

# Describing probability: The limitations of natural language

Dr David A. Hillson PMP FAPM FIRM, Director, Risk Doctor & Partners

## Introduction

Risks are prioritised by assessing two dimensions: probability of occurrence, and impact on objectives. Impact can be simply defined and assessed in terms of measurable effects on objectives. However assessment of probability is more problematic. This is partly due to the subjective nature of the language used to describe probability, as well as inherent difficulties in dealing with uncertainty.

Previous studies have attempted to determine meanings associated with particular common probability-related terms, but the research methodology differed and results were inconsistent and inconclusive. A research exercise was therefore undertaken to confirm and update the previous findings, and clarify their conclusions.

This paper presents the current findings and discusses how varying perceptions of probability are reflected in the language used. Recommendations are made for the description and assessment of probability in the risk management process, to avoid the subjectivity of natural language.

## Dimensions of risk

There is broad consensus over the definition of “risk” among leading national and international standards and guidelines, as well as professional bodies (Hillson 2002; Hillson 2003, 26-31; Raz & Hillson 2005). Although the precise wording of different definitions may vary, all agree that risk has two dimensions (Williams 1996). The first relates to *uncertainty*, since a risk is something which has not yet happened and which may or may not occur. The second is about what would happen were the risk to occur, since risks are defined in terms of their *effect on objectives*. A typical two-dimensional definition of risk in the realm of project management is “An uncertain event or condition that, if it occurs, has a positive or a negative impact on at least one project objective” (Project Management Institute 2004, 238).

It is common to use the terms “probability” and “impact” to describe these two dimensions, with “probability” addressing how likely the risk event or condition is to occur (the uncertainty dimension), and “impact” detailing the extent of what would happen if the risk materialised (the effect dimension). In order for assessments of risk to be consistent and meaningful, attention must be paid to the way in which probability and impact are assessed. Assessing risk impact merely requires defining the situation after the risk has occurred, and then estimating the possible effect on each objective. The other dimension of risk is less amenable to assessment however. Risk practitioners and project teams alike experience repeated difficulty in trying to assess the probability that a given risk might occur (Tricker 2003; Hillson 2003, 109-114). There are a number of reasons for this, which have been addressed elsewhere (Hillson & Hulett 2004), and are summarised below.

- *Terminology*. “Probability” has a precise statistical meaning, for example “a measure of the relative frequency or likelihood of occurrence of an event, whose values lie between zero (impossibility) and one (certainty), derived from a theoretical distribution or from observations” (Collins 1979). However its general usage is less clear, including its use within the risk management process. Confusion can also arise as a result of the use of alternative terms in risk guidelines to describe the uncertainty dimension, such as “frequency”, “likelihood” or “chance”, giving the impression that these are mere synonyms for “probability” when in fact they are distinctly different. If the uncertainty dimension of risks is to be properly assessed and described using the term “probability”, assessors must understand what they are trying to assess.
- *Unique nature of projects*. A project can be defined as “a temporary endeavour undertaken to create a unique product, service or result” (Project Management Institute 2004, 5), or as “a unique process, consisting of a set of coordinated and controlled activities with start and finish dates, undertaken to achieve an objective conforming to specific requirements, including the constraints of time, cost and resources” (British Standards Institute 2000). All projects are *unique* in some respect, and for those elements there is no body of relevant previous experience on which to draw. This is particularly true of project risks, which affect project objectives. Since the objectives of a given project are likely to be different from those of previous projects, the risks affecting a new project are also likely to be different. This means that some (many?) risks on a particular project will be unique to that project, and there will be no relevant data on their probability of occurrence.

- *Non-availability of “risk actuals”*. Of course some risks on a given project will have arisen previously, since not all aspects of every project are completely unique. However even for these risks, data are often not available from previous projects due to the weakness of the project closure process in many organisations. Without an effective “lessons to be learned” process, each new project has to face its challenges without access to the structured experience of past projects. Post-project reviews should identify and capture risk-related lessons to feed forward to future projects. Without such “risk actuals” from previous projects, the task of assessing the probability of risks which recur on a later project is made more difficult.
- *Unknowable risks*. Sometimes risks are identified for which some details are inherently unknowable, including for some risks the probability of occurrence. This can arise where occurrence of the risk is dependent on influences outside the project (such as the decisions and actions of other stakeholders or competitors), or where the project team lack the necessary knowledge to understand and assess the risk, or in the case of uncertain events which are in the realm of pure chance.
- *Estimating vs measuring*. A further problem with assessing risk probability is that risks are possible future events that have not yet occurred, and as such their probability of occurrence cannot be *measured* but can only be *estimated*. However such estimates tend to be influenced by a wide range of subjective and unconscious sources of estimating bias, making them unreliable. Such sources of bias need to be understood and managed if realistic and useful assessments of probability are to be made.

### Defining probability

The previous section describes the problem – risk probability is inherently difficult to assess or estimate, particularly in the project context. One approach to this is to use a range of alternative techniques for probability assessment, as described elsewhere (Hillson & Hulett 2004). Whichever assessment technique is used however, the primary problem remains one of definition. What exactly is meant by probability? Many approaches to assessing risk probability attempt to describe it using natural language terms. However this does not tackle the underlying issue of semantics, when the same word or phrase can be interpreted subjectively and have different meanings to different people.

This has been recognised and addressed by several previous authors, who have explored the range of probability values associated with commonly used natural language phrases. Authors reporting such studies include Lichtenstein & Newman (1967), Moore (1987, 34-35), Boehm (1989, 133), Hamm (1991), and Conrow (2003, 491-513). Each study reached the same conclusion, namely that phrases which might be thought to be unambiguous in fact could support very large variability of interpretation. A comparative review by Theil (2002) of mean ratings from other studies supported this finding.

Unfortunately the research methodology used by these authors differed in significant ways (see Exhibit 1). One study (Moore 1987) took ten terms and asked respondents to rank order them. Boehm (1989), Lichtenstein & Newman (1967) and Conrow (2003) presented survey respondents with a range of terms relating to probability, and asked them to select a corresponding numerical percentage value or range, with Conrow only allowing respondents to choose from a fixed set of values (5/10/15...90/95%). Hamm (1991) proposed an empirically-derived scale relating numerical percentage values to verbal expressions.

Study	Number of terms	Number of responses	Data presentation
Moore 1987	10	250	Rank order list
Boehm 1989	16	20-25	Scatter plot
Hamm 1991	19	n/a	Single value
Lichtenstein & Newman 1967	40	180 per term	Mean + range
Conrow 2003	49	151	Median + quartiles

**Exhibit 1: Previous research studies**

In each of the previous reported studies, the list of probability-related terms was different, and their reported results are inconsistent and inconclusive, making them impossible to compare. Exhibit 2 presents a subset of results from the previous studies (excluding Moore) for those terms which were common to several studies.

Term	Boehm 1989	Hamm 1991	Lichtenstein & Newman 1967	Conrow 2003
	Mode (range)	Proposed value	Mean (range)	Median (quartiles)
Almost certain	90 (85-99)	95	-	95 (85-95)
Highly probable	-	90	89 (60-99)	85 (75-85)
Quite likely	-	80	79 (30-99)	-
Good chance	75 (60-90)	75	74 (25-96)	-
Likely	60 (55-85)	-	72 (25-99)	75 (65-85)
Probable	70 (55-85)	-	71 (1-99)	75 (65-85)
Better than even	52 (50-60)	60	58 (45-89)	65 (55-65)
Possible	-	-	37 (1-99)	55 (35-65)
Unlikely	30 (15-45)	-	18 (1-45)	15 (15-25)
Improbable	20 (15-45)	-	12 (1-40)	15 (15-25)
Seldom	-	15	16 (1-47)	-
Highly unlikely	5 (0-15)	10	9 (1-50)	15 (5-15)
Rare	-	5	7 (1-30)	-

**Exhibit 2: Previous studies on probability-related terms (values as percentages)**

Given the importance of clear definition of probability, and the lack of coherence between previous studies, it was decided to undertake an independent research exercise, to validate and update the previous findings.

### Research method

An email survey was sent on 13 January 2004 to over 5000 people interested in risk management (members of the Risk Doctor Network), seeking their perceived definitions for selected probability-related terms. Personal data was not sought for reasons of data protection, so it is not possible to determine the demographic profile of respondents, but members of the Risk Doctor Network are largely English-speaking and resident in UK, Europe or North America.

Fifteen terms were offered to respondents, taking the thirteen terms which appeared in more than one of the previous studies (as in Exhibit 2), with the addition of “definite” and “impossible”. These latter two were added as control endpoints, since in theory they do not represent any form of uncertainty: one might expect “definite” to be interpreted as 100% probability, and “impossible” as 0%. The fifteen terms included in this survey were therefore :

A good chance; Almost certain; Better than even; Definite; Highly probable; Highly unlikely; Impossible; Improbable; Likely; Possible; Probable; Quite likely; Rare; Seldom; Unlikely.

Candidate phrases were presented in alphabetical order on a protected Excel spreadsheet with checks to validate data entry. Instructions to respondents read: *“For each of the phrases below, please indicate the range of probability values (minimum to maximum) you associate with that phrase, using the scale from 0-100%. For each phrase, write in the minimum & maximum values. Note that there are no ‘right answers’! Your perception is as valid as everyone else’s. Thank you for taking part in this research exercise. On completion please save this file and email it to [research@risk-doctor.com](mailto:research@risk-doctor.com).”* Incomplete responses were excluded from the analysis. A total of 507 usable responses were received (giving a response rate of about 10%).

An earlier paper presented preliminary findings from this research (Hillson & Hulett 2004, Exhibit 1), but the statistical analysis was limited and open to misinterpretation. The results and discussion in this paper supersede the earlier presentation of results and should be used in preference.

### Results and discussion

The basic analysis for the 507 responses is presented in Exhibit 3. Respondents were asked to state the minimum and maximum probability values which they associated with a given term. For each term in the survey, Exhibit 3 presents the “mean minimum” and “mean maximum” values, as well as the “all data mean”, in decreasing order. These results are also presented in graphical form in Exhibit 4.

Descriptive phrase	Mean minimum	All data mean	Mean maximum
Definite	76.6	<b>80.1</b>	83.5
Almost certain	73.6	<b>78.9</b>	84.1
Highly probable	64.2	<b>71.4</b>	78.3
A good chance	54.3	<b>64.5</b>	74.3
Likely	49.9	<b>59.3</b>	68.4
Quite likely	51.8	<b>59.2</b>	66.4
Probable	47.5	<b>57.2</b>	66.7
Better than even	47.1	<b>56.5</b>	65.6
Possible	28.8	<b>43.3</b>	57.5
Improbable	10.6	<b>18.0</b>	25.3
Highly unlikely	9.8	<b>16.6</b>	23.3
Unlikely	6.6	<b>13.6</b>	20.4
Seldom	6.2	<b>11.7</b>	17.1
Impossible	5.5	<b>8.5</b>	11.4
Rare	3.9	<b>8.1</b>	12.2

Exhibit 3: Research results for probability-related terms (values as percentages)

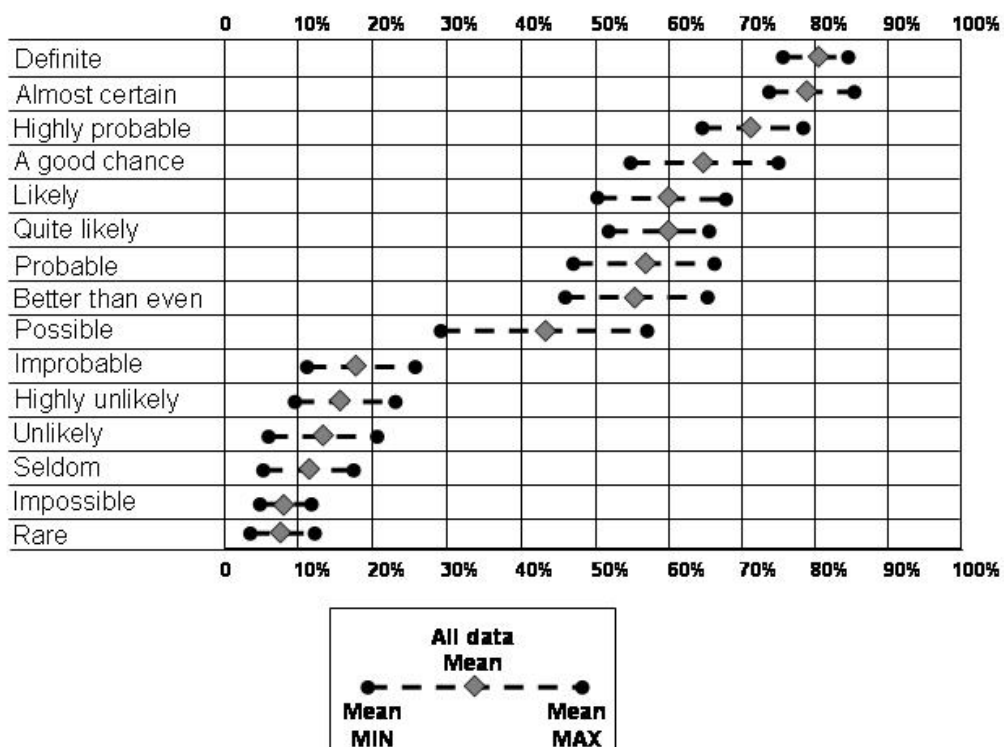


Exhibit 4: Research results for probability-related terms (min/mean/max)

One aim of this exercise was to validate the results of previous studies in defining the meanings ascribed to probability-related terms. Exhibit 5 compares the results from this research (Exhibit 3) with the previous studies (Exhibit 2). Good agreement is shown for terms at the lower end of the probability scale (<50%), as well as for “better than even”, with many of the mean values reported here matching previous results quite closely. However there is much less consistency for terms corresponding to higher probability values (>50%), where mean values from this research are generally lower than comparable values from the previous studies, by 10-20%. In Exhibit 5, results from this research which confirm previous findings are shown in **bold**, and contrary results are in *italics*.

It is possible that a higher degree of validity should be attached to the current results than to previous studies, since the number of respondents (n=507) was much higher than in the earlier work (see Exhibit 1). Also in many cases the spread between mean values of minimum and maximum from this research exercise is smaller than the ranges found in previous studies, and is closest to the results reported by Conrow.

Term	Hillson 2005	Boehm 1989	Hamm 1991	Lichtenstein & Newman 1967	Conrow 2003
	Mean (min/max)	Mode (range)	Proposed value	Mean (range)	Median (quartiles)
Almost certain	79 (74-84)	90 (85-99)	95	-	95 (85-95)
Highly probable	71 (64-78)	-	90	89 (60-99)	85 (75-85)
Good chance	65 (54-74)	75 (60-90)	75	74 (25-96)	-
Likely	59 (50-68)	60 (55-85)	-	72 (25-99)	75 (65-85)
Quite likely	59 (52-66)	-	80	79 (30-99)	-
Probable	57 (48-67)	70 (55-85)	-	71 (1-99)	75 (65-85)
Better than even	<b>57 (47-66)</b>	52 (50-60)	60	58 (45-89)	65 (55-65)
Possible	<b>43 (29-58)</b>	-	-	37 (1-99)	55 (35-65)
Improbable	<b>18 (11-25)</b>	20 (15-45)	-	12 (1-40)	15 (15-25)
Highly unlikely	<b>17 (10-23)</b>	5 (0-15)	10	9 (1-50)	15 (5-15)
Unlikely	<b>14 (7-20)</b>	30 (15-45)	-	18 (1-45)	15 (15-25)
Seldom	<b>12 (6-17)</b>	-	15	16 (1-47)	-
Rare	<b>8 (4-12)</b>	-	5	7 (1-30)	-

**Exhibit 5: Comparing current research with previous studies (values as percentages)**

In addition to the terms in Exhibit 5, this study also included two words which should not be ambiguous in terms of uncertainty, namely “*definite*” and “*impossible*”, expecting these to be interpreted as 100% probability and zero respectively. However ranges of values were also reported for these, and the mean results suggested that even these terms were interpreted as uncertain by most respondents (“*definite*” corresponded to a mean of 80% probability, and “*impossible*” was 8%). Indeed something “*rare*” appeared to be considered slightly less likely than “*impossible*”! This unexpected result raises questions about data quality which cannot be discounted.

Several other interesting patterns emerge from the results. Exhibit 4 appears to show two distinct groups, with six terms having mean values in the range 5-20%, and eight covering 55-80%. Within these two groups, there are closely related subgroups that appear to be synonyms, namely “*seldom/unlikely*” (12-14%), “*improbable/highly unlikely*” (17-18%), and “*better than even/probable/quite likely/likely*” (57-59%). Only one term has a mean reported value in the range 25-50%, namely “*possible*” (mean 43%), and this has a wide min/max spread (from 29-58%). While these results give risk practitioners a useful range of alternative terms when discussing risks with low probability (<20%) and high probability (>55%), there appears to be a lack of useful terminology for risks with probability in the mid-range.

The survey contained two sets of related terms which appear to present a continuum of values, namely “*improbable/probable/highly probable*” and “*highly unlikely/unlikely/quite likely/likely*”. The results for these two sets were inconsistent, with the “probable” set showing regular progression, but with no obvious pattern for terms based on “likely”. Thus :

<i>improbable</i> = 18% (11-25%)	<i>highly unlikely</i> = 17% (10-23%)
<i>probable</i> = 57% (47-67%)	<i>unlikely</i> = 14% (7-20%)
<i>highly probable</i> = 71% (64-78%)	<i>quite likely</i> = 59% (52-66%)
	<i>likely</i> = 59% (50-68%)

These two groups of terms also indicate that adverbs do not have modifying effects in every case. For example there is no significant difference between “*likely*” (min/mean/max 50/59/68 %) and “*quite likely*” (52/59/66 %), nor between “*unlikely*” (7/14/20 %) and “*highly unlikely*” (10/17/23 %). However “*probable*” (47/57/67 %) and “*highly probable*” (64/71/78 %) do appear to be interpreted differently. The adverb “*quite*” may be a special case, since in UK English it can sometimes act as a negative modifier (“Another earthquake is only ‘quite’ likely”), whereas in other contexts it can intensify (“Traffic jams in London are really ‘quite’ likely”); Collins (1979) confirms that “*quite*” can mean either “to the greatest extent, completely” or “partially, somewhat”.

## Conclusions and recommendations

Risk is defined in two dimensions: the uncertainty dimension (assessed as probability of occurrence), and the effect dimension (assessed as impact on objectives). Proper assessment of risks requires appropriate assessment of both probability and impact. The effect on objectives is relatively simple to estimate, as it involves a simple exercise in imagining the situation where the risk happens. Assessing probability of occurrence is less straightforward, but is critical to the effectiveness of the risk process.

Despite the evident importance of the probability dimension, subjective natural language terms are often used to describe probability, leading to a number of problems. However, even where the words are apparently well defined and in common use, individuals translate probability-related terms into percentage values or ranges unreliably, with a range of possible meanings, and with no consensus (as reported previously by Theil, 2002). For example, when someone says a risk is “*unlikely* to occur”, this can be interpreted to mean anything from around 5% probability through to a 20% chance of happening. One person might expect a “*likely*” risk to occur with 50% probability, while another might take this to mean almost 70%.

Three possible remedies can be proposed for this situation. The first is to create and promulgate a standard set of probability-related terms, to be used by all risk practitioners. Secondly, natural language terms can be replaced with ordinal phrases such as “high/medium/low”, though these would need to be defined to avoid ambiguity. Lastly, descriptive terms can be avoided altogether in favour of numerical definitions of probability, using either numbers or ranges. A combination of the latter two options is the preferred solution recommended by several risk standards (including Association for Project Management 2004; and Project Management Institute 2004), and an example is shown in Exhibit 6.

Scale	Range
Very high	>50%
High	26-50%
Medium	11-25%
Low	5-10%
Very low	<5%

**Exhibit 6: Example set of probability definitions**

The results of the research exercise reported here do however provide useful information to the risk practitioner wishing to enhance the risk management process and improve both understanding and assessment of probability. For example, when eliciting probability during a risk interview, terms may be used which are known to be close synonyms to validate previous responses (“This risk *seldom* occurs” – “Would you describe it as *unlikely*?”). Conversely, terms known to have widely different interpretations can be used to challenge assumptions (“You said this risk was *improbable*; do you mean you think there is a *good chance* that it won’t happen?”).

This research has succeeded in validating and extending work previously reported (Lichtenstein & Newman 1967; Moore 1987; Boehm 1989; Hamm 1991; Conrow 2003; Theil 2002), and offers support to risk practitioners in a crucial part of the risk process, namely assessment of the probability of risks occurring. The dangers of relying on subjective terms have been highlighted, and alternative approaches have been outlined. If natural language is used to describe probability, it is *highly probable* that errors will be introduced into the assessment of risk. Replacing these terms with defined ranges makes such errors *unlikely* (though not *impossible*)!

### Author Biography

**Dr David Hillson PMP FAPM FIRM MCMI** is an international risk management consultant, and Director of Risk Doctor & Partners ([www.risk-doctor.com](http://www.risk-doctor.com)). He is a frequent and popular conference speaker and author on risk. David is recognised internationally as a leading thinker and practitioner in the risk field, and has made several innovative contributions to improving risk management. He is well known for promoting the inclusion of proactive opportunity management within the risk process, and has recently been working on applying emotional literacy to risk psychology.

David is an active member of the global Project Management Institute (PMI) and was a founder member of its Risk Management Specific Interest Group. He received the 2002 *PMI Distinguished Contribution Award* for his work in developing risk management over many years. He is a Fellow of both the UK Association for Project Management (APM) and the UK Institute of Risk Management (IRM), as well as being a member of the Chartered Management Institute.

## References

- Association for Project Management (2004) *Project Risk Analysis & Management (PRAM) Guide* (second edition) High Wycombe, Buckinghamshire, UK: APM Publishing
- Boehm, B. W. (1989) *Software Risk Management*. Piscataway, New Jersey, USA: IEEE Computer Society Press
- British Standards Institute (2000) *British Standard BS6079-2:2000 Project Management – Part 2: Vocabulary*. London, UK: British Standards Institute
- Collins (1979). *Collins Dictionary of the English Language*. Glasgow, UK: William Collins Sons & Co Ltd.
- Conrow, E. H. (2003) *Effective Risk Management: Some Keys to Success* (second edition), Reston VA, USA: American Institute of Aeronautics and Astronautics
- Hamm, R. M. (1991) Selection of verbal probabilities: A solution for some problems of verbal probability expression. *Organisational Behaviour & Human Decision Processes*, 48, 193-223
- Hillson D. A. (2002) What is risk? Towards a common definition. *InfoRM, journal of the UK Institute of Risk Management*, April 2002, 11-12
- Hillson, D. A. (2003) *Effective opportunity management for projects: Exploiting positive risk*. New York, USA: Marcel Dekker
- Hillson, D. A. & Hulett D. T. (2004) Assessing risk probability: Alternative approaches. *Proceedings of PMI Global Congress 2004 EMEA, Prague, Czech Republic*.
- Lichtenstein, S. & Newman, J. R. (1967) Empirical scaling of common verbal phrases associated with numerical probabilities. *Psychonomic Science*, 9 (10), 563-564
- Moore, P. G. (1983) *The business of risk*. Cambridge UK: Cambridge University Press
- Project Management Institute. (2004) *A Guide to the Project Management Body of Knowledge (PMBOK®)* (Third Edition). Newtown Square, PA, USA: Project Management Institute
- Raz T. & Hillson D. A. (2005) A comparative review of risk management standards. *Risk Management: An International Journal* (forthcoming)
- Theil, M. (2002) The role of translations of verbal into numerical probability expressions in risk management: A meta-analysis. *J Risk Research*, 5 (2), 177-186
- Tricker L. (2003) Probability – The good, the bad and simple blind faith. *InfoRM, Journal of the UK Institute of Risk Management*, November 2003, 9
- Williams T. M. (1996) The two-dimensionality of project risk. *Int J Project Management*, 14 (3), 185-186