Reliability, Britteness, Covert Understrength Factors, and Fringe Formulas in Concrete Design Codes

Zdeněk P. Bažant, F.ASCE; and Qiang Yu

Abstract: The paper analyzes the reliability consequences of the fact that the current design codes for concrete structure contain covert (or hidden) understrength (or capacity reduction) factors. This prevents distinguishing between different combinations of separate risks due to the statistical scatter of material properties, the error of the design formula, and the degree of brittleness of failure mode, and also makes any prediction of structural reliability (or survival probability) impossible. The covert formula error factor is implied by the fact that the design formula was calibrated to pass not through the mean but through the fringe (or periphery, margin) of the supporting experimental data. The covert material randomness factor is the ratio of the reduced concrete strength required for design to the mean of the strength tests. As a remedy, the covert understrength factor of design formula should be made overt, its coefficient of variation (based on the supporting test data) should be specified, and the type of probability distribution (e.g., Gaussian or Weibull) indicated (which then also implies the probability cutoff). Alternatively, the code could give the mean formula, specify its coefficient of variation and type of distribution, and either prescribe the probability cutoff or overtly declare the understrength factor. The mean of strength tests required for quality control should be figured out from the required design strength on the basis of a specified probability cutoff and the coefficient of variation of these tests. Furthermore, it is proposed that the currently used empirical understrength factor, which accounts mainly for the risks of structural brittleness (or lack of ductility), should be based on the expected maximum kinetic energy that could be imparted to the structure. The reliability integral taking into account the randomness of both the load and structural resistance is generalized for the case of multiple (statistically independent) understrength factors. Finally, it is pointed out that the currently assumed proportionality of the tensile and shear strengths to the square root of compressive strength of concrete is realistic only for the mean, but grossly underestimates the scatter of tensile and shear strengths.

DOI: 10.1061/(ASCE)0733-9445(2006)132:1(3)

CE Database subject headings: Reliability; Britteness; Concrete structure; Design; Codes.

Introduction and Present Status

Improvements in the structure of safety provisions in building design codes offer a much greater potential for advances in safety and economy than improvements in structural analysis. It makes little sense to use, for example, sophisticated finite element analysis to improve the accuracy of structural analysis by 10% if the safety factors that must be applied to the results can be wrong by 50%. However, to realize this potential, the probabilistic considerations behind the safety factors must take into account the new understanding of the failure process, especially with regard to brittleness.

Beginning with the 1971 American Concrete Institute (ACI) concrete design code, two kinds of safety factors have been distinguished: (a) the load factors, and (b) the strength (or capacity) reduction factors, the latter commonly called the understrength factors. In ACI code (ACI 2002), the basic factored load cases are 1.6L+1.2D and 1.4D (D=dead load, L=live load). For very small structures, the live load totally dominates (L>D), and so the load that decides is essentially 1.6L. For very large structures, the self-weight totally dominates (D>L), and so the load that decides is essentially 1.4D.

However, as pointed out by Bažant and Frangopol (2002) and elaborated on by Novák and Bažant (2002), 1.4D is irrational. The actual errors in self-weight of structures (other than sabotage) can hardly justify more than 1.03D, and definitely not more than 1.05D.

Because the self-weight is negligible for small structures and dominates for very large structures, the excessive self-weight factor 1.4 (instead of a rational value =1.05) represents a hidden size effect. Unfortunately, this size effect is applied indiscriminately to all types of failure. This is irrational. For example, ductile failures, such as flexure of underreinforced beams, need no protection from size effect, while the failures in shear, torsion, punching, compression crushing, tension, and bond slip of concrete need an even larger protection in the case of very large structures. As another example, prestressed concrete or high-strength concrete structures receive less protection than un prestressed or normal-strength reinforced concrete structures because they are lighter, while they actually need more protection because they are more brittle. Besides, the shape of the size effect curve (i.e., a curve of the nominal strength of structure versus structure size)
that is generated by imposing an excessive self-weight factor is incorrect (giving too little size effect for medium size structures).

The hidden size effect cannot simply be removed by reducing factor 1.4 to 1.03 or 1.05 because the hidden size effect partially compensates for the current lack of size effect in the code provisions for brittle failures (an exception is the recent introduction of size effect for the pullout of anchors from concrete, in which case the self-weight load factors 1.4 as well as 1.2 should, of course, be reduced).

If, on the other hand, a realistic size effect (i.e., the energetic size effect of quasi-brittle fracture mechanics) was introduced into the code provisions for brittle failures without simultaneously removing the hidden size effect, it would unjustly penalize very large structures and also retain the current unjustified penalization of heavy structures compared to light ones of the same size (e.g., high versus low strength concretes, unprestressed versus prestressed concretes). Therefore, a remedy will require collaboration of statisticians with fracture mechanics experts. This point will not be discussed here any more because it has already been made (Bazant and Frangopol 2002).

However, there are other irrational safety margins hidden in various strength reductions which are implied but not explicitly stated in concrete design codes. The objective of this paper is to clarify them and propose the concepts to remedy them, while leaving the paramount problem of size effect aside for subsequent studies.

Examples of Existing Fringe Formulas in Concrete Design Codes

To figure out the relationship of the design strength (or internal force) required by factored loads to the average internal force measured in failure tests, one must realize [as pointed out in Bažant (2003)] that concrete codes [including ACI (2002)] actually contain two kinds of understrength factors: overt and covert (or hidden). The overt ones are stated in the code explicitly, visible to all users, and are simply called the understrength (or strength reduction) factors. The covert ones [identified in Bažant (2003)] are two: (1) the formula-error factor, \( \psi_f \); and (2) the material randomness factor, \( \psi_m \).

Covert Formula-Error Factor

First consider, as an example, the classical formula for shear of longitudinally reinforced concrete beams without stirrups (or other shear reinforcement). It gives the shear strength (force) contributed by concrete as

\[
v_c = 2 \sqrt{f_c' \psi_f}
\]  
(1)

(better written in a dimensionally proper form as \( v_c = 2 \text{ psi} \sqrt{f_c'} \text{ psi because the formula is valid only in psi} \); \( v_c = V_c / b_{sd} \); \( V_c = \) shear force capacity (or shear strength) provided by concrete (i.e., total shear force \( V \) minus the contribution of steel shear reinforcement); \( b_{sd} = \) web width; and \( d_a = \) depth of beam from its top surface to the longitudinal reinforcement centroid.

Now note the way this currently prescribed code formula (ACI 2002) was calibrated long ago by shear test data. As seen in Fig. 1(a) plotted from ACI (1962), the horizontal line for \( v_c = 2 \sqrt{f_c'} \) was not set to match the optimum (least-squares) fit of the test data, i.e., it was not set to pass through the centroid of the test data cloud in Fig. 1(a). [These small-beam data (194 points) are here plotted directly from a table in that paper. These data also appear in Fig. 7.13 of Park and Paulay (1975), in Fig. 5.5.1 of Wang and Salmon (1998), in Fig. 6.8 of Nawy (2003), and in other subsequent books. But it must be warned that these subsequent plots differ from the original, and also from each other (they were adapted by deletions of many points, and by replacements of groups of points with single points, which produced an increase of the mean above the correct value). Furthermore, one must note that the ordinates in these subsequent plots use \( \sqrt{f_c'} \) for what should be, according to ACI (1962), \( \sqrt{f_c} \) (in current ACI notation, \( f_p = f_p' \)). Eq. (1) was set to fit roughly the lower fringe (or margin) of the test data cloud, which is found to lie at about 65% of the mean (or centroid) of the data cloud, denoted as \( \bar{v}_c \) [see Fig. 1(a)]. Based on the standard deviation of \( v_c \) data and the hypothesis of Gaussian distribution plotted in the figure, this corresponds to a 5% probability cutoff. Hence

\[
\psi_f = \frac{v_c}{\bar{v}_c} = 0.65, \quad p_{cut} = 5\% \quad (2)
\]

This ratio represents a covert understrength factor. It will be called the formula-error factor because it accounts for the scatter in Fig. 1(a) which is caused mainly by the fact that the design formula ignores additional influences such as the size effect, steel ratio, and shear span, and also a host of other poorly understood parameters, such as the concrete mix parameters, aggregate and cement types, etc. [that this is the main cause of high scatter is documented by the fact that, when the main influences are taken into account by a more complex formula, the scatter bandwidth is drastically reduced; see Bažant and Yu (2003, 2005a,b)]. The probability cutoff \( p_{cut} = 5\% \) means that 5% of the data in Fig. 1(a) is expected to be below the line \( 2 \sqrt{f_c} \). No reason for setting the cutoff as 5% is known.

A more accurate design for shear must take into account the size effect, and also the effects of reinforcement ratio \( \rho_s \) and shear span \( d/a \). A formula of this type has recently been proposed and calibrated by Bažant and Yu (2003, 2005a,b); see the solid curve in Fig. 1(b), showing the current database of committee ACI-445 with 398 data points for beams without stirrups [developed as an expansion of an older database of 296 points from Bažant and Kim (1984), and Bažant and Sun (1987)]. The choice of variables in the formula for the solid curve, which reads \( v_c = v_0 (1 + d/a)^{-1/2} \) (in units lb., in., psi), has made it possible to considerably reduce the scatter width (Bažant and Yu 2003, 2005a,b), as seen by comparing this figure with Fig. 1(a), in which \( v_0 = \mu \varphi \rho_{sw}^{1/3} (1 + d/a) \sqrt{\bar{v}_c}, \mu = 13.3, \) and \( d_0 = 3800 \sqrt{\bar{d}_a} / \sqrt{f_c'}^{2/3} \) if \( d_a \) is known. If \( d_a \) is unknown, \( d_0 = 3330 / \sqrt{f_c'}^{2/3} \) (here \( d_{sa} = \) maximum aggregate size in inches, \( \rho_{sw} = \) steel ratio, \( d_{al} = \) shear span). For the mean, we have \( \varphi_f = 1 \), but because the approach of ACI-445 is to pass the design formula at the 5% percentile of data rather than the mean, the solid curve in Fig. 1(b) is reduced by factor \( \varphi_f = 10/13.3 \approx 0.76 \) to the dashed curve corresponding to the 5% probability cutoff (obtained under the hypothesis that the regression errors follow the Gaussian distribution). Note that the improved form of the design formula has made it possible to decrease considerably the scatter bandwidth. Consequently the covert understrength factor, \( \varphi_f = 0.76 \), is much higher than the value 0.65 implied for small-size beams by the current code (which in turn is still too high for large beams designed according to the current code). This documents that factor \( \varphi_f \) takes into
account primarily the error scatter of the mathematical formula. The better the formula, the closer to 1 is its covert understrength factor, i.e., the formula-error factor.

Fig. 1(c) shows a simplification of the formula proposed by Bažant and Yu (2003, 2005a,b). It replaces a smooth curve by two straight lines—the asymptotes of the size effect law, and is defined as

\[ v_c = f_c \left[ 2 - \frac{d}{d_0} \right] \]

If \( f'_c \leq 5,000 \) psi (the first argument governs if \( \delta_f > 505 \) psi, which is usually the case). Here \( \bar{f}_c \) is the average uniaxial compressive strength of concrete measured on 6 in. (151 mm) diameter cylinders (denoted by ACI as \( f'_c \)); and \( \delta_c = \) standard deviation of the compression strength tests of the given concrete. If the strength data for the given concrete are insufficient to assess \( \delta_c \), ACI (2002) code allows taking the difference \( \bar{f}_c - f'_c \) as 1.000 psi for \( f'_c < 3,000 \) psi, as 1.200 psi for \( 3,000 \leq f'_c < 5,000 \) psi, and as 0.1\( f'_c + 700 \) psi for \( f'_c > 5,000 \) psi (1 psi = 6.895 Pa). If the first argument in Eq. (3) governs, the ratio

\[ \varphi_m = \frac{f'_c - \bar{f}_c}{f'_c} = 1 - k_m \delta_f / \bar{f}_c \]

is a covert understrength factor taking into account material randomness, and will therefore be called the material randomness factor. From Article 5.3.2.2 of ACI-318-02 code (ACI 2002), it can be figured out that, approximately

\[ \varphi_m = 0.75 \]
\[ \omega_m = 19\% , \quad \text{Prob}(\text{strength} < \varphi \bar{f}_c) = p_{\text{cut}} \approx 9\% \] (6)

The value of 0.75 coincides with the ACI (2002) code value exactly when \( \bar{f}_c = 4,800 \text{ psi} \). For \( \bar{f}_c = 6,000 \text{ psi} \), the ACI rule gives \( \omega_m = 0.80 \), and for \( \bar{f}_c = 4,200 \text{ psi} \), it gives \( \omega_m = 0.714 \); for \( \bar{f}_c = 3,000 \text{ psi} \), ACI specifies a discontinuity, such that 4,200 psi is reduced to 3,000 psi, and 3,999 psi to 2,999 psi (however, this is a problematic specification, unsuitable for computer analysis, because the discontinuity would cause divergence of design optimization and sensitivity analysis programs). The approximate values \( \omega_m = 19\% \) and \( p_{\text{cut}} \approx 9\% \) represent the coefficient of variation of strength expected by the code and the corresponding approximate cutoff probability, that is, the probability of compression strength being less than \( \varphi \bar{f}_c \). These values were figured out, for \( \bar{f}_c = 4,800 \text{ psi} \), and for a Gaussian distribution of strength, from the fact ACI (2002) implies the 1,200 psi reduction of the cutoff probability, that is, the probability of compression strength being less than \( \varphi \bar{f}_c \) (i.e., less than \( \bar{f}_c \)). These values were figured out, for \( \bar{f}_c = 4,800 \text{ psi} \), and for a Gaussian distribution of strength, from the fact ACI (2002) implies the 1,200 psi reduction of the mean strength to be equivalent to 1.348 \( \delta_f \) where \( \delta_f \) is the standard deviation of compression strength.

As for the cutoff \( p_{\text{cut}} \approx 9\% \) implied by the choice of 1.34, no scientific reason is known, though. It is also unclear why this cutoff should be different from that recently introduced for anchor pullout.

**Overt Understrength Factor in Current Code**

What is the main purpose of the overt understrength factors in the current ACI code, denoted here as \( \varphi \)? It is certainly not to take into account the randomness of strength because this is done by the covert understrength factor \( \omega_m \) (material randomness factor). Neither is it to take into account the formula errors because this is done by the covert understrength factor \( \varphi_f \) (formula-error factor).

What appears to be the main purpose is to take into account the risk of failure and the seriousness of its consequences. Compare the overt understrength factors \( \varphi \) specified by the code for various types of failure: for beam shear failure it now is 0.75 (it was 0.85 until 2002); for column failure, which can also be brittle if concrete is getting crushed (and often puts the entire structure at risk of failure), it is also 0.75; for tensile failures of plain concrete, which is highly brittle, it is 0.55; and for flexural failure of under-reinforced beams, which is nonbrittle, or ductile (and does not put the whole structure at risk), it is 0.9.

**How to Eliminate Deficiencies of Fringe Design Formulas**

There are two potent reasons for eliminating fringe (peripheral, marginal) formulas from the design code: (1) to distinguish properly among the risks of many possible designs and failure modes; and (2) to render statistical reliability analysis of design possible.

**Distinguish Different Risks for Different Designs**

In the allowable stress design, it was impossible to distinguish between load uncertainty and structural resistance uncertainty, and among the different uncertainties for different load combinations and different degrees of brittleness. The purpose of introducing the load factor resistance design (LFRD) four decades ago was to distinguish between these uncertainties.

Now a similar situation arises for the fringe formulas. The random scatter of material strength and the error of the design formula (which is due partly to oversimplification of the mechanism of failure, and partly to incompleteness of its understanding), have different sources and separate effects on design safety. Furthermore, both are different from the effect of the degree of brittleness on the risk of catastrophic failure.

Therefore, separate understrength factors are needed in order to distinguish among these different effects.

**Render Probabilistic Reliability Assessments Possible**

If the type of distribution (e.g., Gaussian, Weibull) is known, the statistical distribution of a random variable is fully characterized by the mean and the coefficient of variation. But if a fringe formula is presented in a design code without specifying:

1. Its probability cutoff (or percentile); and
2. Its coefficient of variation, \( \omega \).

It is impossible to reconstruct the mean and standard deviation (nor the standard error of regression). This in turn makes it impossible to predict the mean structural response and its standard deviation, thus making reliability analysis of the structure impossible.

At reliability conferences, many investigators, unsuspecting of the hidden fringe nature of the existing code formula, simply take that formula as the mean predictor, infer its coefficient of variation intuitively or by some analogy, and proceed straight away to make sophisticated reliability calculations. Unfortunately, such calculations are doomed to be nothing more than a mere mathematical exercise, yielding erroneous and misleading probabilistic estimates of structural reliability.

Based on the current structure of codes, correct estimation of structural reliability would not be easy. The analyst would have to search first for the detailed test data that were originally used by the code-making committee to calibrate its formula, but such data are often hard to find. Then he would have to conduct statistical regression of these data to reconstruct the mean prediction from the fringe formula in the code. Only after that, he could engage in a meaningful reliability analysis of the structure.

To render structural reliability estimates possible, it is inevitable to introduce changes in the design code. For the formula understrength factor, there are two options:

- **Option 1.** Keep the existing fringe design formula unchanged, but make it in the code absolutely clear that the design formula is a fringe estimate. At the same time, specify in each code article the factor \( \varphi_f \) that is necessary to scale up the formula back to the mean of the scatter band of test data, and further indicate in each article the type of probability distribution (e.g., Gaussian or Weibull) and either the probability cutoff associated with \( \varphi_f \) or the coefficient of variation of test data (one of which implies the other).

- **Option 2.** Scale up each current fringe formula so that it would follow the mean of the scatter band of the test data (e.g., the mean beam shear strength or mean anchor strength), and impose a requirement to use additional understrength factor \( \varphi_f \), specifying either its probability cutoff or coefficient of variation.

As for concrete strength tests, their mean value that must be achieved to pass quality control should be determined from the required design strength on the basis of a specified probability cutoff and the coefficient of variation of these tests.

Using tacit about the fringe nature, the code makers unwittingly invite misleading estimates of structural reliability (as well misinterpretation of test results and forensic evidence). As an example, Fig. 1(b) shows the new ACI-445 database for shear failure of beams without stirrups, along with the curves of the
mean prediction and the fringe prediction by a new improved formula proposed for the code by Bazant and Yu (2003, 2005a,b) (this formula, by virtue of taking into account several additional influencing factors, achieves a much reduced coefficient of variation).

Design Criterion Implied by Codes: Apparent and Actual

To sum up, the design criterion according to ACI (2002) appears to have the form

\[ \max(1.6L + 1.2D, 1.4D; \ldots) \leq \varphi F_{\text{red}} \quad (\varphi = 0.75) \]  

where \( L, D = \) internal forces due to live and dead design loads; \( 1.6, 1.2, 1.4; \ldots = \) well-known load factors specified by the code; and \( F_{\text{red}} = \) certain empirically reduced value of the internal force representing the structural resistance with a certain small, but unspecified, probability cutoff. However, based on the foregoing discussion, the actual form of the design criterion implied by the code is

\[ \max(1.6L + 1.2D, 1.4D; \ldots) \leq \varphi_{\text{app}} F_{\text{m}} \quad (\varphi_{\text{app}} = 0.65, \varphi_{\text{m}} = \sqrt{0.75}) \]  

where \( F_{\text{m}} = \) the mean (average) value of the internal force at failure. For reliability analysis, this needs to be translated into failure probability, which we discuss next.

Reliability Analysis with Several Understrength Factors

The objective of reliability analysis is to make all the empirical load factors and understrength factors superfluous by calculating the failure probability \( P_f \) for the given random loads and all random properties of the structures with given statistical distributions. The reliability-based design requires designing the structure so that \( P_f \), under random loads of given statistical properties, would not exceed a given extremely small value such as \( 10^{-7} \) (in other words, if 10 million identical structures were built, not more than one would be expected to fail under these loads).

If we take into account only (1) the randomness of load, corresponding to the load factors; and (2) the randomness of structural resistance, corresponding to a single understrength factor \( \varphi_f \), then the failure probability is, according to Freudenthal et al. (1966), calculated as follows (see also, e.g., Ang and Tang 1984; Madsen et al. 1986; Haldar and Mahadevan 2000)

\[ P_f = \text{Prob}(R < P) = \int_{-\infty}^{\infty} p(\sigma_N) R(\sigma_N) d\sigma_N \]  

where \( P = \) load or loading resultant, \( R = \) resistance of structure in the sense of \( P, \) and \( \sigma_N = c_p P/hD = \) nominal structural strength (where \( D = \) structure size or characteristic dimension, \( b = \) structure thickness or width, \( c_p = \) dimensionless constant chosen for convenience); \( \sigma_N \) represents a load parameter having the dimension of stress [in the example of beam shear in Figs. 1(a–c), \( \sigma_N \) is set to be equal to \( \psi, \) which is the ACI notation for shear strength of concrete]; \( p(\sigma_N) \) is the probability density distribution of load \( P, \) described in terms \( \sigma_N; R(\sigma_N) = \) cumulative probability distribution of structural resistance, which is described also in terms of \( \sigma_N \) and corresponds to the usual ACI understrength factor \( \varphi \) [Fig. 1(e)].

Even though a negative structural resistance makes no sense physically, the lower limit of the integral is written as \( -\infty \) in order to allow the use of Gaussian distribution, which has nonzero (albeit totally negligible) tail values in the negative range (note that if the lower limit were written as 0, the Gaussian density distributions would not be normalized, i.e., their integrals would not be exactly 1). If a distribution with a nonnegative threshold, such as Weibull, is considered, then the lower limit of the integral can be replaced by 0.

Eq. (9) is proven as follows: The probability that a nominal strength is not larger than \( \sigma_N \) is \( R(\sigma_N), \) and the probability that the loading would produce a nominal stress within the infinitesimal interval \( (\sigma_N, \sigma_N + d\sigma_N) \) is \( p(\sigma_N) d\sigma_N. \) So the joint probability of simultaneous occurrence of both is \( R(\sigma_N) [p(\sigma_N) d\sigma_N]. \) Failure occurs if the inequality \( R < P \) is met in any infinitesimal interval. Hence the contributions to failure probability from all these intervals need to be summed, which leads to integral (9).

The deterministic design must ensue as the special case for \( R(\sigma_N) = H(\sigma_N - \mu \sigma_N) \) and \( p(\sigma_N) = \delta(\sigma_N - \varphi_f \varphi_{\text{app}} \sigma_N) \), where \( H \) denotes the Heaviside step function, and \( \delta \) the Dirac delta function; \( \mu = \) load factor (e.g., 1.4 for dead load), \( \sigma_N = \) nominal stress produced by load (as deterministic load parameter), and \( \sigma_N = \) nominal strength characterizing the deterministic structural resistance. Indeed, Eq. (9) in this case yields \( p_f = 0 \) for \( \mu \sigma_N < \varphi_f \varphi_{\text{app}} \sigma_N \) and \( p_f = 1 \) for \( \mu \sigma_N > \varphi_f \varphi_{\text{app}} \sigma_N, \) i.e.,

\[ p_f = H(\mu \sigma_N - \varphi_f \varphi_{\text{app}} \sigma_N) \]  

The type of statistical distribution must be decided by probabilistic fracture analysis. In Bazant (2003), it is argued that \( R(\sigma_N) \) ought to be the Weibull distribution for brittle failures of large concrete structures, while for extrapolation to zero structure size \((D \rightarrow 0), \) \( R(\sigma_N) \) ought to gradually approach the Gaussian distribution. The free parameters of any statistical distribution are its mean and coefficient of variation. This underscores the importance of using formulas that give the mean of the data as well as the coefficient of variation. When only a fringe formula is given, the mean of the distribution is unknown; hence \( R(\sigma_N) \) is indeterminate, the integral in Eq. (9) is in calculable, and thus statistical reliability analysis is impossible.

Eq. (9) is the standard reliability integral. This integral is limited to the case of a single understrength factor \( \varphi \) corresponding to distribution \( R(\sigma_N). \) Taking multiple understrength factors into account, we may write

\[ \sigma_N = \varphi \psi F(\chi, \varphi_f) = f(\psi \psi F(\chi, \varphi_f) \leq \sigma_N) \]  

where \( \varphi_f = \) mean material strength and \( f(\varphi, \psi, \chi) = \) function of three independent random variables; \( \chi \) represents material randomness, assumed to have Gaussian distribution with Weibull tail, \( \psi \) represents the error of the formula (or the theory), which may be assumed to be Gaussian, and \( \varphi = \) structural brittleness factor, corresponding to the understrength factor in the current codes, which may have a Gaussian core with grafted Weibull tail of size-dependent length. The resistance CDF may be expressed as

\[ R(\sigma_N) = \int_D \int_D r(\varphi) r(\psi) r(\chi) F'(\chi, \varphi_f) d\chi d\psi d\varphi \]  

where \( F'(\chi, \varphi_f) = \partial F(\chi, \varphi_f)/\partial \varphi_f \) and \( D \) designates the domain in which \( \psi \psi F(\chi, \varphi_f) \leq \sigma_N, \) \( D = \) complicated domain, hard to specify analytically. Nevertheless, integral (12) may be easily approximated numerically, e.g., by Monte Carlo simulations (in which any random simulation for which \( R > P \) or
The standard deviation, characterizing the central part of the PDF, may be estimated, in view of statistical independence of $\varphi$, $\psi$ and $\chi$, as follows:

$$s^2_n = s^2_\varphi + s^2_\psi + [s_\chi \Phi(f)/\Phi(f)]^2$$

This may be used to anchor the distribution of $R$ if its type is known.

The splitting of failure probability $p_F$ into several probabilities corresponding to the separate understrength factors $\varphi$, $\psi$, and $\chi$ has one benefit—it reduces the length of the far-off tails of the distributions which must be known to calculate $p_F$ (Bažant 2003). Structures must be designed for an extremely small probability of failure such as $10^{-2}$ and $10^{-7}$ (as determined computationally at Northwestern by S. Pang, personal communication, 2003). Failures of such a small probability can hardly be verified by testing. But splitting the understrength probabilities into several random variables $\varphi$, $\psi$, and $\chi$, and assuming their statistical independence, the tails of the individual probability density distributions need to be known only up to about $10^{-2}$, which is within the reach of experimental verification.

**Approximating Brittleness Effect by Reducing Dynamic Strength**

In discussing the (overt) understrength factor $\varphi$, we have accepted the prevailing, though vague, opinion that brittleness (or lack of ductility) causes a reduction of load capacity of softening structures. To put this opinion on a solid foundation, it is necessary to consider the unwanted random disturbances to the structures, which can be either static, such as imperfections, or dynamic, such as the maximum kinetic energy, $K$, that can be randomly imparted to the structure by earthquake, impact, blast, or wind gust.

There are two possible causes of postpeak softening on the load-deflection diagram: (1) either instability (nonlinear geometrical effects), as in buckling of a cylindrical shell under axial compression; or (2) fracturing, as in shear, torsion, punching, or compression crushing of concrete. In this regard, it must be noted that while imperfections have an enormous effect on the static peak load and postpeak response of cylindrical shells [Fig. 2(a)], see Bažant and Cedolin (1991, Chapter 7), they have only a modest effect on the static peak load and postpeak response of fracturing structures [Fig. 2(b)]. Consequently, the static imperfection viewpoint cannot be used to judge the effect of insufficient ductility.

First let us take the dynamic viewpoint and consider the typical load-deflection diagrams of concrete structures, shown in Fig. 3. Fig. 3(a) shows the load-deflection diagram for a structure with unlimited ductility, e.g., for the flexural failure of an underreinforced beam. If a kinetic energy $K$ of any magnitude is imparted to the structure, its peak load would not decrease.

Then consider the load-deflection diagrams with softening shown in Figs. 3(b–d). Assume that the structure is in equilibrium under a constant load at point A. The horizontal line $AB$ = dynamic (nonequilibrium) passage at constant load $P_A$, which occurs if a certain kinetic energy $K$ is imparted to the structure (e.g., by earthquake, impact, blast, or wind gust). The difference in $P$ between the equilibrium curve and the passage $AB$ = inertia force causing deceleration of the structure in the sense of $P$. The area $W_{AB}$ above line AB represents the energy absorbed by the structure during the nonequilibrium load-deflection passage $AB$. If $W_{AB} > K$, the passage $AB$ cannot get completed and thus the structure will not fail under such load $P_A$. If, however, $W_{AB} = K$, the passage $AB$ will get completed and the structure will reach the softening equilibrium state B which is unstable under controlled load (Bažant and Cedolin 1991) and leads to failure.

If the imparted kinetic energy $K$ is fixed, then area $W_{AB}$, and thus also load $P_A$ at which the structure fails, depends on the roundness or sharpness of the peak of the load-deflection diagram. This is clear by looking at Figs. 3(b–d), in which all the

\[
\psi \phi(\chi,f) > \sigma_N
\]

...imparted to the structure by earthquake, impact, blast, or wind gust...

**Fig. 2.** Comparison of imperfection effects in (a) thin shell buckling and (b) quasibrittle fracture. (c and d) Two shapes of load-deflection peaks and load drop based on imparted kinetic energy $K$.

**Fig. 3.** (a–d) Dependence of maximum load and dynamic ductility on peak morphology and on kinetic energy $K$ imparted to the structure. (e–h) Dependence of static ductility $\bar{AB}$ on peak morphology, postpeak, and structural stiffness $C$. 

The dimensionless deflection $\mu = [\text{Deflection}] / [\text{Height}]$...
shaded areas $W_{AB}$ above the passage $AB$ are equal. The sharper the peak, or the steeper the postpeak, the smaller is load $P_A$ at which the structure fails.

If the value (or distribution) of $K$ to be used for comparison of different types of failure is specified (e.g., by some future design code), one could use this simple reasoning for estimating the load capacity decrease caused by brittleness (or lack of ductility). Under dynamic excitation such as earthquake, $K$ cannot exceed the maximum elastic strain energy $U_{el}$ that can be stored in the structure, which equals $P_0^2/2C_γ$ where $P_0 =$ maximum statically applied load [Figs. 3(a and b)] and $C_γ =$ elastic stiffness of the structure. Therefore the imparted kinetic energy $K$ that could be prescribed by the code could be expressed as

$$K = \eta \frac{P_0^2}{2C_γ}$$

Here $\eta$ is some empirical constant not larger than 1 (which could further be generalized to a random variable of a given distribution). In this manner, the reduction of maximum load [or reduction of nominal structural strength, $\Delta K$, Figs. 3(b–d)] approximating the effect of limited ductility could be evaluated precisely. This would make possible a simple but unambiguous comparison of structural designs with different ductility.

It should be noted that the load capacity reduction due to brittleness (or insufficient ductility) comes in addition to the size effect, which is purely static. If the equilibrium load-deflection diagram of a structure exhibits postpeak softening (and if the softening does not have geometric cause such as buckling), then there is always size effect (because distributed fracturing is unstable and must localize into a fracture process zone and a crack, and because the fracture process zone or crack causes stress redistribution before maximum load, with energy release). The size effect, causing a decrease, $\delta\sigma_N$, of nominal structural strength, is illustrated in Figs. 3(c and d).

Note that, in fracture mechanics, brittleness is usually understood as the proximity to LEFM in terms of size effect [see Bažant and Planas (1998) for size effect of type 1; and Bažant and Pang (2005) for types 1, 2, and 3]. This is a more precise measure of brittleness but would be harder to calculate in concrete design.

**Concepts of Static and Dynamic Ductility of Structures**

Since the subject of brittleness, or lack of ductility, inevitably came up, it should be pointed out that stability analysis can be used to formulate an alternative, static, concept of the understrength factor reflecting the differences in brittleness among various failure types. Aside from dynamic ductility, the concept of which was outlined in the preceding section on the basis of Eq. (14), a static assessment of brittleness of a structure or structural member can be made in the manner proposed by Bažant (1976) and further explored by Bažant and Becq-Giraudon (1999) and Bažant et al. (1987) (see also Bažant and Planas 1998; Bažant and Pang 2002). This is illustrated in Fig. 3 too.

Stability is lost when the tangent of the load-deflection diagram has slope $C_γ$ that is equal to $-C_p$, where $-C_p$ represents the stiffness of the rest of the structure, i.e., a structure obtained by extracting the failing member (or softening zone). If the line of slope $-C_p$ is not tangent at any point to the load-deflection diagram [as in Fig. 3(e)], the static ductility is unlimited. In Figs. 3(f–h), the static ductility is finite because a tangent point, B, exists. So, the static ductility may be defined by the ratio of the length $\ell_{AB}$ of horizontal segment AB in Figs. 3(f–h) to the elastic deflection under load $P_A$ at point A, i.e., by

$$\frac{\ell_{AB}}{P_A/C_γ}$$

For the case of insufficient dynamic ductility discussed in the preceding section, simple formulas are also possible if the postpeak load drop $\Delta P$ is not too large. One may discern two possible cases of load-deflection diagrams: (1) a discontinuous change of slope at peak; or (2) a smoothly rounded peak. In the former case, shown in Fig. 2(c)

$$\Delta P = \sqrt{\frac{2K}{C_γ^2 - C_p^2}}$$

and in the latter case, shown in Fig. 2(d)

$$\Delta P = \sqrt{\frac{3}{8} K\kappa}$$

where $C_γ$, $C_p =$ prepeak (elastic) and postpeak (softening) stiffnesses of the structure associated with load $P$; $\kappa = d^2P/du^2$ = curvature of smooth static load-deflection diagram at peak load; and $u =$ load-point deflection of structure.

**True Safety Factor and Its Obscuring Effect on Forensic Evidence**

The overall safety factor $\mu$ is defined as the mean of failure test data divided by the mean (or unfactored) design load. For shear failure of longitudinally reinforced concrete beams without stirrups, the overall safety factor currently is

$$\mu = 1.6/(0.75 \times \sqrt{0.75 \times 0.65}) = 3.8 \quad \text{for small size} \quad (18)$$

$$\mu = 1.4/(0.75 \times \sqrt{0.75 \times 0.65 \times 2.0}) = 1.7 \quad \text{for large size} \quad (19)$$

where the former applies to small beams, totally dominated by the live load, and the latter to large beams, totally dominated by the self-weight. In the latter case, the neglected size effect factor has been considered as 2.0. Factors 1.6 and 1.4 are the load factors; 0.75 is the (over) understrength factor $\varphi$ for shear failure; 0.65 is the covert understrength factor $\varphi_m$ for material strength randomness.

In view of the scatter width seen in Fig. 1(a), the individual safety factors, defined as the ratios of the failure load of individual beams to the unfactored (or mean) design load, vary within the following ranges:

$$\mu_{indiv} = \text{from } 2.3 \text{ to } 6 \quad \text{for small size} \quad (20)$$

$$\mu_{indiv} = \text{from } 1.05 \text{ to } 2.8 \quad \text{for large size} \quad (21)$$

The very large values of these safety factors explain why there have not been many more structural collapses, despite the inadequacy of the design procedure (especially the neglect of size effect). They also reveal that, in concrete engineering (by contrast to aeronautical engineering), one mistake in design or construction is usually not enough to bring the structure down.

The size effect factor can hardly be more than 2, and so the size effect alone does not suffice to cause the collapse of any structure if the material strength and formula error have nearly
mean values. To produce collapse, the material strength and formula error must simultaneously have values of small probability, far from the mean. Thus, at least two, and typically three, simultaneous mistakes or lapses of quality control are needed to make a concrete structure collapse. Such situations might be rare, but they can happen, and doubtlessly will.

For example, in the case of catastrophic sinking of the Sleipner oil platform in a Norwegian fjord in 1991, which was due to shear failure of a thick tricell wall, there were three simultaneous mistakes. The governmental investigating committee recognized (1) an incorrect reinforcement detail in the tricell; and (2) incorrect meshing which caused a simultaneous mistake of about 35% in the shear force at critical location obtained from elastic finite element analysis. But on top of that, there must have been a size effect factor of about 1.4, which was ignored. Although reported by Bažant (on February 14, 1992) to the design firm (Det Norske Veritas), this factor was unfortunately omitted from the governmental report (Jacobsen and Rosendahl 1994). The two aforementioned mistakes seem sufficient to explain collapse only if the covert understrength factors are ignored. If they are not ignored, the third mistake, omission of size effect, becomes necessary to explain collapse.

As documented by the foregoing example, the covert understrength factors, if ignored, have the unfortunate effect of obscuring the forensic evidence after a disaster, especially with respect to the size effect. If the failure can be blamed on one or two simple mistakes other than the neglect of size effect, it is seductive to skip the analysis of fracture propagation and size effect in disaster investigations.

It is conceivable that a number of disasters might not have occurred, despite other mistakes, if the size effect were properly taken into account in design.

How to Judge whether an Experiment Supports the Code

Aside from the original ACI (1962) test data, Fig. 1(a) also shows two data points (Collins and Kuchma 1999; Angelakos et al. 2001; Bentz, personal communication, 2003), obtained in shear tests of rather large beams (of depth d=1.89 m), one without and one with minimum stirrups (as specified by ACI-318-02). The test without stirrups was immediately accepted as a proof that the size effect must be introduced into the code.

On the other hand, the beam with minimum stirrups, otherwise identical, failed at $v_c=1.84\sqrt{f_c}=2.12\sqrt{f'_c}$. This is 6% higher than the required design strength $v_c=2\sqrt{f'_c}$. Based on this fact, many experts in ACI concluded that minimum stirrups suffice to make the design safe without any consideration of size effect. However, a mere look at Fig. 1(a) reveals that such a conclusion is false.

Indeed, when the presence of the covert understrength factors is taken into account, the conclusion is entirely different. Just compare visually the test result for minimum stirrups to the cloud of small beam data in Fig. 1(a), on which the design formula $2\sqrt{f_c}$ was based. If there were no size effect, the test would be expected, on the average, to lie in the middle of the strength distribution shown in the figure. The statistical expectation of a safe value of $v_c$ in the test is not $2\sqrt{f'_c}$ but $2\sqrt{f'_c}/(0.65\sqrt{0.75})$, i.e., $3.55\sqrt{f'_c}$ or $3.1\sqrt{f'_c}$.

So, unless the single test is a chancy rare result lying at the tail of probability distribution (which would be an overoptimistic assumption that would have to be proven by further testing), the correct conclusion is that, even for shear with minimum stirrups, the safety margin is insufficient, much smaller than required for small beams, and that there is a large size effect, albeit not as large as for beams without stirrups. Compared to the average of $v_c=3.55\sqrt{f'_c}$, the reduction of $v_c$ due to size effect on the tested beam with minimum stirrups was about $(3.55-2.12)/3.55$ or 40%, which is not negligible at all.

This real-life example shows that cognizance of the covert (but real and necessary) under-strength factors is essential for interpreting test results.

Incorrect Use of $\sqrt{f'_c}$ in Estimating Scatter in Shear or Tension

The empirical factor $\sqrt{f'_c}$ is in ACI code ubiquitous. Because of approximate proportionality of mean tensile strength $\bar{f}_t$ to $\sqrt{f'_c}$, this factor appears in all the formulas for failures governed by tensile or shear strength of concrete.

Proportionality to $\sqrt{f'_c}$ is often also considered in estimates of statistical scatter. However, this is incorrect. What is overlooked is that the proportionality to $\sqrt{f'_c}$ has been experimentally established only for the mean strength and gives incorrect results if applied to the scatter. Because

$$d\sqrt{f_l}/\sqrt{f'_c} = \frac{1}{2}df/l'c$$

a direct use of factor $\sqrt{f'_c}$ indicates the coefficient of variation of tensile and shear strengths to be one-half of the coefficient of variation of the compression strength. But this is patently untrue, for it is known that the coefficient of variation of tensile and shear strength is actually not less than the coefficient of variation of compression strength of concrete. So, the coefficient of variation of random variable $r_m(\varphi_m)$ in Eq. (4) must not be assumed to be less than the coefficient of variation of compression strength of concrete.

Consequently, the design code should expressly warn that the factor $\sqrt{f'_c}$ in all of its formulas guarding against failures associated with tensile or shear strength of concrete may be used only for calculating the mean load capacity, but not for calculating the standard deviation of load capacity of structure from the standard deviation of compressive strength of concrete.

Conclusions

1. In addition to the explicit understrength (or capacity reduction) factor taking into account mainly the consequences of brittleness of structure, two covert understrength factors are implicit in concrete design codes. One is the formula-error factor, implied by the fact that the design formula (e.g., the formula for beam shear capacity) does not represent the mean fit (least-squares fit) of the data cloud but is made to pass at the lower fringe (margin, periphery) of the data cloud. The other is the material randomness factor implied by designing structures on the basis of the reduced strength of concrete (i.e., the specified compression strength, $f'_c$, lying at the lower fringe of strength data), rather than the mean of material strength data and their variance.
2. One undesirable consequence of the use of fringe formulas in the current codes is that it is impossible to distinguish the different risks of diverse combinations of
   - The degree of brittleness (or lack of ductility) of structural failure;
   - The inevitable random scatter of material strength; and
   - The error of the design formula (stemming from oversimplification and incomplete understanding of the mechanics of failure).

The situation is similar to that which existed prior to the introduction LFRD.

3. Another undesirable consequence is that the use of fringe formulas (with unspecified covert understrength factors and probability cutoffs) renders meaningful probabilistic estimates of structural reliability impossible.

4. To rectify the code, the covert understrength factor of design formula should be overtly stated, its coefficient of variation (based on the test data used by the committee to calibrate the formula) should be specified, and the type of probability distribution (e.g., Gaussian or Weibull) indicated (the type of distribution, coefficient of variation, and the understrength factor then imply the probability cutoff). Alternatively, the code could give the mean formula, specify its coefficient of variation and type of distribution, and either prescribe the probability cutoff or overtly declare the understrength factor. The mean of strength tests required for quality control should be figured out from the required design strength on the basis of a specified probability cutoff and the coefficient of variation of these tests.

5. If all the understrength factors and the associated probability distributions can be assumed to be statistically independent (which seems reasonable), the reliability integral giving the failure probability of structure can be easily generalized.

6. The standard (overt) understrength factor in the current codes reflects mainly the effect of structural brittleness (or lack of ductility). For this purpose, a rigorous definition of ductility is needed. It is proposed to base it on the magnitude of kinetic energy that can be imparted to the structure (by earthquake, impact, blast, or wind gust). This kinetic energy can be assumed to be equal to a certain percentage of the maximum elastic strain energy that can be stored in the structure (which can always be calculated). If this percentage is specified, the reduction of structural strength (i.e., the understrength factor for brittleness) can be described by simple formulas, one for the case of sharp peak with a sudden change of the slope of the load-deflection diagram, and the other for a smoothly rounded peak.

7. The reports evaluating previous disasters should be regarded with reservations because of a possible omission of size effect as a contributing factor. The overall safety factors for brittle failures of concrete structures are so large that several simultaneous mistakes are normally necessary to cause collapse. If the covert understrength factors are disregarded, it is easy to put all the blame on one or two other mistakes, and thus systematically overlook the mistake that does not fit the established thinking.

8. The proportionality of tensile and shear strengths of concrete to \( f_{c}^{*} \), as used in ACI (2002) code, is justified only for the mean. Its use grossly underestimates the statistical scatter, particularly the coefficient of variation of tensile and shear strengths.

Acknowledgments

Financial support by the Infrastructure Technology Institute of Northwestern University under Grant No. 0740-357-A466 is gratefully acknowledged. The theoretical background studies were supported by National Science Foundation Grant No. CMS-0301145 to Northwestern University.

References

American Concrete Institute (ACI). (2002). Building code requirements for structural concrete (ACI 318-02) and commentary (ACI 318R-02), Farmington Hills, Mich.
Fuchs, W., Eligehausen, R., and Breen, J. (1995). "Concrete capacity
safety*, Prentice-Hall, Englewood Cliffs, N.J.
ed., Prentice Hall (Pearson Education), Upper Saddle River, N.J. 
(Fig. 6.6).
Novák, D., and Bažant, Z. P. (2002). "Interaction of size effect and
reliability of design: Case study of flexural strength of concrete." *First
Int. Conf. on Bridge Maintenance, Safety and Management
(IABMAS)*, Universitat Politécnica de Catalunya, Barcelona, 1–8.
New York (Fig. 7.13).
Ed., Addison-Wesley (Longman), Menlo Park, Calif. (Fig. 5.5.1).