Requiem for risk classification matrices

W.C. Hachich, Ph.D.

Professor of Geotechnical Engineering, University of São Paulo, Brazil.

Abstract

Classification matrices are scrutinized for inconsistencies, errors and deficiencies in meaning. Proper definition, measurement and ranking of risks are demonstrated as compelling arguments whenever risk and reliability analyses are required.

1. Introduction: measurement scales

Classification is a most fundamental organizational activity. It may involve, for example, grouping in classes or categories objects which exhibit similar characteristics that distinguish them from others that do not. For such a purpose a nominal scale is enough. Figure 1, for example, presents a classification of tailings dams in a region, grouped exclusively in accordance with the construction procedure.

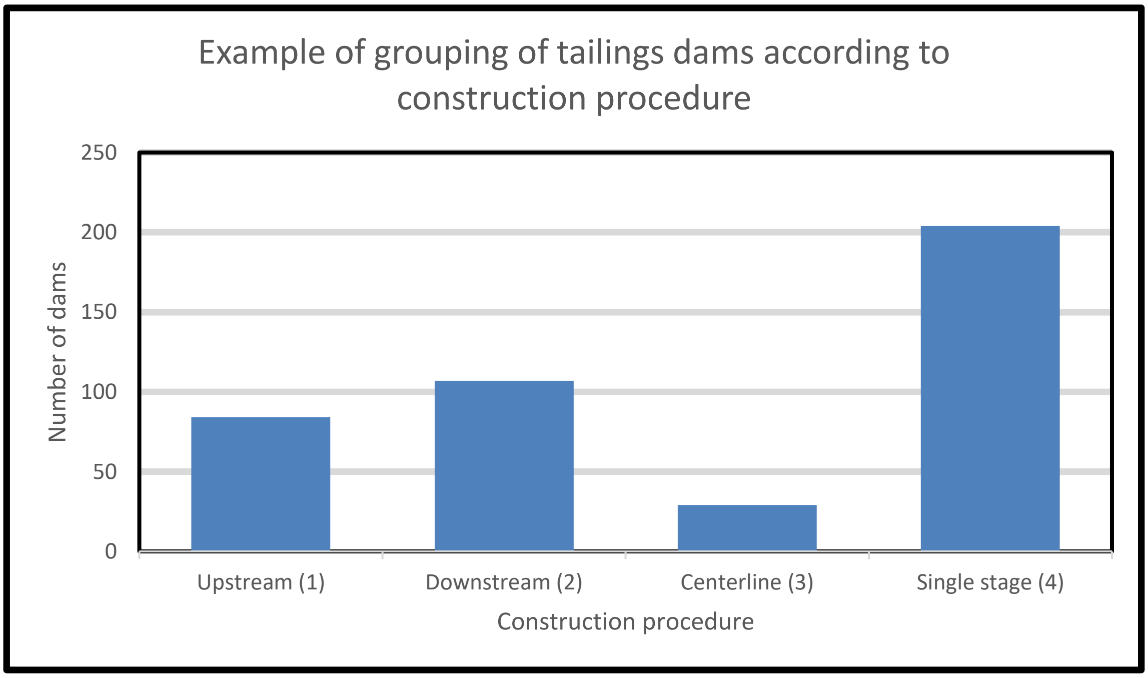

Figure 1 – Example of classification of a group of tailings dams according to construction procedure.

It should be quite clear that even if numbers may be used to identify different categories, none of the usual mathematical operations are valid on these numbers, because they just serve the purpose of nominating classes (thus *nominal* scale).

The data in Figure 1 are presented again in Figure 2, but in order of increasing vulnerability. The term risk is being purposely avoided at this point, while vulnerability is being temporarily proposed as a rather intuitive concept associated with the adopted construction procedure.

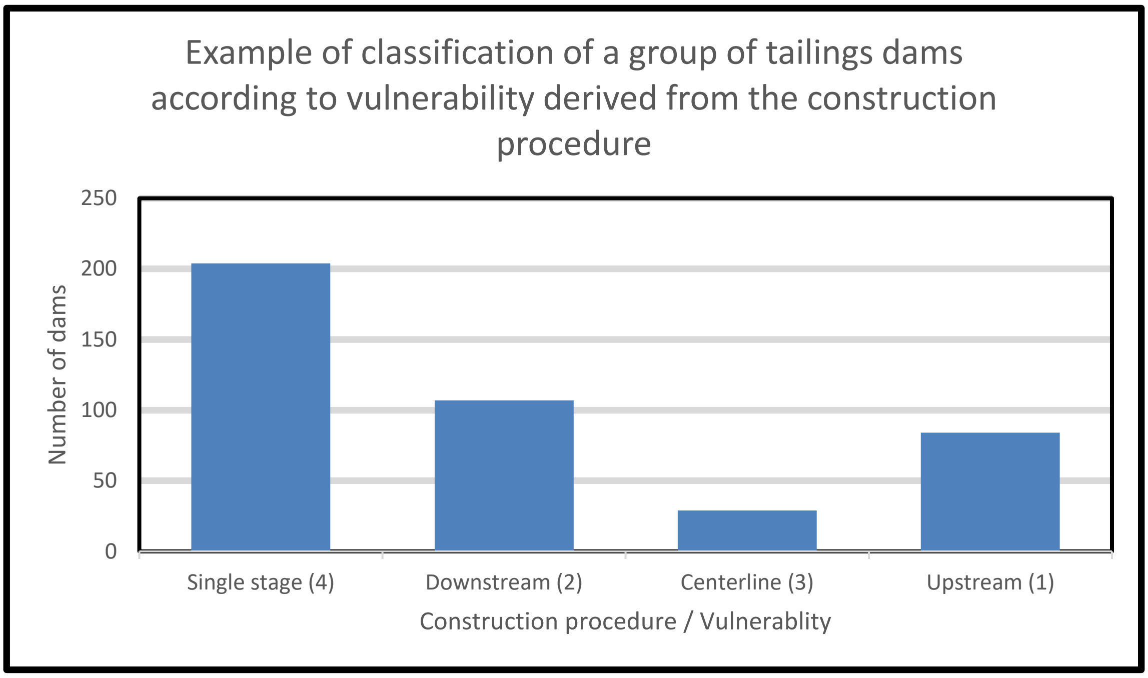

Figure 2 – Example of classification of a group of tailings dams according to vulnerability derived from construction procedure.

Figures 2 and 3 provide evidence of statements by Ackoff (1962) and other theoreticians of measurement:

> *"The use of a letter or a word is no less measurement than is the use of a number,* ***provided that we make explicit, as we must in the case of numbers as well, what operations may be performed on the symbols****.*
>
> *Measurement is a way of obtaining symbols to represent the properties of objects, events, or states, which symbols have the same relevant relationship to each other as do the things which are represented."*

No mathematical operation can or should be performed on those symbols. Figure 3 (a and b) show how one could order, sort or rank objects in accordance with a chosen criterion, such as those 12 dams investigated for risk,

In Figure 3a dams are ordered according to their identification letter (or number). The choice between letters or numbers is irrelevant for this *nominal scale*. Dam H is equally identified as dam 8, but mathematical operators must not be applied to those identifiers.

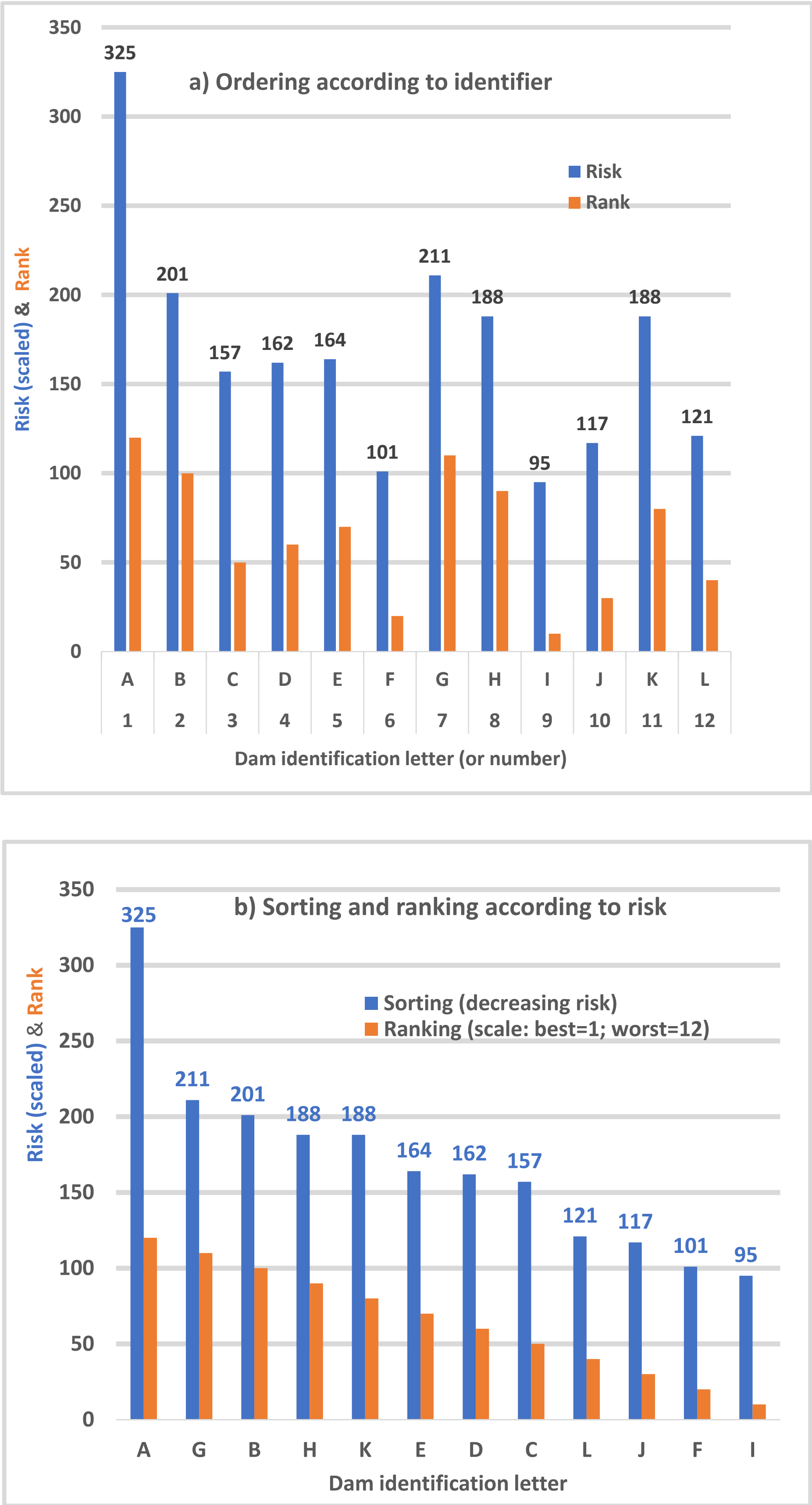

Figure 3 – Example of ordering, sorting and ranking a group of dams.

Risk for dam C, for example, is 157, and it ranks 5.

In its strict sense, measurement involves the use of a constant measurement unit. This unit can be arbitrarily established when there is no natural zero, such as in the case of the Celsius and Fahrenheit scales. In such cases mathematical operations can be performed on intervals, but not directly on the values themselves. Those are called *interval scales*.

When there is a natural zero, such as in the scales of length, weight, and so on, all usual mathematical operations are valid for the numbers that express the measurements and those scales are called *ratio or proportional scales*. 40 centimeters, for example, is twice 20 centimeters. One cannot say, however, that 40 degrees Celsius is twice 20 degrees Celsius, while it is possible to say that the difference in temperature between 20 and 40 is equal to the difference between 40 and 60 degrees Celsius.

A ratio scale is usually preferred over any of the others because it is more informative about the measured quantity. Given our interest in the risk associated with dams, the question is obvious: can a ratio scale be devised to appropriately measure risk?

2. Measurement scale for risk

The answer to that question must be based upon the definition of risk itself, as firmly established in the field of *Risk Analysis*: risk involves a combination (product) of probability of a certain action (or hazard) and the consequences thereof (it is worth noting that the insurance industry uses a different definition).

Thus, Risk Analysis defines risk as the probability of an event multiplied by its consequences (Figure 4, Hachich, 2002). Consequences are seldom just economical. For the sake of conciseness, other types of consequences, such as social and environmental, which are obviously equally relevant from a practical standpoint, are not going to be explored in this paper, given that the fundamental flaw of risk matrices can be demonstrated on the basis of just one type of consequence (Pratt, Raiffa, Schlaiffer, 1965).

As a matter of fact, the proper definition of risk and its use for classification of any group of dams is the crucial point of this paper.

$$\textbf{Risk} = \textbf{E[C]} = \textbf{p} \cdot \textbf{C*}$$

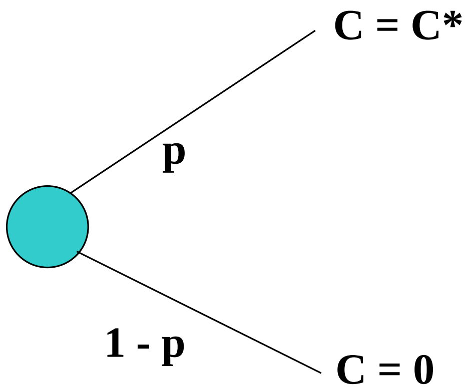

Figure 4 – Risk as the product of uncertainty and consequences. (Hachich, 2002).

When hazards present themselves at several levels, each of them associated with a certain probability and consequence, risk is computed as a weighted average of the consequences, having probabilities as weights (Figure 5, Hachich, 2002). Risk is, therefore, the *expected value* of consequences. The risk associated with the circumstances represented by Figure 6, for example, is quantified by the area below the dotted line.

$$\boldsymbol{Risk = E\,[C] = \sum_{1}^{n} \; p_i \times C_i}$$

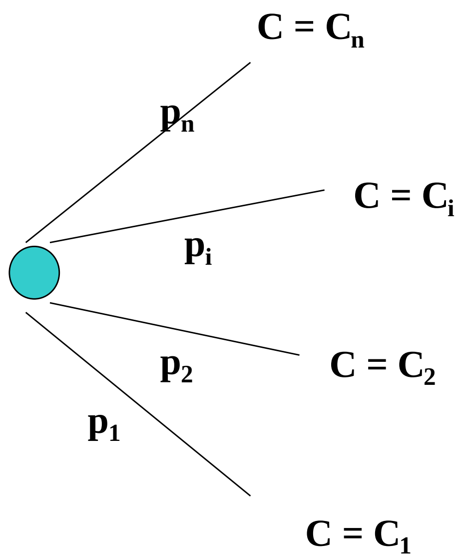

Figure 5 - Risk as the expected consequence. (Hachich, 2002).

The unduly and conceptually wrong use of matrices for risk classification has been criticized for almost 20 years Hachich (2002). The final objective of risk evaluation is to provide guidance as to decisions that have to be made. It is therefore natural that risk be interpreted within the context of *Decision Analysis* (Raiffa, 1968) and *Utility Theory*.

As previously pointed out, the definition in Figure 5 corresponds to the application of the expected value operator to the consequences. If one considers several different sets of circumstances, each with a graphical representation similar to that in Figure 5, the pairs [C;p] define the curve in Figure 6. The area below the curve is a measure of the risk associated with the decision that led to the situation depicted in Figure 5. The decision corresponding to the curve delimiting the smallest area would be preferred over any other. As previously pointed out, those areas need not (or perhaps should not) be restricted to economical values: as a matter of fact, if *Utility Theory* is invoked to assign values of *utilities* to different combinations of economical, social and environmental consequences, Decision Analysis can be applied to more general situations (Keeney and Raiffa, 1976.).

The preference for ratio scales has been previously stated. Probabilities are measured between zero and one in a ratio scale. Consequences are also be measured in a ratio scale, and utilities can also be defined between zero and one. Given the definition of risk, there is no reason whatsoever why it should not be measured in a ratio scale.

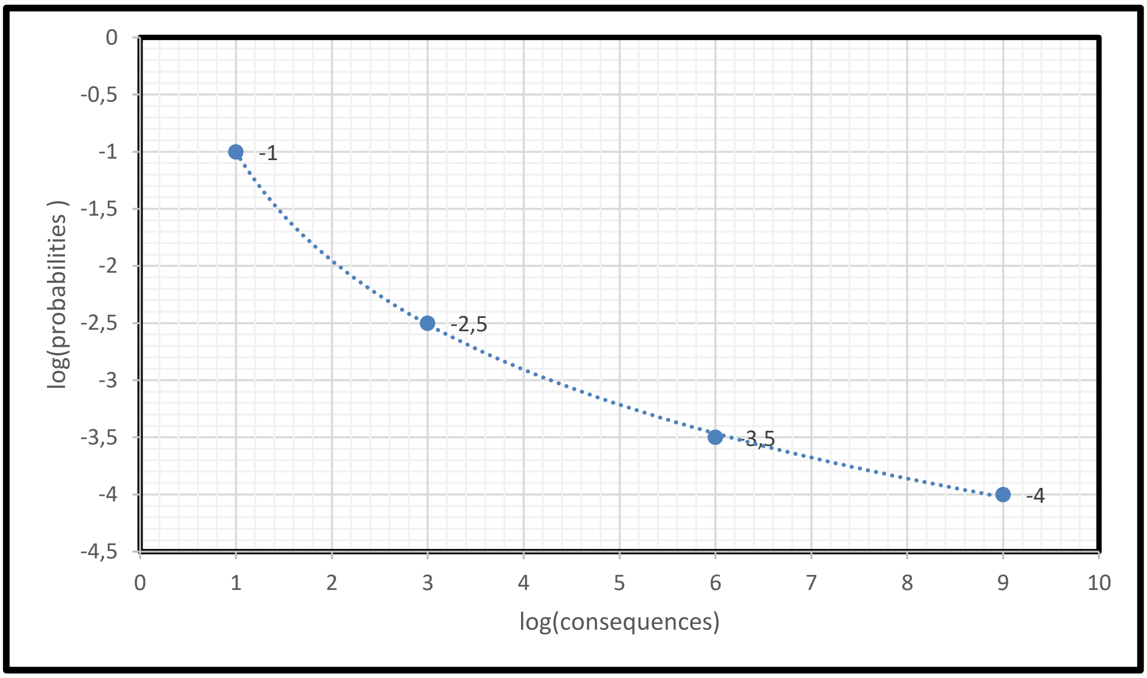

Figure 6 – Example of graphical representation of risk on the probability-consequence space (adapted from Oboni, 1998).

Figure 3b presented the classification of a set of dams on the basis of risks posed by them. Classification must start, of course, with the evaluation of risks, and that is the only way of doing it correctly.

3. “Risk” classification matrices

Our interest is focused, of course, in those dams that pose higher risks: they should be the priority of mitigating actions. Given the definition of risk, its evaluation requires studies of some complexity performed by a team of engineers capable of evaluating probabilities of geotechnical, hydrological, hydraulic and many other engineering-related events, in addition to their consequences (and possibly utilities as well).

In some cases it is known beforehand that risks are not small because of the construction procedure, the lack of information and contingency plans, faulty conservation and many other reasons. In such cases it is usual to see tables like Table 1, published with a wrong definition of risk by the way. The scale adopted for the table is obviously nominal, even if someone decides to exchange symbols for numbers in the cells, such as in Table 2. It follows that mathematical operations performed on those numbers are not acceptable.

The inconsistencies of such an approach are further explored in Hachich (2002). Re-stating Ackoff (1962): used symbols, such as numbers, must have the same relevant relationship to each other as do the things which are represented. In the present case, our interest in risks would require such numbers to be values measured in a ratio scale, so as to represent actually computed risks.

Table 1 – Example of “risk” matrix with the usual type of symbol-based nominal scale (*ad hoc* chosen characters).

	Potential damage		
“Risk”	High	Medium	Low
High	A	B	C
Medium	B	C	D
Low	B	C	E

Table 2 - Example of “risk” matrix with the usual type of symbol-based nominal scale (*ad hoc* chosen digits).

	Potential damage		
“Risk”	> 1000	1 to 1000	< 1
>0.01	5	4	3
0.0001 to 0.01	4	3	2
<0.0001	4	3	1

The possibility of "risk" scales such as those in Table 2 leading to decisions that reflect the decision maker's preferences is never demonstrated, while Decision Analysis and Utility Theory offer mathematical proof (Keeney & Raiffa, 1976). Surprisingly, however, arbitrarily chosen nominal scales is one of the most ubiquitous features of published papers on "risk" assessment. Table 3 is just one such example, borrowed from a real-world situation.

Table 3 - Example of "risk" matrix with the usual type of symbol-based nominal scale (arbitrarily chosen description/classification and "corresponding" digits).

Seepage (e)	**Displacements (f)**	**Flood return period (g)**
Perfectly controlled (0)	No significant displacements (0)	< 500 (0)
Some small areas of surgence downstream but abutments in good condition (3)	Small cracks and settlements undergoing corrective measures (2)	500 (2)
Areas of surgence downstream, slopes and abutments lacking proper corrective measures (6)	Small cracks and settlements lacking proper corrective measures (5)	1000 (5)
Areas of surgence downstream, with increasing flow and material (10)	Cracks, settlements and local instabilities (10)	10000 (10)
EC = Σ (e to **g)**		

The scales in Table 3 are obviously nominal scales. It is indifferent to identify seepage control as "perfect" or to assign the symbol "0" to it. For this reason, the summation presented in the last line of the table has no meaning whatsoever. But supposing, just for the sake of the argument, that the numbers that appear in the cells of Table 3 would have been arrived at by correctly engineered evaluations, the summation would still be completely wrong: Probability Theory (e.g. Benjamin, J.R.; Cornell, C.A., 1970) teaches us that the probability of a joint event is the **product** (not sum) of the individual probabilities, whenever the events can be assumed to be independent from each other (which is not necessarily true for some of the failure modes).

In the original source, however, Table 3 is presented as a table of "risk" classification. As previously discussed, those cell contents cannot be called risks for at least two reasons: their scale is just nominal, and consequences are not taken into account. As far as the latter, Table 4 presents an attempt at classification of at least part of the information that is relevant for the evaluation of the consequences of failure.

Once again, and for similar reasons, the summation presented in the last line of Table 4 is meaningless.

Table 4 naturally implies 4^5 categories (or classes), so that a number between one and 1024 can be assigned to technically classify a given dam. If two dams fall in the same class, they may be considered as "equal" from a technical point of view. When they fall in different classes, however, Table 4 is of no help for deciding which one poses the higher risk.

It is also possible to use just the cell positions to create a 5-digit code number (with a fixed digit position for each property) to identify each technical class. Code 23442, for example, would identify a dam with height between 15m and 30m, crest length between 200m and 600m, design flow lower than 500, upstream construction and monitoring instruments being installed. Neither this classification nor the approach based on the numeric symbols assigned to cells of Table 4 (the summation formula in particular) would support any decision regarding the relative risks of class 23442 versus, for example, class 32341.

4. Sorting a group of dams according to risk

The need to rank a group of dams according the risks they pose is obviously desirable.

Despite having been often and extensively attempted, for the aforementioned conceptual reasons this objective cannot be correctly achieved by means of classification matrices such as tables 3 and 4, let alone by their summaries of summation points.

Again Ackoff (1953) warns that:

> *"We must be careful not to impute automatically to numbers obtained by any process of assigning numbers to objects, events, or properties, the properties which these numbers have as numbers. We can add the numbers of two houses or of two car registrations, but the question is whether or not the sum has any meaning, and if so what."*

Table 4 - Classification matrix with part of the information that is relevant for the evaluation of the consequences of failure (symbol-based nominal scale with arbitrarily chosen description/classification and "corresponding" digits).

Design criteria and maintenance				
Height(a)	**Length (b)**	**Design flow (c)**	**Construction procedure (d)**	**Monitoring (e)**
Height ≤ 15m (0)	Length ≤ 50m (0)	PMF 10000 (0)	Single stage (0)	Monitoring instruments installed according to design (0)
15m < Height < 30m (1)	50m < Length < 200m (1)	1000 (2)	Downstream (2)	Monitoring instruments in the process of being installed (2)
30m ≤ Height ≤ 60m (4)	200 ≤ Length ≤ 600m (2)	500 (5)	Centerline (5)	Monitoring instruments do not follow the design (6)
Height > 60m (7)	Length > 600m (3)	<500 (10)	Upstream (10)	No monitoring instruments (8)
CT = Σ (a to **e)**				

The desired result (ranking a group of dams according the risks they pose) can only be achieved by means of an analysis such as that leading to Figure 3b, where the y-axis represents risks associated with each dam (blue columns). For that purpose, risks must be computed according to the proper engineering definition (Figure 5). Two activities are therefore required:

a. Engineering analysis for the quantitative elicitation of probabilities of failure of dams, usually complemented by extensive historical research in order to generate results which include and extend those in Figure 7;
b. Preview and evaluate failure scenarios and their consequences, in order to generate quantitative results which include an extend those in Figure 8.

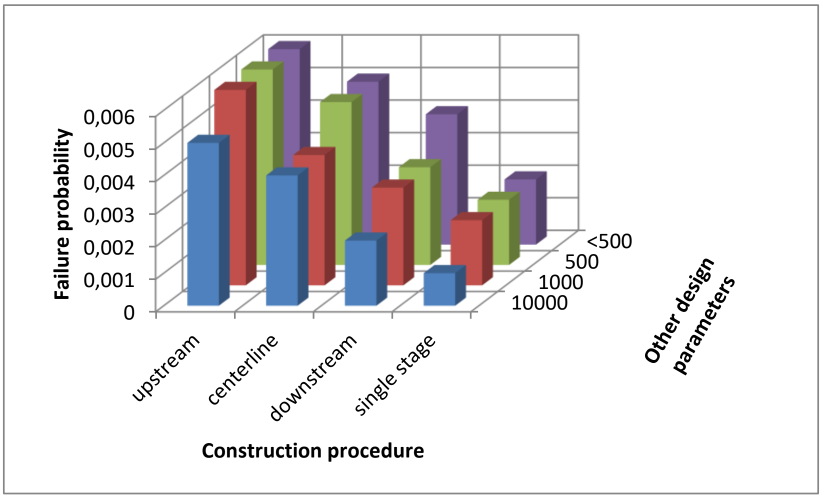

Figure 7 – Example of quantitative results obtained from the elicitation of probabilities of failure of dams.

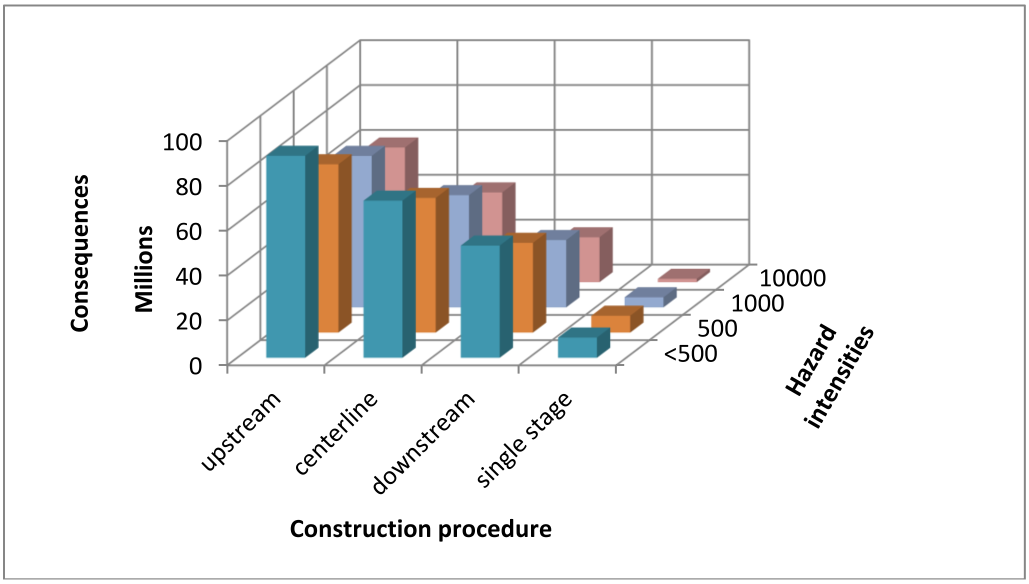

Figure 8 – Example of quantitative results obtained from the evaluation of failure scenarios and their consequences.

5. Conclusions

Decision Analysis and Utility Theory (Pratt, Raiffa, Schlaiffer, 1965) provide a sound theoretical basis for the definition, evaluation and ranking of risks. Results of risks measured in a ratio scale also conform with Measurement Theory.

None of above holds for “risk” classification matrices, which usually ignore or violate well established theoretical principles. Consequently, there is no place for such arbitrary matrices in serious safety and reliability studies.

6. Acknowledgments

The author is deeply indebted and grateful to Bruno Szpigel Dzialoszynski, Luiz Guilherme de Mello and Werner Bilfinger. Their support and help have been essential to the very existence of this paper.

7. References

Ackoff, R.L. Scientific Method: optimizing applied research decisions. New York, 1962, John Wiley & Sons.

Ackoff, R.L. The Design of Social Research. Chicago, 1953, University of Chicago Press.

Benjamin, J.R.; Cornell, C.A. Probability, Statistics and Decision for Civil Engineers, 1970, McGraw-Hill.

Hachich, W. Risk assessment: some views on current practice as evidenced by papers presented to conferences and symposia. Proceedings of the Fourth International Congress on Environmental Geotechnics, Rio de Janeiro, Brazil, August 11-15, 2002, Balkema.

Keeney, R. L. and Raiffa, H. Decisions with Multiple Objectives. 1976, John Wiley & Sons.

Oboni, F. Geo-environmental risk: assessment, analysis, management and planning. Simpósio Brasileiro de Geotecnia Ambiental, São Paulo, Brazil, September 29-30, 1998, ABMS, (CD-ROM).

Pratt, J.W., Raiffa, H.; Schlaiffer, R.O. Introduction to Statistical Decision Theory. New York, 1965, McGraw-Hill.

Raiffa, H. Decision Analysis. 1968, Addison-Wesley, Reading, Mass.