Confidence and risk

Stuart G Reid
University of Sydney
School of Civil Engineering, NSW 2006, Australia

Abstract

The concepts of confidence and risk are discussed with regard to the epistemological (philosophical) basis of risk assessment, and the uncertainty associated with the estimation and quantification of risks. Noting divergent views of the nature of knowledge and science, it is concluded that epistemological issues cannot be objectively resolved or quarantined to protect the 'objectivity' of a risk assessment process, and risk communicators should take account of epistemological issues that affect the way communicated results are interpreted and also affect public trust and confidence. In relation to the treatment of uncertain probabilities of failure, a new measure of 'probabilistic confidence' is proposed to account for the concern that something may be unsafe (in relation to a target level of safety) and may subsequently fail. This measure of 'probabilistic confidence' provides a rational basis for the determination of an appropriate characteristic value for an uncertain probability of failure.

1. Introduction

Risk acceptance and risk communication are critically dependent on the confidence placed in risk assessment results and the trust placed in the specialist communities responsible for the assessment and communication of risks. In the field of quantitative risk assessment, it is often presumed that a scientific approach involving the quantitative modelling of risks will ensure confidence and trust, due to the connotations of objectivity and authority associated with scientific investigations. However, in the context of social policy making, the application of scientific methods and quantitative modelling in areas that extend beyond the natural sciences has been characterised as 'naïve positivism' (where positivism is the philosophical doctrine that maintains that scientific knowledge is the only valid form of knowledge), and there is a general consensus amongst social commentators that 'scientistically' inclined policy makers often fail to recognise the importance of divergent value systems in risk assessment (Stirling, 1998).

A 'scientistic' approach is an approach that makes use of scientific notions, images and methods, not only for the purposes of scientific enquiry, but also for the purpose of invoking the credibility, prestige and authority of 'scientific' knowledge to support an argument or promote a point of view. According to Cameron and Edge (1979): "The concept of scientism implies an attitude to science: those who use scientistic language acknowledge and respect the authority of the scientific community, and wish to capitalise on that authority, in order to make their discourse more persuasive. … However the term is not always (indeed is seldom) used in this purely descriptive or analytic way. Those who espy scientism in others are normally using the concept as a form of criticism. Often scientism is defined as the *illegitimate* use of images drawn from science to add *inappropriate* weight to arguments in which such implicit appeals to scientific authority *should have no place*. … To adopt this evaluative usage of the notion of scientism is to object that the scientific community is imbued with too much authority. … And so we find the common complaint that too much weight and value are attached to the concepts, methods and results of the natural sciences: that, for instance, in trying to make the exercise of political judgement ‘more scientific’ by the use of quantitative techniques purporting to maximise human and social benefits, an aspect of science is being inappropriately carried over into a domain in which its application is illicit. … Alternatively, it is claimed that scientistic thought relies on a mistaken view of the nature of scientific method and practice. The common element in all these formulations is the charge that an excessive respect for the success, prestige and authority of science results in a dangerous misconception of its scope and validity.”

The concept of scientism raises questions concerning the fundamental nature of scientific knowledge and the validity and appropriateness of scientific methods and results. The related question which is addressed in this paper is: What degree of confidence should be placed in a risk estimate or a risk-based decision process?

This is essentially a philosophical question that falls in the domain of epistemology, which is the branch of philosophy that is concerned with the nature, methods, limitations and validity of knowledge and belief. Some general epistemological considerations that influence the effectiveness of risk communication and the degree (or lack) of confidence engendered by the use of probabilistic risk assessments are briefly discussed below.

The paper also discusses confidence in relation to uncertainty in risk estimates, regardless of their epistemological basis. In particular, attention is focussed on safety assessments of structures, for which risk is usually assessed with regard to a Bayesian measure of the failure probability, accounting for all relevant uncertainties. In accordance with Bayesian probability concepts, the total uncertainty is fully characterised by the expected value of the failure probability (which characterises the total uncertainty in the sense of an overall degree of belief), and uncertainty concerning the ‘true’ value of the probability requires no further consideration.

However, in practical applications of risk-based decision-making, the uncertainty or variability of probabilities is often considered to be important. There is a perception that uncertainty in probability estimates adds to the uncertainty represented by the expected probabilities. Furthermore, uncertainties in probability estimates can undermine confidence in probabilistic risk assessments.

To assess uncertainties in Bayesian probability estimates, uncertainties arising from different sources can be treated separately to produce probability distributions of the estimated probabilities. In view of the uncertainty characterised by probability distributions of estimated probabilities, an important question that arises in relation to structural safety is: What is the relationship between an estimated (uncertain) probability of failure and the level of confidence that a structure is ‘safe’?

To account for uncertainty in probability estimates and to promote confidence in the results of probabilistic risk assessments, conservative (safe) estimates of probabilities are sometimes used instead of expected values. However, an objective basis for this approach has not been established and the choice of an appropriate confidence level for the estimation of probabilities (i.e. the level of confidence that the ‘true’ probability is no worse than the estimated value) is essentially subjective. Furthermore, the use of a particular ‘characteristic’ probability does not provide a consistent level of confidence in derived results. To overcome these shortcomings, a different approach is required to characterise the effects of uncertainty in probability estimates.

Alternative treatments of uncertainties arising from different sources are illustrated in the paper with regard to the use of prototype test results for risk-based structural design. The sampling uncertainty associated with prototype test results can be included with the other uncertainties to obtain an estimate of the total (Bayesian) probability of failure. However, different representations can be obtained by treating the sampling variability separately and assessing the statistical confidence associated with the reliability estimates or assessing the probability distribution of the estimated probabilities of failure. In order to characterise the corresponding level of confidence that a structure is safe, a new measure of 'probabilistic confidence' is proposed in this paper and used to determine the appropriate characteristic value of the estimated probabilities of failure to be compared with the specified target value (or target reliability), considering three different prototype testing procedures.

2. Epistemology and confidence

In the field of quantitative risk assessment, it is not uncommon to differentiate between different types of risk and to distinguish between epistemological uncertainty (associated with the limitations of knowledge and understanding) and random variability (associated with natural, intrinsic variability). According to this approach, the epistemological uncertainties involved in quantitative risk analysis can be identified and separated from the other uncertainties, if required, for the purposes of quantitative risk assessment.

However, from a more fundamental philosophical perspective, epistemological issues extend far beyond the identification and treatment of a separable class of uncertainties. Indeed epistemological issues affect every aspect of quantitative risk assessment, even affecting the treatment and interpretation of risks associated with apparent randomness.

Epistemology affects every aspect of any analysis because epistemology is concerned with the nature of knowledge in general, and what it means to know. The term 'epistemology' is derived from the Greek 'episteme' (knowledge) and 'logos' (account or explanation) and was introduced into English by the Scottish philosopher James Frederick Ferrier (1808-1864). However, the origins of epistemology date back to ancient times, and an interpretation of knowledge as justified true belief is sometimes attributed to the dialogue between Socrates and Theaetetus in Plato's *Theatatus* (ca. 360 BC), although a review of the dialogue reveals something a little different:

> **Socrates:** What, then, shall we say of adding reason or explanation to right opinion? If the meaning is, that we should form an opinion of the way in which something differs from another thing, the proposal is ridiculous.
>
> **Theaetetus:** How so?
>
> **Socrates:** We are supposed to acquire a right opinion of the differences which distinguish one thing from another when we have already a right opinion of them, and so we go round and round:-the revolution of the scytal, or pestle, or any other rotatory machine, in the same circles, is as nothing compared with such a requirement; and we may be truly described as the blind directing the blind; for to add those things which we already have, in order that we may learn what we already think, is like a soul utterly benighted.
>
> **Theaetetus:** Tell me; what were you going to say just now, when you asked the question?
>
> **Socrates:** If, my boy, the argument, in speaking of adding the definition, had used the word to "know," and not merely "have an opinion" of the difference, this which is the most promising of all the definitions of knowledge would have come to a pretty end, for to know is surely to acquire knowledge.
>
> **Theaetetus:** True.
>
> **Socrates:** And so, when the question is asked, What is knowledge? this fair argument will answer

"Right opinion with knowledge,"-knowledge, that is, of difference, for this, as the said argument maintains, is adding the definition.

Theaetetus: That seems to be true.

Socrates: But how utterly foolish, when we are asking what is knowledge, that the reply should only be, right opinion with knowledge of difference or of anything! And so, Theaetetus, knowledge is neither sensation nor true opinion, nor yet definition and explanation accompanying and added to true opinion?

Theaetetus: I suppose not.

Socrates: And are you still in labour and travail, my dear friend, or have you brought all that you have to say about knowledge to the birth?

Theaetetus: I am sure, Socrates, that you have elicited from me a good deal more than ever was in me.

Socrates: And does not my art show that you have brought forth wind, and that the offspring of your brain are not worth bringing up?

Nevertheless, the view of knowledge as justified true belief has been long held and is still widely accepted (in a general sense), although divergent views of the concepts of justification and truth have emerged, particularly since the second half of the twentieth century. The divergent views of knowledge can be associated with three general types of epistemology: objectivism, constructionism and subjectivism (Crotty, 1998).

According to the epistemology of objectivism, knowledge is concerned with objective truths that exist externally and independently of the observer. In accordance with the epistemology of constructionism, knowledge is concerned with views that are constructed in the mind, through interaction with a real external world. In accordance with the epistemology of subjectivism, knowledge is concerned with beliefs that are constructed in the mind, independent of any external reality.

Crotty (1998) goes further in specifying the distinguishing characteristics of objectivism, constructionism and subjectivism. According to Crotty, in the epistemology of objectivism, things exist as meaningful entities, independent of human consciousness and experience, and therefore things have truth and meaning residing in them as objects (i.e., objective truth and meaning). In accordance with constructionism, truth and meaning exist only in the conscious mind (of a subject) although they are constructed through engagement and interaction with things (objects) that exist in the real world, and therefore truth is not objective - it can only be constructed and not discovered. Subjectivism differs from constructionism in that meaning is imposed on an object by a subject, without interaction between the subject and object, and therefore meaning is created out of nothing (whilst for constructionism meaning is created out of something - the object).

It is commonly presumed that engineers and physical scientists generally operate (knowingly or unknowingly) in accordance with the epistemology of objectivism. The epistemology of objectivism fits naturally with the ontology (the understanding of what is) of realism (the ontological notion that realities exist outside the mind). Furthermore, the ontology of realism and the epistemology of objectivism are commonly associated with the theoretical perspective of positivism which asserts that objective truth and knowledge can be discovered through scientific study of the real world (independent of the observer).

The philosophical movement of positivism was established in the nineteenth century by the French sociologist Comte who believed that human thought had evolved through religious and metaphysical stages to reach the scientific (or positive) stage, which he attested was the ultimate evolutionary stage of human thought. Comte believed positivism was the final stage in societal and human development, where science and rational explanation for scientific phenomena could provide the basis for continuing social progress. Indeed, Brazil's national motto *Ordem e Progresso* (Order and Progress) was taken from Comte's positivism.

Following Comte, a more critical (less positive) approach to science has developed, and these days positivism is commonly depicted as a naive view that all true knowledge is scientific and that all things are ultimately measurable. The positivist view is sometimes referred to as a scientist ideology and is also associated with technocrats who believe in progress through technological development and science (to the exclusion of subjective human values). This ideological label may also be applied to decision analysts (including risk analysts) who seek to use scientific methods to eliminate subjectivity from societal decision making.

Although many engineers and scientist carry out their work with little regard for philosophical considerations, much has been written about the philosophy of science, including the well known work of Popper (1959) and Kuhn (1970), and to a lesser extent the philosophy of engineering (including Blockley, 1980). Popper is best known for repudiating the classical description of the scientific method that was presumed to be based on careful scientific observation and inductive reasoning to arrive at scientific explanations. Instead Popper proposed that the distinguishing feature of the scientific method (i.e., the feature that distinguished true science from pseudo-science) was the criterion of empirical falsifiability, which requires that any scientific conclusion must be based on empirical evidence that could have shown if the conclusion was false. Kuhn is best known for introducing the concept of "paradigm shifts" to explain major advances in science associated with abrupt and distinct scientific revolutions. Kuhn argued that scientific observations are interpreted in accordance with prevailing "paradigms' (or "world views") and acceptance or rejection of any paradigm is not simply based on logic, but is also dependent on social interactions within the scientific community. In relation to the philosophy of engineering, Blockley discussed the use of imprecise information and imprecise theoretical models for the purposes of design, and examined the implications for engineering safety with regard to the limitations of scientific knowledge (including references to the work of Popper and Kuhn).

In view of the above, it is clear that the underlying philosophies of science and engineering involve much more than a realist/objectivist/positivist stance. The ontology of realism (referring to an external reality) is generally accepted for the physical sciences. However the concept of scientific realism is often extended to include the view that scientific knowledge is a true representation of the real world, and this is less widely accepted. Similarly the objectivist view that the true state of the real world can be known by an observer (independent of the prevailing paradigm) is not widely accepted. Also most scientists and engineers would disagree with the positivist view that there is nothing more to life than science and engineering.

Therefore the epistemology of constructionism may appear to provide a more appropriate basis for science and engineering. Indeed in engineering it is widely acknowledged that we use mathematical models that are constructed to provide an approximate representation of real behaviour. However a constructionist view of science is not widely accepted amongst scientists, because if a consistently constructionist stance is adopted, all understandings - scientific and non-scientific alike – are put on the same footing. They are all constructions: none is objective or absolute, and there is no objective basis to argue that a scientific construction is better than a non-scientific one. To avoid these complications, the scientific community prefers to invoke the alternative concept of instrumentalism (in which perceptions, and scientific models do not necessarily provide an accurate representation of the real world, but are useful instruments to explain and predict real behaviour) or the concept of pragmatism (dating back to the sceptics in classical antiquity who denied the possibility of achieving authentic knowledge regarding the real truth and taught that we must make do with plausible information adequate to the needs of practice).

Also, it is important to recognise that although the scientific approach may be well-suited to the modelling of physical systems that are real and independent of consciousness (in accordance with the ontology of realism) the scientific approach is not well suited to the modelling of abstract things that exist only in a conscious mind. In the absence of an observable external reality, it is not possible to test and ‘verify’ results in accordance with Popper’s principle of empirical falsifiability, and therefore the results cannot be scientific. Attempts to develop scientific models to describe things other than physical systems may result in models that have the general appearance of scientific models, but without the authority that comes from testing in accordance with the scientific method, such models are not truly scientific – they are merely “scientistic” (Cameron and Edge, 1979).

Therefore it must be acknowledged that work in the field of risk assessment cannot be truly scientific because it includes not only the modelling of physical systems, but also the assessment of probabilities and risks which are abstract concepts that exist only in a conscious mind. Furthermore, decision analysis inevitably involves choices that depend on value judgements, and attempts to include such judgements in mathematical models will inevitably produce “scientistic” results.

If scientistic models are used, then it is necessary to address concerns about the confidence that can be placed in the results obtained from such modelling. Confidence cannot be generated by simply asserting that a scientistic model should be treated as if it were as authoritative as a truly scientific model.

The concept of confidence is complex and multi-facetted. Confidence is related to trust, and trust in risk communication can only be achieved on the basis of complete honesty and full disclosure of all information that might be relevant to the assessment of risks, with regard to the interests and concerns of all interested parties. The discovery of any misrepresentations or omissions will result in a loss of confidence in the entire risk assessment process and will erode the authority of all risk assessment results.

These issues have been addressed by workers in the field of social epistemology which is the study of the social dimensions of knowledge. Although the term social epistemology was first introduced in 1970, an antecedent can again be found in one of Plato’s dialogues, *Charmides*, in which Plato posed the question of how a layperson can determine whether or not someone who claims to be an expert can be believed. Dependence on experts or authorities is a topic that falls within the scope of social epistemology, and there is a considerable volume of work that highlights the problems of incompatible epistemologies (e.g. epistemic relativism versus objectivism) and other work (notably from members of the Edinburgh School) that attempts to debunk or undermine the authority of science. Clearly it is relatively easy to identify problems, but there seem to be no easy answers.

Also it should be noted that there is another branch of epistemology that challenges the authority of science and technology: the feminist epistemology. Feminist epistemology challenges ways in which dominant conceptions and practices of knowledge attribution, acquisition and justification systematically disadvantage women and other subordinated groups. According to feminist epistemology, science and technology may be viewed as instruments of established elites that are used to perpetuate the established androcentric structures of power and authority. Feminist epistemology also finds fault with the objectivity and detachment of traditional science, and seeks to incorporate feminine values consistent with feminine cognitive styles (intuitive, holistic, emotional etc).

Clearly, there are many epistemological considerations that can affect the way that the results of a risk assessment will be interpreted. The aim of this discussion is not to establish which epistemologies and interpretations are right or wrong, but to demonstrate the range of views that should be expected, and taken into account, when engaging in the process of risk communication.

Adherents of any of these views will have preconceived ideas about the risks to be considered in a probabilistic risk assessment, and a risk assessment report will only change their own assessments of the risks if that report addresses their concerns in a way that is transparent and acceptable to them. All interested parties, regardless of their epistemic proclivities, will update their own assessments of the risks based on informal Bayesian updating procedures involving not only the communicated results of a risk assessment, but also their own 'prior' assessments and subjective 'likelihood functions' that reflect their level of confidence in the reported results. Unless there is a high level of confidence in the reported results, a preconceived view will not change. Moreover, if a report confirms the presence of an anticipated (perceived) bias, informal Bayesian updating will simply serve to reinforce the preconceived view.

3. Prototype testing

Confidence is also affected by the degree of uncertainty associated with risk assessment results, including probability estimates. The uncertainty associated with a probability estimate can be represented by a probability distribution of the estimated probability. The concept of a probability of a probability is reasonably straightforward, but the corresponding effects on confidence are not obvious.

The effects of uncertainty in probability estimates are discussed below with regard to structural design procedures based on prototype testing. These procedures yield estimates of design strengths, each of which corresponds to a nominal probability of failure for the structure. A reliability-based procedure for design based on prototype testing can be calibrated to achieve a target level of safety (target probability of failure) for the overall population of structures designed in accordance with that procedure, but sampling variability will result in a different nominal probability of failure for each structure.

The uncertainty in the probability of failure for particular structures is described below for structures designed in accordance with a fully-probabilistic procedure (in which the sampling uncertainty is included in the reliability model) and a statistical procedure (in which the sampling uncertainty is taken into account with regard to the concept of statistical confidence).

The paper then introduces a new measure of 'probabilistic confidence' to characterise the safety of the structures designed using the various prototype testing procedures.

3.1. AISI procedure for prototype testing

The AISI specification for cold-formed steel structures (AISI 1990) includes reliability-based load testing provisions for determining structural performance where strength calculations cannot be made in accordance with the AISI specification. To determine structural performance without theoretical strength calculations, load tests must be carried out on at least four identical specimens.

The design load carrying capacity of the tested elements R_D is given by the product of the sample mean capacity R_p and a capacity reduction factor ϕ:

$$R_D = \phi R_p \tag{1}$$

$$\phi = 1.5 M_m F_m \exp(-\beta_o V_o) \tag{2}$$

where M_m and F_m denote the mean values of a material factor and a fabrication factor, respectively; β_o is the target reliability index; and V_o is the coefficient of variation of the safety margin, given by:

$$V_o = (V_M^2 + V_F^2 + C_P V_P^2 + V_Q^2)^{0.5} \tag{3}$$

where V_M, V_F, V_P and V_Q denote the coefficients of variation of the material factor, the fabrication factor, the sample strengths and the relevant load effect, respectively; and C_P is a 'sample variance correction factor' dependent on the sample size n (i.e. the number of test results):

$$C_P = (n-1)/(n-3) \tag{4}$$

The expression for the resistance factor ϕ (Eq. 2) is obtained from reliability calibration studies based on lognormal distributions of the random variables, including approximations that apply when the coefficients of variation are small.

The sample variance correction factor C_p (Eq. 4) is based on an assumption that strengths are log-Student T distributed and the factor C_p is set equal to the variance of the standardised Student T distribution with m=(n-1) degrees of freedom (Pekoz and Hall, 1988).

3.2. Australian Standard procedure for prototype testing

Several Australian Standards include prototype test provisions based on proof load tests of one or more prototype test specimens. The proof load is normally specified as the product of the ultimate design load and a proof load factor that is given as a function of the estimated strength variability (coefficient of variation) and the size of the test sample (possibly as small as one).

The proof load factor is defined as the factor $k(p,c,n)$ that relates the minimum sample strength $R_{min,n}$ (from a sample of size n), to the characteristic strength estimate $R_{p,c}$ (for a particular characteristic strength fractile p and statistical confidence c that the true value of the characteristic strength R_p is not less than the estimated value $R_{p,c}$).

$$R_{p,c} = R_{\min,n} / k(p,c,n) \tag{5}$$

Based on a result obtained for a Weibull distribution of strengths, the proof load factor $k(p,c,n)$ is taken to be (Leicester, 1987):

$$k(p,c,n) = \{\ln(1-c)/[n\ln(1-p)]\}^{V_R} \tag{6}$$

where V_R is the coefficient of variation of the resistance.

For design purposes, the estimated characteristic strength $R_{p,c}$ is multiplied by a capacity reduction factor ϕ_p to obtain the effective design strength R_d. For reliability based design, the capacity reduction factor ϕ_p is calibrated with regard to the relevant target reliability index β_o. Hence, the effective design strength R_d is related to the minimum sample strength $R_{min,n}$ by means of an effective reliability-based Proof Strength Factor (*PSF*) such that:

$$R_d = R_{\min,n} / PSF \quad (7)$$

where the effective Proof Strength Factor (*PSF*) is:

$$PSF = k(p,c,n) / \phi_p \quad (8)$$

For Australian Standards, the proof load factor *k(p,c,n)* has typically been evaluated for p = 0.05 and c = 0.9, and the coefficient of variation of the resistance has been assigned assumed values dependent on the structural type and mode of failure.

For the purposes of this paper, Proof Strength Factors have been used for a Weibull distribution of strengths with a coefficient of variation of 0.224 (relevant to shear strengths of reinforced concrete culverts). Factors were determined for confirmation of the characteristic strength (p=0.05) with alternative statistical confidence levels c of 90% and 50% (for a Lognormal distribution of loads with a coefficient of variation of 0.3, a target safety index β_o of 3.0, and a capacity reduction factor ϕ_p related to the theoretical design point).

3.3. Simulated design strengths based on prototype testing

To illustrate the influence of sampling variability, prototype test results have been simulated for a population of structures with Weibull-distributed strengths, with a fixed mean and coefficient of variation of 0.224 (consistent with the Proof Strength Factors described above). Considering samples of size n=5, design strengths were obtained by simulation in accordance with the AISI procedure and the Australian Standard procedure (considering c equal to 90% and 50%).

3.4. Distribution of the probability of failure for simulated design strengths

The distribution of the variable probability of failure p_f associated with the variable design strengths was determined for each prototype test procedure, and the cumulative distribution functions of the nominal probabilities of failure are shown in Figure 1. It is evident that the distributions of the nominal probability of failure p_f are very skewed and have relatively long upper tails. The mean value of the probability of failure was 0.0054 for the AISI procedure, and for the Australian Standard procedure the mean values were 0.00042 and 0.0017, for c of 90% and 50% respectively (noting that the target probability of failure was 0.00135).

3.5. Confidence in the estimation of the probability of failure

The Australian Standard procedure for prototype testing is intended to provide a high level of statistical confidence (c= 90%) that the design strength will not be less than the target value. However, the results shown in Figure 1 indicate that the use of c=90% results in a very (excessively?) conservative design process. On the other hand the use of c=50% yields an expected probability of failure close to the target value, but one cannot be confident that any particular design will achieve the target safety level.

Results obtained from the AISI procedure are shown to be both unconservative (with respect to the mean probability of failure) and also relatively uncertain (with a very long upper tail). These results are not dependent on the selection of a level of statistical confidence in the design process.

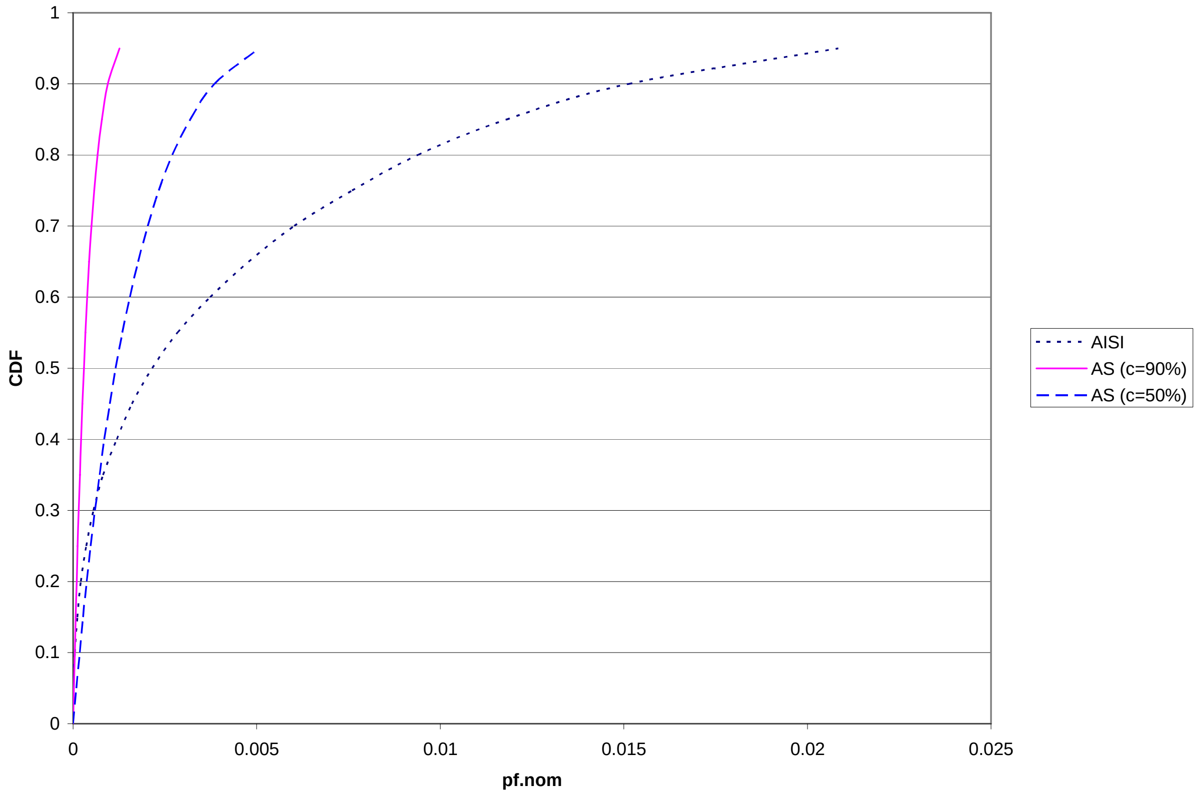

Figure 1: Cumulative probability distribution functions for the nominal probability of failure $p_{f,nom}$ for designs based on prototype testing in accordance with the AISI procedure and the Australian Standard (AS) procedure with c of 90% and 50%.

Although there is no objective basis for selecting an appropriate level of statistical confidence, it is clear that the dependability of the safety assessment and the corresponding confidence in the result is enhanced by considering a conservative 'characteristic value' of the estimated probability of failure, corresponding to a high 'characteristic fractile' of the estimated distribution of the failure probability. However, because the distribution of p_f is highly skewed and has a relatively long upper tail, the value of a 'characteristic probability of failure' would be very sensitive to the choice of an arbitrary 'characteristic fractile', and a conservative 'characteristic probability of failure' would not be a good indicator of the overall level of safety.

Nevertheless, if a characteristic fractile of the distribution of p_f were to be selected to provide confidence in the estimated value of p_f, it would be desireable to select a fractile that was significant with regard to some measure of overall uncertainty. The relative significance of all the fractiles of the distribution of p_f has previously been evaluated with regard to density functions of the relative contributions to the expected probability of failure for designs based on prototype testing (Reid, 2004). It was shown that the contributions to the expected probability (and expected entropy) are dominated by the high fractiles of the distributions of p_f. Accordingly, the expected values are relatively sensitive to the probabilities of failure associated with high fractiles of the distribution of p_f and therefore considerable 'confidence' might be placed in the use of the expected values (which correspond to the 67th percentile and the 63rd percentile for the AISI and Australian Standard results, respectively).

However, the significance of the statistical confidence attached to any particular characteristic fractile of a distribution of p_f is unclear. In addition to the arbitrariness of the selection of a characteristic fractile, the specification of a characteristic fractile corresponding to a statistical confidence of the order of 90%, is not consistent with a target reliability of the order of 99.9% (yet the overall effect has been shown to produce a very conservative basis for design).

In view of the above, there is a clear need for a measure of confidence that could be applied consistently and could characterise the required level of confidence implicit in the specification of a target reliability for a structure with a well-defined probability of failure. Such a measure of ‘probabilistic confidence’ is proposed in the following.

4. A probabilistic measure of confidence for uncertain probability distributions

Statistical measures of confidence are based on the probability that a random value will fall outside some limits. For example, a characteristic fractile of a distribution of p_f is associated with a corresponding level of statistical confidence that p_f will not exceed that value. Although this is the accepted meaning of ‘confidence’ in the context of statistics, in the context of risk-based decision-making ‘confidence’ has a different meaning and is based on a different concept.

In the context of risk-based decision-making, confidence is related to the concern that something might go wrong and produce undesireable consequences. In relation to a design procedure based on prototype testing, the concern is not simply that sampling variability could lead to a design that is less safe than it should be (in relation to a target value of the probability of failure p_{ft}), but rather the concern is that sampling variability could lead to a design that could be unsafe (with $p_f > p_{ft}$) **and** could subsequently fail. Therefore, it is proposed that the probability that a structure is unsafe (with $p_f > p_{ft}$) **and** will fail should be used as a measure of lack of confidence in the safety of the structure. Thus a probabilistic measure of ‘lack of confidence’ that a structure is safe with regard to a target probability of failure p_{ft} can be determined from the probability distribution of the probability of failure $f_{p_f}(p)$, giving:

$$C' = \int_{p_{ft}}^{1} (f_{p_f}(p) \cdot p) dp \tag{9}$$

A complementary measure of ‘probabilistic confidence’ C that a structure is safe (in relation to p_{ft}) is then given by:

$$C = 1 - C' = 1 - \int_{p_{ft}}^{1} (f_{p_f}(p) \cdot p) dp \tag{10}$$

For a structure with a well-defined probability of failure (i.e., a probability of failure that can be estimated with negligible uncertainty), the distribution of the probability of failure may be approximated by a uniform distribution confined to a small interval centred on the expected value. Then if the estimated probability of failure coincides with the target probability p_{ft}, the ‘probabilistic confidence’ that the structure is ‘safe’ (in relation to p_{ft}) is:

$$C = 1 - C' \approx 1 - \int_{p_{ft}}^{1} (f_{p_f}(p) \cdot p_{ft}) dp = 1 - p_{ft}/2 \tag{11}$$

This level of confidence is implicitly required in the specification of p_{ft} as an acceptably safe value for a well-defined probability of failure, and it is proposed that this same level of confidence should also be required when dealing with uncertain probabilities of failure. Hence, for any distribution of p_f, not only should the expected value not exceed the target value p_{ft}, but also the probabilistic confidence C should be not less than (1-p_{ft}/2). (In fact, the requirement for probabilistic confidence will always govern.)

For any distribution of p_f, a characteristic value p_{fc} can be determined such that the required probabilistic confidence is achieved if p_{fc} is not greater than the target probability of failure p_{ft}. This characteristic fractile of the distribution of p_f is given by:

$$\int_{p_{fc}}^{1} (f_{p_f}(p) \cdot p)dp = p_{fc}/2 \tag{12}$$

The characteristic values p_{fc} of the uncertain distributions of p_f obtained from the various prototype testing procedures are given in Table 1, together with the expected values of the probability of failure $E[p_f]$, the fractiles $F_{pf}(E[p_f])$ corresponding to the expected values, and the fractiles $F_{pf}(p_{fc})$ corresponding to the characteristic values p_{fc}. The results given in Table 1 show that the characteristic values of the probabilities of failure p_{fc} that should be compared with the target value p_{ft} are significantly greater than the expected values of p_f. Similarly, Table I shows that the expected values of p_f obtained from the AISI and AS methods of prototype testing (for the examples considered) correspond to the 67th and 63rd percentiles, respectively, whilst the characteristic failure probabilities p_{fc} correspond to the 76th and 72nd percentiles of the distributions of p_f obtained from prototype testing.

Table 1: Statistics of the uncertain distributions of p_f obtained from prototype testing.

	$E[p_f]$	$F_{pf}(E[p_f])$	p_{fc}	$F_{pf}(p_{fc})$
AISI	0.0054	0.67	0.0078	0.76
AS (c=90%)	0.00042	0.63	0.00053	0.72
AS (c=50%)	0.0017	0.63	0.00214	0.72

5. Conclusion

The related concepts of confidence and risk have been discussed with regard to the epistemological basis of quantitative risk assessment. It is concluded that epistemological issues affect every aspect of a risk assessment, and contrary to the expectations of those who claim that the processes of quantitative risk assessment are scientific and objective, epistemological issues cannot be objectively resolved or conveniently quarantined to protect the purported objectivity of a risk assessment process. Furthermore, if risk communicators want to communicate effectively with an audience that includes people from outside the risk assessment community, then they need to understand and take account of epistemological issues that affect the way the communicated results will be interpreted. In particular, risk analysts and risk communicators need to understand the importance of earning the trust and confidence of the general public, taking due account of divergent epistemological stances.

The concepts of confidence and risk have also been discussed with regard to the uncertainty associated with the estimation of risks. In particular, uncertainty in estimated probabilities of failure has been discussed with regard to prototype testing procedures for reliability-based structural design. These procedures yield estimated design strengths, each of which corresponds to a nominal probability of failure that satisfies an overall target level of safety for the population of structures, but sampling variability results in a different nominal probability of failure for each structure. The uncertain nominal probabilities of failure have been described by probability distributions of the nominal probabilities, and an associated measure of ‘probabilistic confidence’ has been proposed to reflect the concern that sampling variability could lead to the design of a structure that could be unsafe (in relation to the target level of safety) and could subsequently fail. The level of ‘probabilistic confidence’ implicitly required in the specification of a target reliability has been evaluated for structures with well-defined probabilities of failure, and it has been proposed that this same level of ‘probabilistic confidence’ should also be required when dealing with uncertain probabilities. The paper describes how the required level of ‘probabilistic confidence’ can be achieved with an uncertain probability of failure by ensuring that an appropriate characteristic value of the uncertain probability (defined in the paper) does not exceed the relevant target probability.

References

AISI (1990) Load and Resistance Factor Design Specification for Cold-Formed Steel Structural Members, Research Report CF 90-1, AISI, Washington, DC, USA.

Blockley, D.I. (1980) The Nature of Structural Design and Safety, Ellis Horwood, Chichester, UK

Cameron, I., and Edge, D. (1979) Scientific Images and their Social Uses: an Introduction to the Concept of Scientism, Butterworths, London, England.

Crotty, M. (1998) Foundations of Social Research: Meaning and Perspective in the Research Process, Allen & Unwin, Australia.

Kuhn, T.S. (1970) The Structure of Scientific Revolutions, University of Chicago Press, Chicago,IL, USA

Leicester, R.H. (1987) “Load factors for proof and prototype testing.” First National Structural Engineering Conference, IEAust, Melbourne, Australia, 546-550.

Pekoz, T., and Hall, W.B. (1988) “Probabilstic evaluation of test results.” Proceedings of the Ninth International Specialty Conference on Cold-Formed Steel structures, University of Missouri-Rolla, Missouri, USA.

Popper, K.R. (1959) The Logic of Scientific Discovery, Routledge, Taylor and Francis Group, London, England

Reid, S.G. (1992) “Acceptable risk.” Engineering Safety, D.I.Blockley (ed), McGraw Hill, 138-166.

Reid, S.G. (1998) “Prototype load testing procedures for structural design.” Structural safety and Reliability (Proceedings of the Seventh International Conference on Structural Safety and Reliability, Kyoto, Japan, 1997), Shiraishi, et al (eds), Balkema, Rotterdam, 573-580.

Reid, S.G. (1999) “Perception and communication of risk and the importance of dependability.” Structural Safety, 21(4), 373-384.

Reid, S.G. ((2004) “Uncertain probability estimates and an entropy-based measure of uncertainty.” Reliability and Optimisation of Structural Systems, M.Maes & L.Huyse (eds), Taylor and Francis, London, 79-86.

Stirling, A. (1998) “Risk at a turning point?” Journal of Risk Research, 1(2), 97-109.