

Reflections on Common Misunderstanding When Using AHP and a Response to Criticism of Saaty's Consistency Index

Claudio Garuti
Fulcrum Ingenieria
claudiogaruti@fulcrum.cl

ABSTRACT

This essay deals with misunderstandings which, in this author's opinion, seem to prevail in the minds of some AHP critics. The main apparent problem is the lack of necessary knowledge about the scale concept including what a scale is, how it is mathematically defined, and what one can or cannot do with a scale. This knowledge is critical for a good understanding of AHP. This study will review the concept of scales, discuss common misconceptions, introduce the compatibility index G and explain how it addresses one of the previous fallacies perpetuated about AHP.

Keywords: AHP misconceptions; scales; compatibility index G

1. Introduction

There are different kinds of scales and what one can do with each one will depend on the properties of the scale. For instance, it is clear that with an ordinal scale you cannot make arithmetic operations (they are not allowed), but with a ratio scale you can. A scale is mathematically defined by its invariant function; the core of the scales and its properties come from this mathematical function.

Mathematical Definition of a Scale: Mathematically, a scale is a triplet composed of a set of numbers, a set of objects, and a transformation from the numbers to the objects. The scale also has a more abstract interpretation that only refers to the nature of the numbers and not to the objects, or how the numbers are assigned to the objects. The kind of transformation or the forms to create the numbers admissible for a particular measurement define what is called a scale of measurement for a measurement operation.

We have different kinds of transformation (invariant) and scales as follows:

The description of these scales can be found in Saaty, L.T. (2001), *The Analytic Network Process: Decision making with dependence and feedback* and in Garuti, C. and Escudey M. (2005), *Toma de Decisiones en Escenarios Complejos*.

Nominal Scale: Invariant under a one-to-one correspondence; for example, when a name or telephone number is assigned to an object, there is one and only one name and telephone number assigned to each object in the set.

Ordinal Scale: Invariant under monotone transformation, where the numbers order the objects, but the magnitude of those numbers are only useful for defining whether the

order is increasing or decreasing; for example, when assigning numbers 1 and 2 to two people to indicate that one is taller than the other, without including more information about their real height. The minor number can be assigned to the taller person or vice versa.

Interval Scale: Invariant under a positive linear transformation. For instance, the linear transformation $F = (9/5) + 32$, which transforms readings of temperature from Celsius to Fahrenheit. Notice that it is not possible to add two measures x_1 and x_2 in an interval scale, because then $y_1 + y_2 = (ax_1+b) + (ax_2 + b) = a(x_1+x_2) + 2b$, which takes the form of $(ax_3 + 2b)$ which is not in the form of $(ax + b)$ anymore. However, we can take the average of both readings because after dividing by 2 we are back to the original form. This is why 10 degrees of temperature plus 15 degrees of temperature does not produce 25 degrees of temperature (at most, its sum can make 15 degree).

Proportional Scale: Invariant under homogenous transformation, $y = ax$, $a>0$. An example is a transformation from pounds to kilograms with the transformation $K = 2.2P$. The proportion of the weights of two objects is the same and does not depend on whether the measurement was in pounds or kilos. The zero has no correspondence with any measurement of a real object; it is only applied to objects that do not present the property. It is not possible to divide by zero and get back a result we can interpret. We also may note that we can add two measures of the same scale $a(x_1+x_2) = a(x_3)$ which have the same form of ax . We can also multiply and divide different readings from the same proportional scales. When dividing two measures of the same proportional scale, the ratio of any two measures within it belongs to a proportional absolute scale. For example, $6\text{kg} / 3\text{kg} = 2$. The number "2" belongs to a proportional absolute scale, showing that the object weighing 6kg is double the weight of the object weighing 3kg. The number 2 is an absolute number because it cannot be transformed to any different number. The idea of an "absolute number" gives us the entry to present the following scale.

Absolute Scale: Invariant under the identity transformation $y = x$, (with $a = 1$), coming from the ratio: $y_1/y_2 = ax_1/ax_2 = 1x_3$. Examples of this scale are the numbers used to count people in a room, and the natural and real numbers (that is, those used to resolve equations). These absolute numbers are defined in terms of correspondence and equivalence classes of one-to-one correspondence following the postulates of the great Italian mathematician Peano, not in terms of some unit of measure starting from an origin at zero.

For the last 3 scales (interval, proportional and absolute), it is important to not confuse or mix the concept of the invariant of transformation (the concept that defines the scale) with the linear equation within the invariant of transformation. In doing so, we may fall into different conceptual errors. For instance, to believe that the absolute scale is just another proportional scale no different from any other ratio scale using as argument that the invariant of transformation of an absolute ratio scale (the identity function $y=x$) is just a particular case of the ratio scale ($y=ax$ with $a=1$). Although, this is true from the mathematical function point of view, when talking about scale and invariant of transformation it has a different meaning. An absolute scale is a different scale from a proportional scale; it has different properties and a different way of being built. For

instance, an absolute scale is dimensionless and its numbers are absolute numbers which means they cannot be transformed into another (number 3 is number 3 no matter what).

Another possible error (a kind of extrapolation of the last exposed error) is to believe that a proportional scale is just a particular case of the interval ratio scale, thinking that the transformation $y=ax$ is just a particular case of the transformation $\Delta^1y=a\Delta x$. But, we know that interval and proportional ratio scales are different scales since they have a different invariant, and thus different properties. If we fail to keep these differences in mind when working with the scales, then we may have several misconceptions, such as believing that the core of a ratio and absolute ratio scale is defined by the existence of a natural or absolute zero, but the zero (when existing) is just a consequence of the invariant transformation not its cause. For example, in thermodynamics, the absolute zero for temperature is defined as the detention of movement at the molecular level and it is represented as 0 degrees Kelvin or 0 K. This is the absolute (and also natural) zero for temperature; there is no temperature below 0 K. However, the presence of this natural zero doesn't make the Kelvin scale of temperature a proportional scale, it is still an interval ratio scale like any temperature scale (3K degree + 5K degree do not make 8K degree) and this is due to the characteristics of the temperature function.

On the other hand, suppose that we have built a scale for beauty. Does that mean that I know the zero for beauty? No, it doesn't mean that necessarily. It just means that I know the ratio between the elements of the set and nothing else. We have to note that this scale was built from singular elements (probably using a fundamental scale) thus is not a continuous function, but a discrete one. In this way, the zero value does not need to exist (it could exist but it doesn't need to). An important corollary is that the scale of ugly is not necessarily the inverse of the scale of beauty (they probably are 2 different scales). There is no need to reach one from the other passing through zero. Simply put, when I say that A is 3 times more beautiful than B, it doesn't mean (necessarily) that B is 3 times as ugly as A.

None of the absolute scales have a need for a unit or zero. In spite of the fact that it is said that a proportional scale has an absolute zero, this is only a supposition that makes it easier to work with the scale. Nowhere in the mathematical definition of a scale is it stated that it should have a unit and an origin with a zero value. (Saaty, 2001; (Garuti & Escudey, 2005). Therefore, the definitions that come from sources like "Scales for Dummies" are just an incomplete list of properties of scales and do not help to really understand what a scale is. In order for a fuller understanding, we need to know the invariant of the scale. There is the core of the scale.

2. Common errors in AHP use

With the scales definition clear, let's start discussing the first common error. The first error or confusion I have found was thinking of Saaty's fundamental scale as an ordinal scale. This initial confusion leads to the following errors:

¹ Δ stand by difference or interval

Error 1- Apply the Arrow's impossibility theorem to AHP numbers.

Arrow's theorem says that it is not possible to combine 3 different ordinal rankings (A, B, C and B, C, A and C, B, A) in a way that the 5 properties of the ranking are obeyed. Of course if we are working on a ratio scale (as in the case of AHP), then the impossibility becomes possible. Indeed, it is easy to combine different ratio scales into just one. Moreover, it is easy to demonstrate that a total inverse ordinal ranking (A, B, C and C, B, A for instance) can be more compatible (close) than another that has the same order (A, B, C and A, B, C) when A, B and C are numbers based in a ratio scale. This shows that trying to apply Arrow's theorem on cardinal numbers is a fallacy. Arrow's theorem is based on an ordinal scale (as Arrow correctly pointed out in his demonstration) and AHP is based on a ratio-absolute scale. Thus, Arrow's theorem is not applicable for AHP operations and results.

Error 2- How is it possible that 2 moderates (or weak) comparisons can make an extreme one?

This is very common mistake which is presented as: If $A=3B$ and $B=3C$, then $A=9C$ (two moderate or weak make an extreme). First, we have to remember that number 3 means moderate or weak and 9 means extreme in the verbal mode of Saaty's fundamental scale. But, Saaty's fundamental scale is an absolute ratio scale. Keeping this in mind, it becomes obvious that two moderate comparisons (3) must make an extreme (9), if you want to be totally consistent (or close to 9 if you want to be close to fully consistent). This is because in an absolute scale (as in any ratio scale as well) 3 times 3 is 9, not 4, not 5, not 6. Again, the problem here seems to be in the belief that Saaty's fundamental scale is an ordinal kind of scale. I think this confusion may come from the verbal mode of Saaty's scale. Some people seem to believe that the verbal mode of the scale makes the scale itself an ordinal one. That is not true; the verbal mode is just a way to make the numeric scale more operative over qualitative criteria. (By the way, this can also be made through a graphical mode and this mode or appearance does not take away any cardinality from the original scale).

Error 3- Why does Saaty's fundamental scale not have a zero?

This misinterpretation revolves around the claim about the lack of zero in Saaty's fundamental scale. Saaty's fundamental scale is based on an absolute ratio scale, and no absolute ratio scale has a zero (it can exist, but is not a necessity) because the absolute ratio scale is a scale built from the ratio of two ratio scales and zero is not a proportion to anything. Also, the neutral of the scale is the number 1 not zero. Thus, when I say that $a=b$, I'm really saying that $a/b=1$, and not necessarily that $a-b=0$. For instance, suppose that Helen and Betty are equally beautiful (or smart). When I say that Helen is as beautiful as Betty, I'm not defining an absolute zero for beauty; I'm just saying that the ratio of beauty between Helen and Betty is one, or that Helen is as beautiful as Betty. By the way, Saaty's scale is intended only for pair-comparison purposes. It is not made for evaluation purposes (I have seen different applications of AHP where the author uses the fundamental scale to evaluate the alternatives). The fundamental scale is a scale to build scales of measurement not to evaluate alternatives. Thomas Saaty himself said, "Build scales from measurement, not measurement from scales" (ISAHP2009, Pittsburgh 2009).

Error 4- What happens beyond 9 in Saaty's fundamental scale?

The next misinterpretation is a little bit trickier. It says: If $A=3B$ and $B=4C$, then $A=12C$ (due to the absolute-ratio scale consideration). But Saaty's fundamental scale only goes until 9; so, what happens beyond 9? To answer this objection, it must first be noted that Saaty's fundamental scale was built to be used by humans, and humans have a problem when trying to compare two very different objects. A person's precision decays exponentially as the ratio of comparison grows. Indeed, if anyone tries to guess the number of baseballs that fit into a one meter cubic box, they will have serious trouble getting the right number. However, their precision will increase rapidly if the size of the box is reduced to $1/3$. This is because now the size of the balls are relatively comparable with the box (relatively comparable means one order of magnitude, no more). Therefore, Saaty's 1-9 scale is a human capacity issue, not an AHP issue. If someone wants to use larger numbers for a quantitative criteria it is possible, but they must be warned about the difficulties. This last misinterpretation is also related to the second axiom of AHP, the Homogeneity Axiom. This axiom states that you have to compare homogenous objects, which means objects that are within one order of magnitude. When two objects do not belong to the same order of magnitude in some criterion, then the AHP model must be adjusted in a way that does not break the second axiom.

This last error can produce many different "new errors", even wrong demonstrations that AHP is incorrect or even that the Saaty's consistency index diverges. One such criticism is described in "On the Convergence of the Pairwise Comparisons Inconsistency Reduction Process" (Koczkodaj & Szybowski, 2015). For instance, a classic error puts criteria that are very important to the problem with others that have low importance or even irrelevant to the problem in the same level or cluster. This is a very typical "design error" in the AHP/ANP models. When this happens, it makes it highly probable that the output results (the priority ranking) will be incorrect. In fact, there was just such a case where a professor wrote a paper trying to show that the AHP was wrong. However, the professor was not aware of Axiom 2, and built a wrong model with heterogeneous elements in the same cluster. Of course, with a wrong model anything can happen (garbage in garbage out). By the way, in this example when the error (heterogeneous criteria) in the model is corrected the right results for the alternatives emerge.

It is important to mention, that this list of errors should be seen not only as errors produced by misunderstandings on AHP, but also a list of good practices for creating a model. This list of errors can be used to understand how to use Saaty's fundamental scale and also to know what kind of scale is being created when creating new scales of measurement so we can know what to do with that scale. These errors (and others) can be found at www.ResearchGate.net, under AHP issues and different papers published.

3. Consistency index misinterpretation and compatibility index G

Another error that can be produced due to the misuse of Axiom 2 involves Saaty's consistency index. The same basic example is used as in the criticism made by W.W. Koczkodaj and J. Szybowski (2015). This criticism says that Saaty's consistency index is too wide (infinitely wide indeed). The mathematical and logical errors of this argument against Saaty's consistency index will be discussed.

The logical argument is related to Axiom 2 of AHP that is using numbers beyond 9 in the PC-matrix. The mathematical argument is related with the reduction to the absurdum, using the compatibility index G for measuring in weighted environments combined with the logical fact that a zero vector is not a valid point of comparison with any vector in weighted environments.

First, the compatibility index G must be introduced. This index is useful in determining closeness (similarity) between vectors in a weighted environment (where the Euclidean measurement is not effective).

G is mathematically defined as: $G = \frac{1}{2} \sum_i (a_i + b_i) \text{Min}_i(a, b) / \text{Max}_i(a, b)$

G is a continuous real function that returns values within (0 – 1) range, with 1 representing total compatibility (A = B, parallel vectors) and 0 total incompatibility (A ⊥ B, perpendicular vectors). A, B are normalized priority vectors that belong to an absolute ratio scale within a weighted environment.

More details about the Compatibility Index G can be found in Garuti (2012), *Measuring in weighted environments: Moving from metric to order topology*, Garuti (2014) *Compatibility of AHP/ANP vectors with known results* and Saaty, L.T. (2010), *Group decision making: Drawing out and reconciling differences*. It is also interesting to read about the Jaccard index in Jaccard, P. (1901) *Distribution de la flore alpine dans le basin des dranses et dans quelques regions voisines*) since the reader may find that the G index seems to be a mathematical generalization of the J index (point to point).

Table 1 shows the meaning of ranges of compatibility in terms of index G and its description.

Table 1
Ranges of compatibilities and its meaning

Degree of Compatibility	Compatibility value range (G%)	Description	Compatible
Very High	≥ 90%	Very high compatibility Compatibility at cardinal level (Compatible vectors)	YES
High	85 – 89.9	High Compatibility (Almost compatible vectors)	
Moderate	75 – 84.9	Moderate compatibility (Not compatible vectors)	NO
Low	65 – 74.9	Low level of compatibility (Not compatible vectors)	
Very Low	60 – 64.9	very low compatibility (Almost incompatible vectors)	
Null (random)	< 60%	Random level of compatibility (Totally Incompatible vectors)	

The basic example (the criticism):

The critic says, "The consistency index of AHP (Saaty's index) is wrong since it may let some values (comparisons) pass that are not acceptable by common sense".

Description of the example:

Suppose there are three equal bars of the same length like the ones in Figure 1.



Figure 1 Bar length

Of course, the correct pair-comparison matrix (PC matrix) for this situation is the following (consistent) PC- matrix:

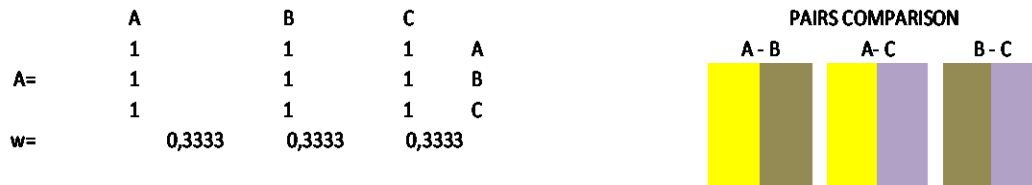


Figure 2 Bar comparisons 1

The obvious (correct) priority vector "w" is given by $w = \{1/3, 1/3, 1/3\}$ with 100% consistency (CR=0). This is because the bars are all equally long. Suppose now that (due to some visualization mistake) the new appreciation about the bars is as shown in Figure 3.

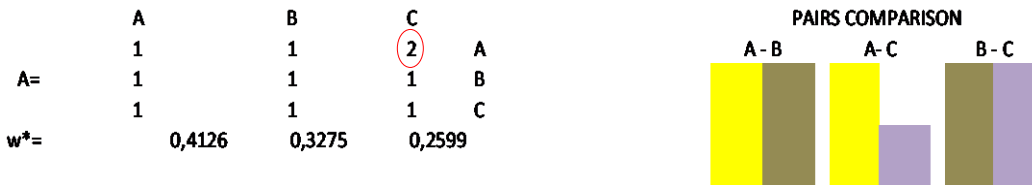


Figure 3 Bar comparisons 2

The new perturbed priority vector is: $w^* = \{0.4126, 0.3275, 0.2599\}$, with CR = 0.05 (95% of consistency) which according to the theory is the maximum acceptable CR for a 3x3 PC- matrix. Also, 2 is the maximum possible value if we want to stay within 95% consistency. The critic claims that the A-C bar comparison has a 100% difference (100% of error) which is not an acceptable or tolerable error (easy to see even with the naked eye). Also, the global error (deviation) in the priority vectors is 15.85%, calculated with the common formula: $e = \sum \text{Abs}(w^* - w)$, for each coordinate and then added over the

coordinates. But, Saaty's consistency index says that $CR=95\%$ (or 5% of inconsistency), is a tolerable limit for a 3x3 PC matrix. Hence, the critic claims that Saaty's consistency index is wrong (useless and mathematically unsound to precisely quote the critic).

The Response:

The critic misunderstands two important things. First, the $CR=CI/RI$ (the Saaty's index of consistency) comes from the eigenvalue-eigenvector problem, so it is a systemic approach. Thus, it is not concerned with any particular comparison, like comparison in the (1, 3) matrix position in this case. Second, the possible error should be measured by its final result (the resulting metric) not in the prior or any middle step.

The first misunderstanding is self-explanatory (systemic approach). For the second one, before any calculation, we need to understand what kind of numbers we are dealing with (in what environment we are working) because it is not the same for errors and deviations to be closer to a big priority than to a little one. This is a weighted environment and the measure of closeness (proximity) and thus possible errors has to be considered in this situation. We must work in the order topology domain to correctly measure the closeness of two vectors or rule of measurement on this environment. To do this task correctly, two aspects of the information have to be considered, the intensity (the weight or priority) and the degree of deviation between the two priority vectors (geometrically it can be seen as the projection between the vectors). The only index that takes good care of these two factors simultaneously is the compatibility index G.

The Explanation:

Graphically we are trying to make the following reasoning by demonstrating the reduction to the absurdum.

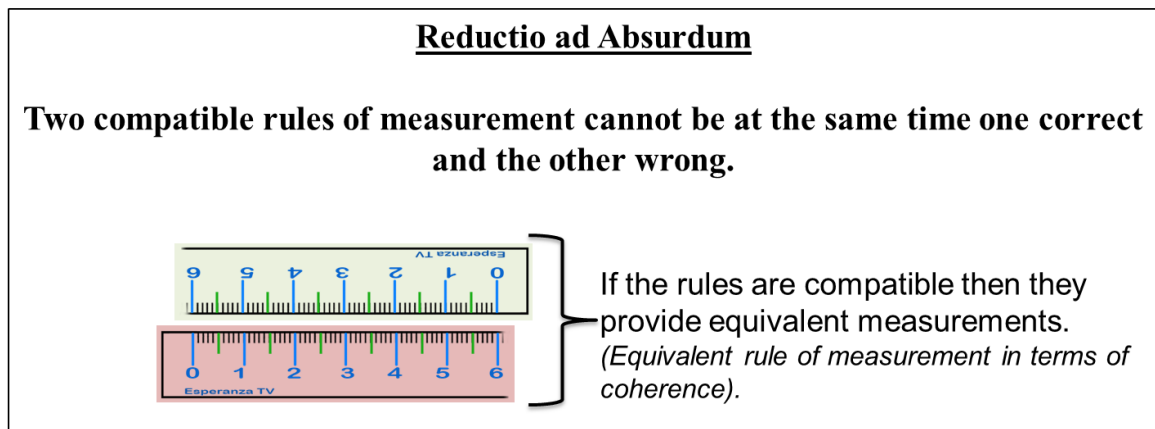


Figure 4 Compatible rule of measurements

If Saaty's consistency index is wrong, then the reference and perturbed metrics (w and w^*) cannot be compatible (cannot measure the same), since we know one is correct (w) and the other is supposed to be wrong (w^*).

Summarizing the vectors of the correct and perturbed metric:

Correct metric (priority vector w) :	0.3333	0.3333	0.3333
Perturbed or approximated metric (priority vector w*) :	0.4126	0.3275	0.2599

Thus, the basic question is, “how close is the perturbed metric to the correct metric?” To respond to that question it is necessary to evaluate the similarity (or closeness) of the two priority vectors, and this job is correctly performed by the index of compatibility (G). Assessing G (Correct vs Perturbed), the value obtained is: G=85.72% which in numerical terms represents almost compatible metrics. (See Table 1). G=90% is a threshold to consider two priority vectors as compatible vectors. Also, G=85% is an acceptable lower limit value. Hence, the two metrics are relatively close (close enough considering that they are not physical measures).

We have shown that both metrics are compatible that is they are measured almost with the same rule (as shown in Figure 4). One rule cannot be correct and the other wrong if they are compatible rules. Thus, the second rule (w*) is an acceptable rule of measurement, and by reduction to the absurdum, the criticism about Saaty's consistency index is incorrect.

It is important to note that the same exercise was performed with 4x4 to 9x9 PC matrices. The value (n-1) was put in the position cell (1, n), (n= matrix dimension) and even better results were obtained for the compatibility index G than in the 3x3 dimension matrix. This is shown in the cells in bold in Table 2.

Table 2
Compatibility indices for perturbed matrices from range 3x3 to 9x9

Inconsistency												
5%	3x3	0,333333333	0,333333333	0,333333333	1							
	(1,3)==>2	0,4126	0,3275	0,2599	1							
		0,30131416	0,324634375	0,231272015		85,7%						
6%	4x4	0,25	0,25	0,25	0,25	1						
	(1,4)==>3	0,331	0,2407	0,2407	0,1888	1,0012						
		0,219410876	0,23622298	0,23622298	0,16569088		85,8%					
6%	5x5	0,2	0,2	0,2	0,2	0,2	1					
	(1,5)==>4	0,277	0,1906	0,1906	0,1906	0,151	0,9998					
		0,172202166	0,1861209	0,1861209	0,1861209	0,1325025		86,3%				
5%	6x6	0,1667	0,1667	0,1667	0,1667	0,1667	0,1667	1				
	(1,6)==>5	0,2392	0,1582	0,1582	0,1582	0,1582	0,1279	0,9999				
		0,14139725	0,15418172	0,15418172	0,15418172	0,15418172	0,11302523		87,1%			
4%	9x9	0,111111111	0,111111111	0,111111111	0,111111111	0,111111111	0,111111111	0,111111111	0,111111111	0,111111111	1	
	(1,9)==>7	0,1717	0,1055	0,1055	0,1055	0,1055	0,1055	0,1055	0,1055	0,0897	0,9999	
		0,091506863	0,102836125	0,102836125	0,102836125	0,102836125	0,102836125	0,10283613	0,10283613	0,08105741	0,89241714	89,2%

The outcome of Table 2 is not a surprise since it comes from a matrix built within an intrinsic systemic behavior. The pair comparison process in the matrix produces highly

related elements among the pairs. When searching for the equilibrium point of the matrix (the eigenvector that represents the metric of the matrix) this process of relations and interconnections can be perceived as a growing complex system (as the graph theory shows). Thus, the analysis of the quality of the consistency index must be done considering this relevant fact (complex system) with many connections and redundancies (Garuti, C., Salomon, V. & Spencer, I., 2008).

Redundancies (redundant judgments) are necessary and very important because they give more reliability to the system (any system without redundancy is a vulnerable or fragile system). In addition, these redundancies give more stability to the system because they allow for an acceptable result even with a cell in the matrix having a very bad pair comparison value. When a system allows for redundancies it has the capacity to receive new information that may or may not be consistent with the old one. This characteristic allows the system to evolve by connecting old and new data in a peaceful way.

For instance, in Table 2 for the case of the 9x9 matrix in position (1, 9), there is a 8 instead of 1; that is a "very large error" of 800%, (following the bar example, the bar would be visualized as 8 times bigger on this specific comparison) and in spite of that we still obtained a healthy outcome of 89.2% for compatibility index G (even better than the rest of the G values). Even more, when performing the hypothetical case for n=15 (15x15 matrix) with a value of 14 in the cell position (1, 15), (a huge 1400% error), the outcome for compatibility index G is 99.96% (almost 100% of compatibility). This means that both vectors (correct and perturbed) are almost identical in terms of measurement. Thus, the trend of compatibility clearly shows that the divergence in the value of a cell (1, n), (or any other cell by the way) does not produce any decay in the quality of the generated metric. By the way, the Saaty's consistency index for this case was 3%, a very good consistency index for a 15x15 matrix (a very large and inadvisable matrix).

It is interesting to know what happens if we leave "n" (the size of the matrix and the value for the pair comparison position (1, n)) free to grow beyond 9. In this case, the acceptable error in Saaty's consistency index may diverge (go to infinity). This means that in a very large PC matrix you can put a pair comparison number as big as you like and still get an acceptable ratio of consistency. This abnormal behavior for Saaty's consistency index is also revealed by the compatibility index G. As "n" the size of the matrix and the pair comparison value in the (1, n) position increase, G decreases making the priority vector (the final or perturbed metric) more and more incompatible with the reference vector. For instance, in a 15x15 PC matrix the value for cell (1, 15) can be as large as 50 (5000% "error"), a very large value (beyond one order of magnitude) and still have an acceptable consistency (10%). For this case, the output value for G is 79.7%, which according to Table 1 is not an acceptable value for the quality test (not compatible vector). Nevertheless, it is interesting to understand what it means to leave "n" free to grow in a weighted environment.

We have to remember that our reference vector or metric is defined by: $\{1/n, 1/n \dots 1/n\}$, then, as n grows the reference vector tends toward the vector $\{0, 0 \dots 0\}$, the null vector. The null vector (or zero vector) is not a point of reference for anything in a weighted environment (using this vector as a reference point is like dividing by zero in a mathematical demonstration). Thus, the useful mathematical concept of limit analysis

behavior is not applicable for this situation. In conclusion, we cannot leave “n” free to grow as presented in the demonstration that Saaty’s consistency index is wrong. It is also interesting to reflect here about the great relevance of Axiom 2 of AHP (the homogeneity axiom) which states that one must not make comparisons beyond one order of magnitude. This problem or inconsistency is possible in very large matrices (that is, when we leave “n” free to grow).

There are several important conclusions that can be made from this example. From the first misunderstanding, that the $CR=CI/RI$ (the Saaty’s index of consistency) comes from the eigenvalue-eigenvector problem and thus is a systemic kind of approach we are not concerned with any particular comparison but the full PC-matrix (the whole set of comparisons). Also, redundancies are important and need to be captured (as in any systemic approach, redundancies are fundamental in the system behavior). For the second misunderstanding that says the possible error should be measured by its final result (the resulting metric) not in the prior or any middle step, it is important to note that we need to understand what kind of numbers we are dealing with (what kind of environment are we working in before any calculations. This is because (and just as an example) for errors and deviations it is not the same to be close to a big priority than to a little one. This is a weighted environment and the measurement of the closeness (proximity) and thus possible errors has to be considered in this situation. So, we cannot just determine the difference of coordinates between two vectors to find their closeness. Moreover, in general, MCDM belongs to the order topology domain, and we must work in this domain to adequately measure the closeness in this environment. To do this task correctly, two aspects of the information have to be considered, the intensity (the weight or priority) and the degree of deviation between the two priority vectors (the projection between the vectors). The only index that considers these two factors simultaneously is the compatibility index G .

Additional Caveats:

1. When possible, all pair comparisons in the matrix have to be done, and all of the pair comparisons have to take into account the rights and the wrongs, which by the way are indistinguishable, to correctly assess consistency and priority (the weighted metric of the matrix). This is where the eigenvector operator (and its principal eigenvalue) perform the best.
2. The analysis of the behavior of an isolated element to characterize the whole system (this would be a kind of basic mechanical analysis) is not valid. The PC-matrix represents a highly related system. Thus, it is not possible to evaluate a complex system of behavior (the PC matrix) by the behavior of only one of its elements; there are redundancies that are not well captured in the one isolated element analysis.
3. If you want to have a representative consistency index (without a very large bad comparison), you should never go beyond a matrix size of 9 (9x9 matrices) in any PC matrix. This is because it is very probable that you may have comparisons beyond 9 inside the matrix as the size of the matrix increases. By the way, this is aligned with Axiom 2 of AHP which is in keeping with the homogeneity factor (not going beyond one order of magnitude between the elements to be compared). Breaking Axiom 2

- may produce a loss of consistency as well which is another source of error that normally comes from a poor modelling process.
4. It is not possible to apply the common formula for error measurement as: $e = \frac{\sum \text{Abs}(w^* - w)}{n}$ within a weighted environment since it is not the same to be close in a high weighted element as in a low one. (This common formula works OK in the Euclidean or flat space, but not in a weighted one).
 5. Finally, it seems that the threshold of 10% of Saaty's consistency index could be too lax for some cases. Indeed, this threshold is 5% for 3x3 matrices and if we want to keep an acceptable level of compatibility the threshold should not go beyond 6 or 7% in matrices of superior order (instead of the current 10%). Nevertheless, this also depends on the kind of problem being solved. However, we think that more investigation and numerical tests should be carried out in this line of research.

A final important comment about consistency:

Of course, better consistency (100% for instance) can always be achieved. The question is, “do we really obtain a better result when being totally consistent?” The answer is probably no. Because, in real problems we never have the “real” answer (the true metric “w” to use as reference). Experience shows that pursuing consistent metrics per se may provide less sustained results. For instance, in the presented problem, one could answer that: $A/B=2$, $A/C=2$, and $B/C=1$, as shown in Figure 5, and he/she would be totally consistent (but consistently wrong).

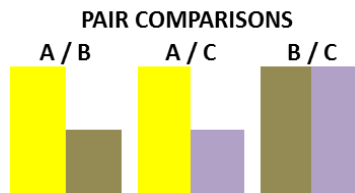


Figure 5 total consistent pairwise matrix

In Figure 5, the new priority vector is $w^{**} = (0.5, 0.25, 0.25)$, with $CR=0$ (totally consistent), and $G= 71.5\%$, which means not compatible vectors (low compatibility). Thus, a totally consistent metric is incompatible with the correct result. Hence, in the end it is better to be approximately correct than consistently wrong. The consistency index is just a thermometer not a goal. It is also important to note the fact that the quality of the metric of any PC-matrix is not found in some specific PC judgment of the matrix. It should be found in its final interaction, which means after all the relations and redundancies have played their part in the search of an equilibrium point of the system (the principal eigenvector). Therefore, defining the quality of the metric that comes from a PC-matrix directly from the matrix values (the judgments), as presented in this criticism example, is not a good idea.

4. Conclusions

The list of errors that have been presented have to be seen not only as errors produced by misunderstandings, but also as a list of good practices for creating a model. They can help show how to correctly use Saaty's fundamental scale and also show how to know what kind of scale is being created and thus what can and cannot be done with that scale.

The bar example presented in the paper shows that systemic behavior cannot be analyzed by its elements in a separate way. The PC matrix is a system and the pair comparison judgments are its elements. The G index shows that one poor comparison (very poor indeed) can be present and an acceptable quality metric can still result (an acceptable priority vector). The relevant conclusions from this example were presented earlier in the paper and are a clear rebuttal for the criticism that claims Saaty's consistency index is useless and mathematically unsound.

REFERENCES

- Garuti, C. (2012). Measuring in weighted environments: Moving from metric to order topology. In F. De Felice, A. Petrillo, T. Saaty (Ed.), *Applications and Theory of Analytic Hierarchy Process - Decision Making for Strategic Decisions* (pp. 247-275) Santiago, Chile: Universidad Federico Santa Maria. Doi: 10.5772/63670
- Garuti, C. (2014b). Compatibility of AHP/ANP vectors with known results. *ISAHP2014*
- Garuti, C., Salomon, V., Spencer, I. (2008). A systemic rebuttal to the criticism of using the eigenvector for priority assessment in the Analytic Hierarchy Process for decision making. *Computación y Sistemas (Review)*, 12(2), 192-207.
- Garuti, C Escudey M. (2005). *Toma de decisiones en escenarios complejos*. Santiago-Chile: Editorial Universidad de Santiago Publications.
- Jaccard, P. (1901). Distribution de la flore alpine dans le basin des dranses et dans quelques regions voisines. *Bulletin de la Société Vaudoise des Sciences Naturelles* 37, 241-272. Doi: <http://dx.doi.org/10.5169/seals-266440>
- Mahalanobis, P.C., (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Science of India*, 12, 49-55.
- Saaty, L.T. (2001). *The Analytic Network Process: Decision making with dependence and feedback*. Pittsburgh, PA: RWS Publications.
- Saaty, L.T. (2010). *Group decision making: Drawing out and reconciling differences*. Pittsburgh, PA: RWS Publications. Doi: <https://doi.org/10.13033/ijahp.v9i3.533>
- Koczkodaj, W.W., Szybowski, J. (2015). *On the convergence of the pairwise comparisons inconsistency reduction process*.