaty's 1–9 scale across all evaluated properties. Several alternatives outperform the benchmark on individual criteria, but those advantages are offset by weaker performance elsewhere. The evidence therefore supports a conditional rather than universal interpretation of scale superiority: scale choice can be criterion-specific, while Saaty's 1–9 scale remains the most balanced option across the complete twelve-criterion assessment.

CONCLUSION

This study compared seven AHP measurement scales using twelve evaluation criteria and treated Saaty's 1–9 scale as the benchmark. The numerical tests showed that alternative scales can outperform the benchmark on particular properties, including uniformity, midpoint error, discreteness, finiteness, or discrimination, but no alternative maintained that advantage across the complete evaluation set. Ranking outcomes were often stable across scales, while priority intensity, sensitivity, and consistency varied with the scale used.

Overall, Saaty's 1–9 scale displayed the most balanced performance across the twelve criteria and therefore remained the strongest general-purpose choice among the scales examined. The finding is consistent with the study's benchmark comparisons and with earlier work cited in the manuscript. For AHP applications, the results support retaining the 1–9 scale when no problem-specific evidence favors another scale, while allowing alternative scales when a particular evaluative property is especially important.

REFERENCES

Azadfallah, M. (2017$_a$). The impacts of aggregation rules, measurement scales and normalization methods on ranking of alternatives in AHP. *National Journal of System and Information Technology, 10(1)*, 1-22.

Azadfallah, M. (2017$_b$). The effect of scale ranges on priorities and discrimination level of alternatives in Analytic Hierarchy Process (AHP). *Journal of Business Applied Information Science, 5(1)*, 24-30.

Azadfallah, M. (2017$_c$). The impact of the scale elements alteration on priorities in Analytic Hierarchy Process technique. *International Journal of Business Analytics and Intelligence, 5(1)*, 43-51.

Azadfallah, M. (2025). Alternative Measurement Scales in Multi Criterion Decision Making Context, Journal of Management Research, 24(2),100-115.

Barzilai, J. (2006). On the mathematical modeling of measurement. http://arxiv.org/math.GM/0609555, Canada, 1-4.

Brans, J. P. and Mareschal, B. (2005). PROMETHEE methods. Ch. 5, in: Multiple criteria decision analysis- state of the art survey, (eds.): J. Figueira, S. Greco, and M. Ehrgott, Springer.

Cox M. A. A. (2009). Multidimensional scaling as an aid for the Analytic Network and Analytic Hierarchy Process. *Journal of Data Science, 7 (2009)*, 381-396.

Denscombe, M. (2017). The good research guide: for small-scale social research projects. McGraw-Hill Education (UK), 2017.

Dong, Y., Xu, Y., Li, H. and Dai. M. (2008). A comparative study of the numerical scales and the prioritization methods in AHP. *European Journal of Operational Research, 186(2008)*, 229-242.

Forman, E. and Selly, M. A. (2001). Decision by objectives. World Scientific Press.

Franek, J. and Kresta, A. (2014). Judgment scales and consistency measure in AHP. *Procedia Economics and Finance, 12(2014)*, 164-173.

Goepel, K. D. (2019). Comparison of judgment scales of the Analytic Hierarchy Process- A new approach. *International Journal of Information technology & Decision Making, 18(02)*, 445-463.

Grunig. R. and Kuhn, R. (2005). Successful decision making, a systematic approach to complex problem. Translated from German by: A. Clark and C. O'Dea, Springer-Verlag Berlin Heidelberg.

Harker, P. T. and Vargas, L. G. (1987). The theory of ratio scale estimation: Saaty's Analytic Hierarchy Processes. *Management Science, 33(11)*, 1383-1403.

Ishizaka, A. (2019). Analytic Hierarchy Process and its extensions, in: M. Doumpos et al. (eds.), new perspective in multi criteria decision making. Springer Nature Switzerland AG 2019.

Ishizaka, A. and Labib, A. (2009). Analytic Hierarchy Processes and expert choice: benefits and limitations. *OR Insight, 22(4)*, 201-220.

Ishizaka, A., Balkenborg, D., and Kaplan, T. (2011). Influence of aggregation and measurement scale on ranking a compromise alternative in AHP. *Journal of the Operational Research Society, 62(4)*, 700-710.

Ji, P. and Jiang, R. (2003). Scale transitivity in the AHP. *Journal of the Operational Research Society, 54(8)*, 896-905.

Lipovetsky, S. (2021). AHP in non linear scaling: from two-envelope problem to modeling by predictors. *Production, 31(2021)*, e20210007. doi: 10.1590/0103-6513.20210007.

Lootsma, F. A. (1993). Context-related scaling of human judgment in the multiplicative AHP, SMART, and ELECTRE, *working paper*, IIASA, Austria, 1-22.

Luce, R. D. (1997). Quantification and symmetry: commentary on Michell, quantitative science and the definition of measurement in psychology. *British Journal of Psychology, 88(1997)*, 395-398.

Lukinskiy, V., Lukinskiy, V., Bazhina, D. (2025). Refining the analytic hierarchy process: A statistical and methodological exploration of expert judgments. Alexandria Engineering Journal, 125(2025), 526-536.

Mirhedayatian, S. M. and Saen, R. F. (2011). A new approach for weight derivation using data envelopment analysis in the analytic hierarchy process. Journal of the Operational Research Society, 62(2011), 1585-1595.

Monat J. P. (2009). The benefits of global scaling in multi criteria decision analysis. *Judgment and Decision Making, 4(6)*, 492-508.

Mukherjee, K. (2017). Supplier selection-an MCDA-based approach. Springer (India) Pvt. Ltd. 2017.

Olson, D. L. (1996). Decision aids for selection problems. Springer-Verlag New York, Inc.

Ozsahin, I., Ozsahin, U. and Uzun, B. (2021). Applications of multicriteria decision making theories in: *healthcare and biomedical engineering, Ch. 2*, (eds.): I. Ozsahin, U. Ozsahin, and B. Uzun, Elsevier Inc.

Poyhonen, M. A., Hamalainen, R. P. and Salo, A. A. (1997). An experiment on the numerical modeling of verbal ratio statements. *Journal of Multi Criteria Decision Analysis, 6(1997)*, 1-10.

Qu, Z., Wan, C., Yang, Z., and Lee, P. T. W. (2018). A discourse of multi criteria decision making (MCDM) approaches. *(eds.): P T. W. Lee, Z. Yang, multi criteria decision making in maritime studies and logistics*, Springer International Publishing AG 2018.

Romanchak, V. (2023). The Problem of the Adequacy of the Analytic Hierarchy Process and Its Solution. In: Analytic Hierarchy Process - Models, Methods, Concepts, and Applications ,Chapter metrics overview, Chapter 5, Edited by: Fabio De Felice and Antonella Petrillo, DOI: 10.5772/intechopen.1001308, 2023.

Royo, C. S. and Chiclana, F. (2026). Contribution of the ratio scale of expert judgments in the analytic hierarchy process, Expert Systems with Applications,296, 2026, 128938, 1-11.

Saaty, T. L. (2000). Fundamentals of decision making and priority theory. RWS Publication, 6.

Saaty, T. L. (2001_a). Decision making with dependence and feedback, Vol. IX of the AHP series, RWS Publication.

Saaty, T. L. (2001_b). Deriving the AHP 1-9 scale from first principles. *Proceedings of the 6th ISAHP 2001 Bern*. Switzerland, pp. 397-401, August, 2001.

Saaty, T. L. (2008). Relative measurement and its generalization in decision making. Why pair wise comparisons are central in mathematics for the measurement of intangible factors the Analytic Hierarchy/ Network Process. *Rev. R. Acad. Cien. Serie A. Mat., 102(2)*, 251-318.

Saaty, T. L. and Daji, E. (2015). When is a decision making method trustworthy? Criteria for evaluation multicriteria decision making methods. International Journal of Information Technology & decision Making, IJITDM-D-15-00031R1, 1-18.

Saaty, T. L. and Kearns, K. P. (1991). Analytical planning, the organization of systems. Vol. IV, the AHP series, RWS Publications Pittsburgh, USA.

Saaty, T. L. and Vargas, L. G. (1991). The logic of priorities. RWS Publication, Pittsburg, Pennsylvania.

Salo A. and R. Hamalainen (1997). On the measurement of preferences in the Analytic Hierarchy Process. *Journal of Multi Criteria Decision Analysis, 6(1997)*, 309-319.

Sarie, W. S. (1997). Measurement theory: Frequently asked question. Version 3, Sep 14, 1997. URL: ftp://ftp.sas.com/pub/neural/measurement.html, USA.

Sato Y. (2001). The impact on scaling on the pair wise comparison of the Analytic Hierarchy Process. *ISAHP, 2001, Berns Switzerland, August 2-4*, 421-430.

Sequra, M., Maroto, C., Belton, V., Ginestar, C. and Marques, I. (2019). Collaborative management of ecosystem services in natural park based on AHP and PROMETHEE. *In: S. Huber, M. J. Geiger and A. T. D. Almedia (Eds.)*, Springer Natural Switzerland AG 2019.

Shinohara, M., Miyake, C., and Ohsawa, K. (2001). Why not use the Entropy method for weight estimations. *Proceedings-6th ISAHP 2001, Berns*, Switzerland, August 2-4, 2001, 431-434.

Siraj, S. (2011). Preference elicitation from pair wise comparison multi-criteria decision making. PhD thesis, University of Manchester, Supervisor: L. Mikhailov and J. Keane, UK.

Starczewski, T. (2017). Remark on the impact of the adopted scale on the priority estimation qualities. *Journal of Applied Mathematics and Computational Mechanics, 16(3)*, 105-116.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103(2684)*, 677-680.

Triantaphyllou, E., Lootsma, F. A., Pardalos, P., and Mann, S. H. (1994). On the evaluation and application of different scales for quantifying pair wise comparisons in fuzzy sets. *Journal of Multi Criteria Decision Analysis, 3(1994)*, 133-155.

Triantaphyllou, E., Shu, B., Sanchez, S. N., and Ray, T. (1998). Multi-criteria decision making: An operations research approach. Encyclopedia of Electrical and Electronics Engineering (J. G. Webster, Ed.), John Wiley & Sons, Newyork, NY, 15(1998), 175-186.

Voogd, J. H. (1982). Multicriteria evaluation for urban and regional planning. PhD thesis, Gouda, The Netherlands, 1982.

Wedley W. C. (2001). AHP answers to problems with known composite values. *ISAHP 2001*, Berns Switzerland, August 2-4, 2001, 551-560.