Designing Against Size Effect on Shear Strength of Reinforced Concrete Beams Without Stirrups: I. Formulation

Zdeněk P. Bažant and Qiang Yu

Abstract: The shear failure of reinforced concrete beams is a very complex fracture phenomenon for which a purely mathematical approach is not possible at present. However, detailed modeling of the fracture mechanism is not necessary for establishing the general form of the size effect. The first part of this paper shows that the general approximate mathematical form of the size effect law to be calibrated by experimental data can be deduced from two facts: (1) the failure is caused by cohesive (or quasibrittle) fracture propagation; and (2) the maximum load is attained only after large fracture growth (rather than at fracture initiation). Simple dimensional analysis yields the asymptotic properties of size effect, which are characterized by: (1) a constant beam shear strength \( v_c \), (i.e., absence of size effect) for sufficiently small beam depths; and (2) the linear elastic fracture mechanics size effect \( v_c \sim d^{-1/2} \) for very large beam depths \( d \). Together with the recently established small- and large-size second-order asymptotic properties of the cohesive (or fictitious) crack model, this suffices to unambiguously support a size effect formula of the general approximate form \( v_c = v_0 (1 + d/d_0)^{-1/2} \) (where \( v_0, d_0 \) are constants), which was proposed in 1984 for shear failure of beams on the basis of less general and less fundamental arguments. Verification and calibration are left for Part II of this paper which follows.

DOI: 10.1061/(ASCE)0733-9445(2005)131:12(1877)

CE Database subject headings: Size effect; Shear strength; Shear failure; Concrete beams; Concrete, reinforced; Fracture; Design; Safety.

Introduction

Although a provision for size effect in shear failure of reinforced concrete beams was incorporated into some design codes more than a dozen years ago, compelling experimental evidence obtained by properly scaled tests of large-size beams has become available only during the last few years. The case has now become clear—the design formula must include the size effect. The problem is how to best interpret the test results, and how to best describe the size effect mathematically in a sufficiently simple and practical manner, without violating certain restrictions that have crystallized from theoretical research during the last two decades. To clarify this timely question, is the objective of this paper.

The size effect is measured in terms of the nominal strength, generally defined as \( \sigma_N = F/ld \), where \( F \) is the maximum (or ultimate) load (or load parameter), \( b \) is the structure width, and \( d \) is the characteristic dimension (or size) of the structure. The size effect is characterized by comparing \( \sigma_N \) for geometrically similar structures of different sizes \( d \). According to the classical allowable stress design, as well as the theory of limit states (or plastic limit analysis) which underlies the current design codes for reinforced concrete structures, the nominal strength \( \sigma_N \) is independent of the structure size. We say that in this case there is no size effect. It has been generally proven that the size effect is absent from all structural analysis methods in which the material failure at a point of the structure is decided exclusively by the stress and strain tensors at that point, and that a size effect inevitably arises if the material failure criterion involves energy. This is the case of fracture mechanics provided that, at maximum load, the crack, or the fracture process zone (i.e., microcracking zone), or both, is not negligible compared to structure dimensions.

A size effect is exhibited by all the theories of failure which involve some material characteristic length \( l_0 \). This is always the case when the failure criterion involves both the stress and fracture energy because the ratio of energy per unit area to stress has the dimension \((\text{Nm/m}^2)/(\text{N/m}^2) = \text{m})\). An example is the cohesive crack model, as well as the crack band model (which is almost equivalent, Bazant and Planas 1998), and also the nonlocal damage model. The cohesive (or fictitious) crack model, originated by Barenblatt (1959), developed by Rice (1968), and pioneered for concrete by Hillerborg et al. (1976), is an approximation now generally regarded as the best compromise between simplicity and accuracy.

The salient property of these models is that attainment of the strength limit at a material point means only that fracture can
Method of Interpretation of Existing Experimental Database

Why can’t a size effect formula be determined purely empirically? We need to address this question to place the problem in proper perspective. Many formulas in concrete design codes can of course be developed purely empirically because it is possible to obtain adequate test data for the entire range of practical interest, and to conduct experiments that sample this range statistically uniformly, without bias. An example is the expression for the elastic modulus in terms of the compressive strength, or the effect of reinforcement ratio in various code specifications. Unfortunately, the size effect is not a problem of that kind.

Fig. 1 shows the histogram of the number of test data versus beam depth d according to the database compiled by subdivision 445F of the American Concrete Institute (ACI), also called the ESDB database (Reineck et al. 2003). The size effect is of practical concern mainly for beam depths ranging from 1 to 15 m (the depth of the failed Koror Bridge girder in Palau was 14.2 m). Unfortunately, 86% of all the available test data pertain to beam depths less than 0.5 m, and 99% to depths less than 1.1 m. The coefficient of variation ω of the deviations of a size effect formula from the points of the database will therefore be totally dominated by small beam depths for which the size effect is unimportant. Thus it can easily happen that some formula that gives the smallest ω for such data could be completely wrong for very large sizes while another formula that might give a higher ω could be much more realistic for large sizes. So, data fitting alone is not the way to develop a size effect formula for the ACI code. For an unbiased, purely empirical, validation of a formula, the test data would have to be distributed uniformly over the entire range of interest. In view of the costs of large scale tests, we cannot even hope to acquire such a database. Therefore, we must extrapolate. But the extrapolation, visualized by Fig. 1, cannot be accomplished empirically. In keeping with the motto of former ACI president’s inaugural message (Izquierdo-Escaranión 2003), “ars sine scientia nihil est,” a sound scientific support is required. The scientific theory, in turn, should be verified by properly scaled size effect tests on one and the same concrete, and especially by reduced-scale model tests in which the dimensionless size (characterized in a shape-independent manner by the brittleness numbers β, explained in Bažant and Planas 1998; Bažant 2002; Bažant 2004) can attain the highest possible values.

A size effect formula for shear strength, motivated by fracture mechanics, was proposed by Bažant and Kim in 1984. Since then a number of other formulas have appeared; see Fig. 2 which includes:

1. The size effect law based on fracture mechanics and energy release arguments (Bažant 1984; Bažant and Kazemi 1991; Bažant and Planas 1998; Bažant 2002; Bažant 2004);
2. An extended form of that law for fractures in which the cohesive stresses are never reduced to zero but exhibit a finite residual strength (Bažant 1987);
3. A formula resulting from an enhancement of the modified compression field theory based on a hypothesis that crack spacing causes size effect (Vecchio and Collins 1986; Collins et al. 1996; Collins and Kuchma 1999; Angelakos et al. 2001);
4. The CEB-FIP formula, introduced empirically (Comité Euro-International du Béton 1991);
5. Similar Carpenter’s (1994) formula (MFSL);
6. The formula of Japan Concrete Institute (Japan Society of Civil Engineers 1991), motivated by Weibull statistical theory—a theory that assumes the failure to occur right at the initiation of a macroscopic crack, and the size effect to be caused by randomness of local material strength;
7. A power law of exponent −1/2 corresponding to linear elastic fracture mechanics (LEFM) and supplemented with an upper bound (small-size cutoff); and
8. A purely empirical power law of exponent −1/3 proposed in 2003 by an ACI subcommittee.

For comparison, Fig. 2 shows all the data points of the existing ACI-445F database. What is striking in this figure is that the very different curves of the aforementioned formulas look almost equally good (or equally bad). The reason is that the size range covered by the data is not broad enough and the scatter is enormous. The size range cannot be significantly extended without very large financial outlays. The main cause of the enormous scatter is that test data for different concretes, different shear spans, different reinforcement ratios, etc., are mixed in one and the same database. This man-made scatter cannot be filtered out to a significant extent because its causes are poorly understood. The problem is compounded by the fact that most of the data sets included in the database involve only a single beam size (depth) or a negligible size range because their original purpose was to clarify influences other than the size effect.

Some efforts are being made to choose among various formulas by comparing the coefficients of variation of the errors of each formula, calculated for the existing ACI-445F database. But such
efforts are futile. The coefficients of variation of the deviations of the formula from the data points are almost the same for all the formulas, good or bad. The arbitrariness of such a comparison then inevitably leads a committee to a political choice.

The most serious obstacle to extracting size effect information from the ACI-445F database is the fact that the vast majority (more than 97%) of its 398 data points come from tests motivated by different objectives (such as the effect of concrete type, reinforcement, shear span, etc.), in which the beam depth was not varied at all. To document the problem, see Fig. 3, which shows some such test data (marked by an oval) in comparison with the size effect law (Bažant 1984) and with the data points from two test series of the broadest available size range. These data contaminate the database by irrelevant scatter, caused by influences that cannot be eliminated because they are poorly understood. Such contamination widens the scatter band of the database and masks the size effect trends of the individual data series with a significant size range. When this database is fitted with a power law (a straight line in Fig. 4), the best-fit exponent (slope of the straight line) will depend on the beam size distribution in the database; see Fig. 4, which illustrates how the shifting of a hypothetical data cloud to smaller or larger sizes can yield any power law with exponent between 0 and −1/2. Obviously, such statistical inferences are not objective; they depend on the frequency of test data in various size intervals, which is a subjective choice of the experimenters, influenced by the funds available.

So, is the existing ACI-445F database useless? Not at all. But it should be used only for calibrating the size effect formula after the basic form of that formula has already been selected, and nothing else. The selection of the best form of the formula must be based on a sound theory. The theory must be such that it could also capture similar size effects in other types of failure with the same physical source, occurring in concrete as well as other quasibrittle materials. The theory should of course be experimentally validated. This can be done only by comparisons with individual test series in which, ideally, no parameter but the beam size is varied and the size range is broad enough in relation to the inevitable random scatter. Because of the high random scatter in beam shear tests, the size range should be at least 1:8. In particular, the concrete must be one and the same, and geometrical similarity of the beams of various sizes should be maintained as closely as possible, so as to prevent polluting the data set by uncertain influences other than those of the size.

Currently there exist only about 11 data series, from eight investigator teams (Leonhardt and Walther 1962; Kani 1967; Bhal 1968; Iguro et al. 1985; Bažant and Kazemi 1991; Shiyoa and Akiyama 1994; Collins and Kuchma 1999; Angelakos et al. 2001), satisfying these criteria at least approximately (a few more have a significant size range but grossly depart from geometrical similarity). Only two of them, namely the 1991 Northwestern tests (Bažant and Kazemi 1991) and the recent Toronto tests (Collins and Kuchma 1999; Angelakos et al. 2001), satisfy these criteria quite closely. Of these two, the Northwestern ones are reduced-scale model tests, which have the advantage that they achieve (thanks to the reduced scale) the largest dimensionless size so far, as measured by the brittleness number $\beta$.

Recently, while limiting consideration to test series in which the beam depth was significantly varied, the aggregate of such test series was evaluated jointly, without paying attention to the trends of the individual test series. However, such an approach is again misleading. To illustrate it, consider Fig. 5 showing bilogarithmic
plots of $\log \nu_c$ versus $\log d$ (on the left) for two sets of four hypothetical data series with the same range of beam depth $d$ (in log scale), generated so as to match exactly the size effect law for four different concretes; (b), (d) incorrect inference made by regression of the combined data set. Note that shifting of the chosen size range of data can yield any desired slope of regression line.

Fig. 5. Example of fallacious statistical analysis: (a), (c) hypothetical perfect data generated so as to match exactly the size effect law for four different concretes; (b), (d) incorrect inference made by regression of the combined data set. Note that shifting of the chosen size range of data can yield any desired slope of regression line.

Fig. 6. (a) Softening stress–separation curve of cohesive (or fictitious) crack model and (b) geometry of reinforced concrete beam

Variables in Problem of Failure and Size Effect

Problems of quasibrittle fracture in the normal range of interest are very difficult to solve. However, the asymptotic properties are simple. For very small structure sizes, the asymptotic solution can be obtained by plastic limit analysis, and for very large sizes, by linear elastic fracture mechanics [(LEFM)—a theory in which the fracture process zone is a point and the structure, which may be inelastic, uploads during fracture propagation elastically]. These asymptotic situations are often outside the size range of practical interest (being approached closely, for example, only for concrete beams 1 mm deep and 100 mm deep, respectively). It is nevertheless very helpful to know these asymptotic situations because a good approximate solution for the intermediate practical size range can be obtained by some sort of interpolation between these asymptotic cases, called asymptotic matching.

Although the detailed mechanism of shear failure of reinforced concrete beams is complicated, it clearly consists of some complex form of cohesive softening fracture. The material softening due to distributed fracturing may also be taking place but because it must localize into narrow bands, it may be approximated as cohesive softening fracture. This fact alone suffices to establish the basic asymptotic properties of size effect.

The problem of shear capacity of the cross section involves the following variables:

1. The shear strength $\nu_c$ to be predicted, defined as $\nu_c = V/b_w d$ ($\nu_c$ here plays the role of nominal strength of structure, normally denoted as $\sigma_0$); $V$=shear force, $d$=depth from the top face of beam to the centroid of longitudinal reinforcement, and $b_w$=width of the web, which equals the total beam width $b$ if the cross section is rectangular.

2. The characteristic size of the structure, chosen as the beam depth $d$ (distance from top surface to reinforcement centroid).

3. Parameters $G_f$ and $\sigma_0$ of the cohesive crack model, which automatically exhibits size effect [Bažant 2002; Bažant et al. 2002]. This model, developed in the works of Barenblatt 1959, 1962; Leonov and Panasyuk 1959; Rice 1968; Palmer and Rice 1973; and others (Knauss 1973, 1974; Smith 1974; Wnuk 1974; and Kfouri and Rice 1977), was pioneered for concrete by Hillerborg et al. (1976) under the name fictitious crack model (Pettersson 1981; Bažant and Planas 1998). For the purpose of maximum load analysis, it is useful to define the fracture energy $G_f$ as the area (Fig. 6) under the initial tangent of the softening curve of cohesive stress versus crack–face separation (Bažant et al. 2002). The shape of the softening curve is assumed to be fixed, which means that all the other parameters of the softening curve, such as the total fracture energy $G_F$, representing the area under the entire curve, are related to $\sigma_0$ and $G_f$.

4. Geometry parameters $\rho_w = A_t/b_w d$, $L_1/d$, $L_2/d$, $...$. 

1880 / JOURNAL OF STRUCTURAL ENGINEERING © ASCE / DECEMBER 2005
which represent the reinforcement ratio ($A_r =$ cross section area of steel) and the ratios to $d$ of all dimensions $L$, $L_1$, $L_2$, ... defining the span $L$, the beam length, the cover thickness, and the distances defining the locations of loading points. The beam width $b$, need not be included since its effect on $v_c$ is known to be negligible. Some of the parameters $L_i$ characterize the crack shape and crack tip location at maximum load. Normally the cohesive crack model is considered to describe the tensile fracture, in which case $\sigma_{0}\equiv f_c'$ = tensile strength of concrete. However, the shear failure of the beam is doubtless triggered by shear-compression fracture of concrete in the region above the tip of the main crack and delamination fracture of the concrete cover (as a result of dowel action). Bond fracture (propagation of a crack along the steel–concrete interface, with subsidiary cracks emanating from bar deformations) may also play some role. Note that shear-compression fracture, or compression crushing (which is an essential criterion for brittle failure in a realistic strut-and-tie model) exhibits a pronounced size effect and is also described by the crack-band model (in the form of triaxial softening damage) or by the cohesive crack model (adapted for compression). Likewise, the cover delamination and propagation of bond fracture also require a cohesive crack model.

In the case of crack band or cohesive crack models for compression fracture, the strength limit is the compressive strength $\sigma_{0}\equiv f_c'$, and in the case of bond fracture it is the bond strength. For our general dimensional analysis, we do not need to distinguish among these diverse strength limits (and that is why the strength is here generally denoted as $\sigma_{0}$). The fracture energy together with $\sigma_{0}$ imply a material characteristic length

$$l_0 = \frac{EG}{\sigma_0^2}$$

introduced by Irwin (1958) in the context of metals and by Hillerborg et al. (1976) for concrete; $l_0$ is proportional to the length of the fracture process zone at the tip of a crack or crack band, and is approximately equal to it when the structure is very large; see Bažant and Planas (1998).

The solution represents a functional relation which may generally be written as

$$\Phi(v_c, \sigma_0, EG_f, d, \rho_w, L, L_1, L_2, ...L_{m-1}) = 0$$

(2)

Here we multiply $G_f$ with Young’s elastic modulus $E$ because it is known that maximum loads in fracture mechanics do not depend separately on $E$ or $G_f$ but only on the parameter $\sqrt{EG_f}$ which, according to Irwin, defines fracture toughness $K_c$. In Eq. (2), we have a total of $n_1=5+m$ parameters.

Based on the fracturing truss model [Eqs. (9)-(11) in Bažant (1997)], an example of function $\Phi$ is

$$\Phi(v_c, ... ) = k_c v_c^2 (1 + a^2/d^2) - EG_f/(1 + d/\kappa_0 l_0) = 0$$

(3)

where $k_c$ and $\kappa_0$ = empirical functions of $f_c'$ and $\rho_w$. However, although this model yields a realistic form of size effect, it turned out impossible to estimate realistically the effective slope and cross section of the compression “strut.” Consequently, this model per se appears to be insufficient to capture realistically the effects of other influencing parameters, including $a/d$, $\rho_w$, $f_c'$, and $d_0$. This is why the size effect must be justified by general dimensional analysis, which is valid regardless of the detailed failure mechanism and of influences other than the size.

**Small- and Large-Size Asymptotes Dictated by Dimensional Analysis**

The number of unknowns may be reduced by introducing dimensionless parameters. According to the II theorem of dimensional analysis (Vasny 1892; Riabouchinsky 1911; Buckingham 1914), the number $n$ of independent dimensionless parameters is equal to the total number $n_1$ of all parameters minus the number of independent physical dimensions. Combining this theorem with the known physical meaning of $l_0$ will suffice to determine the asymptotic behaviors.

First, consider that $l_0 \ll d$. Since $l_0$ is known to characterize the size of the fracture process zone, it follows that this zone becomes infinitely smaller than $d$. This means that, in relation to the beam depth, all of the fracture is happening at only one point, the propagating fracture tip. This is the situation in LEFM, in which the stress at the fracture tip is infinite, making it impossible to use material strength $\sigma_0$ to judge load capacity. So, $\sigma_0$ must be removed from the list of parameters in Eq. (2). Function $\Phi$ of 5 $+m$ parameters then becomes a function, $\Phi$, of 4 $+m$ parameters, i.e.,

$$\Phi(v_c, EG_f, d, \rho_w, L, L_1, L_2, ...L_{m-1}) = 0$$

(4)

There are only two independent physical dimensions—length and stress—because the dimensions of all the other parameters in Eq. (2) can be obtained as products or the ratios of length and stress (for instance, the physical dimension of $EG_f$ is, in SI units, Pa $\times$ 1/m$^2$ or N/m$^3$). Therefore, according to the II theorem, the problem of failure can be recast as a functional relation among 2 $+m$ dimensionless parameters, which may be written as follows:

$$\Phi(u_c^2 d/EG_f, \rho_w, L/d, L_1/d, L_2/d, ...L_{m-1}/d) = 0$$

(5)

If the structures of different sizes $d$ are geometrically similar (which includes the condition that the loadings, supports, and the main cracks in specimens of different sizes must be similar), then $\rho_w$, $L/d$, $L_1/d$, $L_2/d$, ... are all constant, and so the first parameter in the foregoing equation must also be constant, i.e., $u_c^2 d/EG_f$= constant. It follows that

$$v_c = C_0 \sqrt{\frac{EG_f}{d}} \left(\frac{\text{constant}}{\sqrt{d}}\right)$$

(6)

where $C_0$=constant depending on the geometry parameters.

Note that the use of $G_f$ in the last condition requires the fracture length at the moment of failure to be non-negligible. When the failure occurs right at fracture initiation from a smooth surface (as in the modulus of rupture test of flexural strength of reinforced beams), then the last equation does not apply because an infinitely short crack has a vanishing energy release rate (i.e., vanishing derivative of the complementary energy of the structure with respect to the crack length; Bažant and Planas 1998).

Eq. (6) is a power scaling law that is characteristic of LEFM and is observed in the case where the cracks at maximum load are large and geometrically similar, or where the structures contain geometrically similar notches. In a plot of log $v_c$ versus log $d$, this asymptotic scaling is represented by a straight line of slope $-1/2$ (note that the power law with exponent $-1/3$, i.e., $v_c \propto d^{-1/3}$, which has recently been proposed on the basis of fitting of the entire ACI-444F database, as shown by the data variation of highly
scattered variables other than $d$, would be justified by dimensional analysis only if the fracture energy had the irrational dimension of $J/m^{3/4}$ instead of $J/m$ where $J=N \cdot m=Joule$).

Second, consider that $l_0 \gg d$. In this case, the fracture process zone occupies the entire cross section, and so there can be no fracture propagation. Hence, the failure load must be independent of $K_c$ or $EG_r$. According to the $\Pi$ theorem, the problem again reduces to a functional relation among $2+m$ dimensionless parameters, which may be chosen as follows:

$$\bar{\Phi}(v_c, \sigma_0, p_w, L/d, L_1/d, L_2/d, \ldots, L_m/d) = 0 \quad (7)$$

Noting again that, for geometrically similar structures, $\rho_w, L/d, L_1/d, L_2/d, \ldots$ are constant, we conclude that the first parameter in the foregoing equation must also be constant, i.e., $v_c/\sigma_0=constant$, and so

$$v_c = C_1 \sigma_0 = constant \quad (8)$$

where $C_1=constant$. So, in this asymptotic case the size effect is absent. This is characteristic of plastic limit analysis as well as any theory in which the material failure condition is expressed solely in terms of the stress and strain tensors.

From the foregoing dimensional analysis, it follows that the size effect law must be some sort of a transition between two asymptotes of slopes 0 and $-1/2$. Nothing more can be learned from dimensional analysis alone. To learn more, we must take into account the second-order asymptotic properties (Bažant 2002, 2004).

**Second-Order Asymptotic Properties**

Based on the form of the differential equations and boundary conditions governing failure, it has been generally proven (Bažant 2001, 2002) that the first two terms of the small-size and large-size asymptotic expansions of size effect based on the cohesive (or fictitious) crack model must have the following form:

for $d \to 0$

$$v_c = 1 - \frac{d}{d_1} - \ldots \quad (9)$$

for $d \to \infty$

$$v_c = \sqrt{\frac{d}{d_0}} \left[1 - \frac{d}{2d_0} + \ldots \right] \quad (10)$$

[Bažant (2002), Eqs. (9.125), (9.126)]; $d_0, d_1, C_0$, and $C_1$ are constants with respect to the size effect (i.e., they depend on structure geometry but not on size $d$). These asymptotic properties hold true under the condition that the softening stress-separation curve of the cohesive crack model begins its descent with a finite slope, which is known to be true for concrete.

The asymptotic properties in Eqs. (9) and (10) apply to all types of failure due to cohesive fracture or localization of distributed damage, provided that either there are large notches (which is not our case) or large geometrically similar fractures that develop in a stable manner prior to the maximum load. The fact that the fracture patterns in small and large beams are approximately similar is documented by many laboratory experiments as well as finite element simulations. If, for example, the depths of fracture at maximum load in small and large beams were 20 and 80% of cross section depth, respectively, or if the fracture in small beams were running almost vertically and in large beams almost horizontally, then this assumption would not apply. But, from experience, this is not the case.

The small-size asymptotic properties in Eq. (9) have been analytically derived by transformations of the differential equations of continuum mechanics and boundary conditions of the cohesive crack model. The large-size asymptotic properties in Eq. (10) have been derived by asymptotic expansion of: (1) equivalent LEFM or (2) the $J$ integral, or (3) the smeared-tip method (Bažant 2002). Knowing these properties, one can extend dimensional analysis to obtain a simple expression for the transition between the asymptotes, which we do next (following the general procedure in Bažant 2004).

**Size Effect Transition Between Asymptotes**

The fact that there are only two independent physical dimensions (stress and length) means that, according to the $\Pi$ theorem, the number of independent dimensionless variables is $n=3+m$. So, the governing equation can be expressed in the form

$$F(\Pi_1, \Pi_2, \ldots, \Pi_n) = 0 \quad (11)$$

Many diverse choices of dimensionless variables $\Pi_j$ ($i=1,2,\ldots,n$) are possible. However, as shown in Bažant (2004), it is convenient to make a choice for which $\Pi_1=0$ when $D/l_0 \to 0$ and $\Pi_2=0$ when $D/l_0 \to \infty$. It is natural to limit consideration to dimensionless monomials, and then the most general choice is

$$\Pi_1 = \left( \frac{v_c}{\sigma_0} \right)^\varrho \left( \frac{d}{l_0} \right)^\varrho \quad (12)$$

with $\Pi_2 = p_w, \Pi_3 = L/d, \ldots, \Pi_n = L_{m-1}/d$ ($\varrho, \varrho_u, u, v = four \ unknown \ real \ constants$). If we let $\Pi_1=0$ correspond to $D \to 0$, then $F(0, \Pi_2, \ldots) = 0$, which implies that $\Pi_2=constant$ or $v_c^\varrho d^\varrho=constant$ for $d \to 0$. This must be the case of no size effect; hence $v=0$. If we let $\Pi_2=0$ correspond to $D \to \infty$, then $F(\Pi_1, 0, \ldots) = 0$, which implies that $\Pi_1=constant$ or $v_c^\varrho d^\varrho=constant$, or $v_c \propto d^{-u/p}$ for $d \to \infty$. This must be the LEFM scaling; hence $u/p=1/2$ or $u/p=2$. To find $p$ and $q$, we truncate the Taylor series expansion of $F$ after the linear terms

$$F(\Pi_1, \Pi_2, \ldots) \approx F_0 + F_1 \Pi_1 + F_2 \Pi_2 = 0 \quad (13)$$

or

$$F_1 \left( v_c \sqrt{d \sigma_0 l_0^\varrho} \right)^\varrho = -F_0 \quad (14)$$

where $F_1 = \left[ \frac{\partial F}{\partial \Pi_1} \right]_{\Pi_2=0}$ and $F_2 = \left[ \frac{\partial^2 F}{\partial \Pi_1^2} \right]_{\Pi_2=0}$ (evaluated at $\Pi_1 = \Pi_2 = 0$) and $F_0 = F(0, 0) (F_{01} = F_{02} = 0)$. For general $p$ and $q$, the last equation cannot be solved explicitly for $v_c$, but it can for $d$

$$d = l_0 \sigma_0^\varrho (F_0/F_1)^{2/[q+1]} \left(1 + (F_1/F_0) v_c^\varrho \right)^{2/p} \quad (15)$$

This may be compared to the inverse expansion of the large-size asymptotic expansion [Eq. (10)], which can be shown to have the form

$$d = \frac{l_0}{\sigma_0\varrho} (F_0/F_1)^{2/[\varrho+1]} \left(1 + (F_1/F_0) v_c\varrho \right)^{2/p} \quad (15)$$
\[ d = Bv_c^2(1 - Cv_c^2 + \ldots) \]  
(16)

for \( v_c \rightarrow 0 \) (\( B,C \) positive constants). Evidently, matching of the first two terms of this expansion by Eq. (15) requires that \( p = q = 2 \). Then Eq. (14) can be solved for \( v_c \). This furnishes the size effect law of the following general form:

\[ v_c = \frac{v_0}{\sqrt{1 + d/d_0}} \]  
(17)

in which \( v_0 = \sigma_0(1 - F_0/F_2)^{1/2} \) and \( d_0 = l_0F_2/F_1 \). For \( D/l_0 \rightarrow 0 \), Eq. (17) has (according to the binomial series expansion) the approximation

\[ v_c \approx v_0(1 - d/d_0) \]  
(18)

This verifies that the form of the second term of the small-size expansion Eq. (9) can be matched too. The fact that Eq. (17) indeed satisfies the first two terms of the large-size asymptotic series expansion dictated by Eq. (10) can be readily checked by the following approximation (according to the first two terms of the binomial series expansion in terms of \( \xi = d/d_0/d \)):

\[ (1 + 1/\xi)^{-1/2} = \xi^{1/2}(1 + \xi)^{-1/2} = \sqrt{\xi}(1 - \xi/2) \]  
(19)

To sum up, the known asymptotic properties of cohesive fracture require that \( p = q = 2, u = 1, \) and \( v = 0 \). The asymptotic matching based on the first-order approximation in Eq. (13) then unambiguously leads to the size effect law in Eq. (17), proposed in general by Bažant (1984) on the basis of a much more restrictive and simplified mathematical argument, and applied to beam shear failure by Bažant and Kim (1984). The smooth transition between the aforementioned two asymptotes is centered at \( d_0 \), called the transitional structure size, which separates sizes \( d > d_0 \), for which the failure is predominantly brittle, from sizes \( d < d_0 \), for which the failure is predominantly ductile (for this reason, the ratio \( \beta = d/d_0 \) is called the brittleness number; see Bažant 1987; Bažant and Planas 1998).

Note that the foregoing derivation of size effect law has not relied on continuum mechanics. Even if a specimen or structure is not large enough (compared to the aggregate size) for a continuum approximation to hold, the foregoing derivation remains valid because, even for random discrete lattice or particle models, the material strength \( \sigma_0 \) and fracture energy \( G_f \) can be defined statistically.

**Question of Uniqueness of Size Effect Law**

As a possible generalization, one might consider that the softening stress-displacement law of cohesive fracture could terminate with a finite residual stress \( \sigma_r \), representing some frictional-plastic residual strength of the material. In that case, the fracture energy would be the area between the softening curve and the horizontal line \( \sigma = \sigma_r \) and \( v_c \) would have to be replaced in the foregoing derivation with \( v_r - v_c \) where \( v_r \) is a finite residual shear strength of the cross section. The dimensional analysis combined with asymptotic matching would thus yield, instead of Eq. (17), a more general formula

\[ v_c = v_0(1 + d/d_0)^{-1/2} + v_r \]  
(20)

Shear-compression fracture can, in principle, exhibit a finite residual strength \( \sigma_r \) (Bažant 2002), but only under conditions of high triaxial confinement, which could hardly be provided in the zone above the tip of the main diagonal crack. The data for beam shear strength within the size range tested so far give no evidence of a finite residual shear strength \( v_r \) and so, in the interest of safety, it will be assumed that \( v_r = 0 \).

More complex formulas of the same asymptotic accuracy, but more flexible in data fitting, can be obtained by replacing Eq. (12) with dimensionless polynomials or various monotonic functions of \( \Pi_1 \) and \( \Pi_2 \). Such formulas can make a significant difference only if the size effect needs to be modeled for a size range exceeding about \( 1:30 \) (Bažant 1999, 2002).

For instance, consider what happens if one chooses \( \Pi_1 = (v_r/\sigma_0)^{d/l_0} \) and \( \Pi_2 = v_r/l_0 \). Matching of the small-size and large-size asymptotes would then yield the result \( v_c = v_0(1 + \sqrt{d/d_0}) \), which is different from Eq. (17). However, the second-order asymptotic terms in Eqs. (9) and (10) would not be matched.

On the other hand, nothing would, for example, prevent us from taking \( \Pi_1 = v_r^2/\sigma_0^2(d/l_0)(1 + d^2/l_0^2) \). In that case, the asymptotic terms in Eqs. (9) and (10) would be satisfied up to the second order, but the size effect expression would become considerably more complex.

There are also physical arguments that lead to dimensionless variables \( \Pi_1 \) and \( \Pi_2 \) as chosen. These variables characterize failure in terms of energy. For \( p = q = 2, u = 1, \) and \( v = 0 \), \( \Pi_1 \) represents (except for a geometry-dependent factor) the ratio of the energy release rate of a crack to \( G_f \), which is what controls the propagation of a crack when the structure is much larger than the fracture process zone. \( \Pi_2 \) represents (except for a geometry-dependent factor) the ratio of the strain energy density to its value at the strength limit of the material, which is what controls failure when the structure is smaller than a fully developed fracture process zone.

For data fitting, the size effect law in Eqs. (17) or (20) has the advantage that it can be algebraically converted to a linear regression plot

\[ Y = Ad + C \]  
(21)

where

\[ Y = \frac{1}{(v_c - v_r)^2} \]

\[ v_0 = \frac{1}{\sqrt{C}} \]  
(22)

\[ d_0 = \frac{C}{A} \]

(Part II, Fig. 1 right). The regression is a convenient way to identify \( v_0 \) and \( d_0 \) from size effect test data.

For less complex problems in which the fracture path is simple and known, the size effect law in Eq. (17) has also been derived by several other methods, e.g., by asymptotic expansion of equivalent LEM, asymptotic expansion of the \( J \) integral, and asymptotic expansion of the smeared-tip method (Bažant 2002). The first of these methods further yields expressions for the size effect parameters in terms of the energy release function of fracture mechanics. Such expressions, however, could hardly be obtained for the beam shear problem because the LEM crack pattern corresponding to infinite size extrapolation is unknown.

The size effect law in Eq. (17) has been verified numerically for many problems by finite element simulations based on the
crack band model, nonlinear damage model, cohesive crack models, and lattice or particle models of microstructure. Experimental verifications now include many types of failure of reinforced concrete structures, as well as rocks, toughened ceramics, fiber composites, brittle foams, snow slabs, and sea ice (the last up to a record size of 80 m × 80 m × 1.8 m) (Bažant and Planas 1998; Bažant 2002, 2004). In view of all this evidence, accumulated over the past 2 decades, it would be extremely surprising if the size effect law in Eq. (17) did not provide a good approximation for the shear failure of reinforced concrete beams.

Closing Comment

The next step is the verification and calibration of the theoretically derived formula. This step, a discussion of previously proposed alternative formulas, and the formulation of conclusions, are relegated to Part II of this paper which follows.

Appendix. Question of Difference Between Notched Specimens and Unnotched Structures with Large Cracks

Application of the size effect law Eq. (17) to unnotched structures with large cracks rests on two hypotheses:

- **Hypothesis I.** The major cracks at failure in small and large structures are geometricaly similar.
- **Hypothesis II.** The fracture process zones at the tips of a notch or a crack give approximately the same energy dissipation rates.

According to the cohesive (or fictitious) crack model, both hypotheses are asymptotically exact for d → ∞ and thus are good approximations for large enough beams. These hypotheses are justified by the following theorem: *As long as the fracture problem has a unique solution, sharp (LEFM) cracks in geometrically similar structures have similar paths and, at maximum load (as well as other corresponding stages of loading), also geometrically similar lengths.* This theorem can be rigorously proved by scaling transformations of all the differential equations, boundary conditions, and crack face conditions of equivalent LEFM, which is an approximation (widely used in nonlinear fracture mechanics) based on the assumption that a crack with a large fracture process zone is approximately equivalent to a sharp (LEFM) crack with a tip located in the middle of the fracture process zone.

The cohesive crack model shows that the cohesive stress $\sigma_{\text{fail}}$ at the tail of the fracture process zone is zero for a naturally growing crack. But when the fracture process zone is attached to the notch tip, $\sigma_{\text{fail}}$ is, for a finite beam size, nonzero (approaching zero only if the size tends to infinity). The consequence is that the size effect with second-order asymptotic accuracy for large sizes is slightly different from Eq. (17),

$$v_c = v_0 \left( \frac{d_1}{d_1 + d} \right)^{1/2}$$

(23)

[see Bažant 2001 and in detail Eqs. (9.77) and (9.109) in Bažant 2002]. Here $d_0$ and $d_1$ are constants, and $d_1$ must be larger than $d_0$. Obviously this law has the same asymptotes as Eq. (17) (which is of course required by the dimensional analysis presented here) and for $d_1 \rightarrow \infty$ becomes identical to Eq. (17). For finite $d_1$, the transition between the asymptotes is more abrupt (with a sharper curvature) than for Eq. (17), and the $v_c$ values in the intermediate range are always larger. But they are only slightly larger. Based on preliminary numerical cohesive crack simulations, $d_1 \approx 5d_0$ to $10d_0$, but then the difference from Eq. (17) is negligible compared to the uncertainty due to the scatter of test results. This is one justification of the use of the simpler size effect law Eq. (17) for beam shear. Besides, Eq. (23) has one more unknown parameter than Eq. (17), and the data scatter makes it next to impossible to identify it experimentally.

Another justification stems from the fact that, before an overload to failure, the beam may be subjected to cyclic loading. Such loading reduces the cohesive stresses in a naturally grown crack to almost zero, which means that a preexisting fatigued crack is stress free and thus acts like a notch. Assuming that cyclic loading may occur is on the side of safety.

Finally, those who think that the difference between a notch and a natural crack is important should note that the growth of the diagonal shear crack is usually not what causes the load to peak. Rather, it is the growth of shear-compression fracture across the ligament above the tip of the diagonal shear crack. How the difference between a notch and a naturally grown crack might affect such fracture is not known. But the present asymptotic considerations based on dimensional analysis, which led to Eq. (17), circumvent the question of detailed failure mechanism and must apply to shear-compression fracture as well.

References


Designing Against Size Effect on Shear Strength of Reinforced Concrete Beams Without Stirrups: II. Verification and Calibration

Zdeněk P. Bažant¹ and Qiang Yu²

Abstract: After theoretical derivation of the general form of the size effect formula for beam shear in the preceding Part I, this Part II presents experimental verification by least-square fitting of those existing individual data sets that have a broad size range. Subsequently, empirical prediction formulas for the size effect parameters, consisting of the asymptotic small-size strength $\nu_0$ and the transitional size $d_0$, are calibrated by least-square regression of (1) a recent American Concrete Institute database with 398 data points, and (2) a combination of this database with large-scale Japanese tests and Northwestern reduced-scale model tests. Previous alternative proposals for dealing with the size effect in beam shear are also discussed.

DOI: 10.1061/(ASCE)0733-9445(2005)131:12(1886)

CE Database subject headings: Size effect; Fracture; Shear strength; Concrete beams; Calibration.

Verification by Comparisons with Individual Size Effect Data

The form of the size effect formula must be checked by comparison with data sets that are not significantly contaminated by variation of parameters other than size $d$. Ideally, the beams tested, as well as the reinforcing bars and the bar cover, must be geometrically scaled, cover a size range broad enough compared to the scatter band width, and use one and the same concrete (ideally cured and tested under the same environmental conditions). Only two data sets came close to this ideal situation:

- Reduced-scale tests in 1991 at Northwestern University (Bažant and Kazemi 1991), in which beams having depths from 0.021 to 0.33 m, maximum aggregate size 4.8 mm, and reduced-scale bars with standard ASTM deformations (bought from PCA, Skokie) were tested (with three identical beams for each size); and

- Recent normal-scale tests at the University of Toronto (Collins and Kuchma 1999; Angelakos et al. 2001), with beam depths from 0.11 to 1.89 m and maximum aggregate size 10 mm (only one specimen of each size was tested, with an interval of up to two years elapsing between subsequent tests).

Eq. (17) in Part I leads to a very good agreement with both data series, as can be seen from the optimum fits of these data.

This is documented by Fig. 1, and validates the correctness of the form of the formula. Three kinds of comparisons are shown: those with the Northwestern data alone (on top), those with the Toronto data alone (in the middle), and those with both data combined (at the bottom). In the last case, the optimum fit has been obtained under the constraint that the minimum relative values of $\nu_0/\nu_{0_0}$ would be the same in both cases (because the smallest specimens tested were in both cases the smallest specimens that could be cast, for the respective aggregate size).

The optimum fits in Fig. 1 (on both the left and right) have been obtained by least-square fitting in the logarithmic scale using the Levenberg–Marquardt nonlinear optimization algorithm. The optimum parameters $\nu_0$ and $d_0$ of the size effect formula are given in Fig. 1, along with the coefficient of variation of errors which is defined later. All the fits are shown both in the plots of $\log(\nu_0/\nu_{0_0})$ versus $\log(d/d_0)$ (on the left), and in linear regression plots of $(\nu_0/\nu_{0_0})^2$ versus $d/d_0$ (on the right). The top two rows show the individual fits of each data set, and the last row shows a combined fit of both data sets obtained by the constrained optimization already mentioned.

Although linear regression in transformed variables has not been used to get the optimum fit, the linear regression plots are also shown on the right of Fig. 1, along with their coefficients of variation. The results of nonlinear optimization and linear regression are close, but not identical. The nonlinear optimization is slightly preferable because it implies a better weighting of the data (Bažant and Planas 1998).

What is particularly noteworthy is that the test series at both Northwestern University and University of Toronto verify very well the asymptotic behavior of fracture mechanics required by the foregoing asymptotic arguments coupled with dimensional analysis [Eq. (6) in Part I]. This confirms that the asymptotic slope of size effect curve in a doubly logarithmic plot must be $-1/2$.

In addition, it must be emphasized that the well-known Japanese tests (Iguro et al. 1984; Shioya and Akiyama 1994) do not at all contradict this asymptotic slope, even though they were previously interpreted by a power law of exponent $-1/4$; see Fig. 2.

¹McCormick Institute Professor and W. P. Murphy Professor of Civil Engineering and Materials Science, Northwestern Univ., Tech-CEE, 2145 Sheridan Rd., Evanston, IL 60208. E-mail: z-bazant@northwestern.edu
²Graduate Research Assistant and Doctoral Candidate, Northwestern Univ., Tech-CEE, 2145 Sheridan Rd., Evanston, IL 60208. E-mail: qiangyu@northwestern.edu

Note. Associate Editor: Khalid M. Mosaalam. Discussion open until May 1, 2006. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on October 7, 2004; approved on July 24, 2005. This paper is part of the Journal of Structural Engineering, Vol. 131, No. 12, December 1, 2005. ©ASCE, ISSN 0733-9445/2005/12-1886–1897/$25.00.
Fig. 1. Two test series of geometrically similar beams with a significant size range, fitted by size effect law [Eq. (17)] in Part I. Top: Optimum fit of data of Bažant and Kazemi (1991, Northwestern). Middle: Optimum fit of data of Podgornik-Štanik (1998, Toronto) and Yoshiida (2000, Toronto). Bottom: Constrained optimization of combined Northwestern and Toronto data. Left: Plots used in nonlinear optimization. Right: Optimization results shown in linear regression plots (note the disagreement with power laws of exponents \(-1/3\) and \(-1/4\)).

The comparisons with all the relevant test data thus confirm that the correct explanation of the size effect lies in fracture mechanics. This, in turn, verifies that the correctness of choosing the fracture energy and the material strength as the governing material parameters in dimensional analysis.

The fact that the test data are much closer to the linear elastic fracture mechanics (LEFM) asymptote of slope \(-1/2\) than to the horizontal asymptote means that the shear failure of beams is highly brittle. This further implies that the fracture energy is a more important material parameter than the material strength, and that, if finite element programs are used, they must be based on fracture mechanics [as another consequence, the importance of introducing without delay a standardized fracture test of concrete is thus highlighted; see Bažant et al. (2002)].

Several more data sets with at least a limited size range exist (Leonhardt and Walther 1962; Kani 1967; Bhal 1968); see Fig. 3. Some parameters of these tests deviated significantly from geometric similarity, which probably is the cause of increased scatter in Fig. 3. Nevertheless, the overall size effect trend of these data is seen to be modelled quite well by the present formula, and the asymptotic slope of \(-1/2\) is seen not to be contradicted.

Fig. 2. Size effect in two classical series of shear tests of very large beams conducted in Japan (Cov=coefficient of variation)

Expressions for Size Effect Law Parameters and Their Justification

The question now is how to predict in design the values of size effect parameters \(v_0\) and \(d_0\). In 1984, expressions for these constants were set up on the basis of certain simplifications based on the beam flexure theory and the concept of arch action (Bažant and Kim 1984; 1985). However, these expressions were superseded by later research. Further study (Bažant 1997) led to an energetic fracturing generalization of the classical truss model conceived by Ritter (1899) (recently renamed as the strut-and-tie model). This generalization was based on a simplified estimation of the energy release rate from either the potential energy or the complementary energy (under the hypothesis that the failure at maximum load is triggered by propagation of a compression damage band of a fixed width across the compression "strut"). Such analysis led rigorously to the size effect law in Eq. (17) of Part I, which furnished additional support for the general form of this law, and also gave an intuitive explanation as to why a size effect must arise. Recent comparisons with numerous test data from the databases compiled at Northwestern University and in American Concrete Institute (ACI) Subcommittee 445E have nevertheless revealed that the detailed mechanism of failure assumed in Bažant (1997) for the fracturing truss model was too simplified for capturing all of the dependence of coefficients \(v_0\) and \(d_0\) on parameters other than size \(d\). particularly the dependence on \(a/d\) and \(p_w\). For example, the inclination of the compression strut is steeper than predicted by a simple strut-and-tie model [see Fig. 7(a)] or by the composite beam action. Therefore, semiempirical expressions are introduced:

\[
d_0 = c_0 f_{c}' d_0^p \left( a/d \right)^q, \quad v_0 = k_0 f_{c}'^{w_1} d_0^{w_2} \left( k_1 + (a/d)^{w_3} \right)^{w_4}
\]

where \(k_0, k_1, c_0, r_1, \ldots, r_7 = \) constants; \(a/d = \) shear span (Part I, Fig. 6); \(p_w = \) steel ratio; \(f_{c}' = \) compressive strength of concrete; and \(d_0 = \) maximum aggregate size. Note that this expression is chosen in a form that makes negative values impossible. Could not the Pi-theorem be used to reduce these expressions to a dimensionless form?—No, because not all the influencing parameters are known. Let us now discuss four influences on the size effect parameters.

Material Strength

The higher the concrete strength \(f_{c}'\), the higher the brittleness of failure. Because a higher brittleness means greater proximity to LEFM, \(d_0\) must decrease with increasing \(f_{c}'\), and so \(r_1 < 0\). To agree with the current ACI design formula [which reads
beam action due to bond between concrete and steel bars must also be taken into account. According to Bažant and Kim (1984, 1985) one may write \( M = T(jd) \) where \( T = T(x) \) = tensile force in steel, \( j = j(x) \) = arm between the tensile and compression resultants, and \( x \) = distance of cross section from the simple support; then \( V = dM/dx = V_A + V_B \) where \( V_A = T(jd/dx) \) = shear force due to arch action, and \( V_B = d(T/dx)j/dx = \) shear force due to composite beam action. Making some reasonable simplifying assumptions, and setting \( x = d \) (which is the critical cross section), one finds that \( V_A/dx \) and \( V_B/dx \) = positive constant, denoted as \( k_1 \) [see Eqs. (6) and (9) in Bažant and Kim (1984, 1985)]. This argument justifies the term \( k_1 + \text{const.} (a/d) \) in Eq. (1), showing that \( r_1 = 1 \), i.e. \( v_0 = (k_1 + d/n) \), and further implies that \( r_3 = 0 \).

It is impossible to determine the effects of \( a/d \), \( d_{cr} \), and \( \rho_w \) merely by least-square fitting of the entire ACI-445F database because the trends are obscured by variation of many other parameters, in a similar way as already explained for the effect of size \( d \) by Figs. 3 and 5 of Part 1 (when \( a/d \) and \( d_{cr} \) are excluded, the \( \omega \) value for the best fit of this database increases from 15 to 16%, which is insignificant). Fortunately, the database contains a few test series focused only on \( a/d \) or \( d_{cr} \).

Figs. 4(a and b) present a series of 14 tests by Kani (1967) that has a significant range of \( a/d \) and gives a clear trend. In these tests, \( a/d \) varied from 3.5 to 8, for three different beam depths \( d \). Despite the scatter, it is clear from Fig. 4 that \( a/d \) does have a significant effect, and that the effect is similar for various sizes \( d \). In view of the scatter, one can justify nothing more detailed than a linear function of the inverse shear span, \( v_0 \propto (d/a+\text{constant}) \), independent of size \( d \) [this simple function was shown to result from the hypothesis that \( V_c \) is the sum of contributions from arch (or strut) action and composite beam action; Bažant and Kim (1984, 1985)].

**Fig. 3.** Optimum fits by size effect law [Eq. (17) in Part I] of remaining existing data that have a non-negligible size range but have gross deviations from geometrical similarity (all coordinates are log \( d, d \) in mm)

\[ v_0 = 2 \sqrt[3]{d} \] and is motivated by Pauw's (1960) observation that the tensile strength is approximately proportional to \( \sqrt[3]{d} \), we choose \( r_3 = 1/2 \).

**Shear Span \( a/d \)**

The fracturing truss model (Bažant 1997) leads to the expression \( v_0 \propto (d/a+\text{constant}) \). In the practical range, \( d/a \) ranges roughly from 3/10 to 3/8. These values are negligible compared \( a/d \), which ranges from 8/3 to 10/3, and so we may just as well write \( v_0 \propto (d/a+\text{small constant}) \). To better assess this constant, focusing on truss action (or strut action) is insufficient, and the composite

**Fig. 4.** (a) Kani's (1967) data showing the effect of shear span ratio \( d/a \) on \( v_0 \); (b) Kani's data replotted in variables that give linear regression according to proposed Eq. (4) for \( v_0 \); (c) Kling's 1985 data (see Reineck et al. 2003) showing the dependence of shear strength \( v_0 \) on steel ratio \( \rho_s \); and (d) data of Shioya and Akiyama (1994) showing the effect of changing maximum aggregate size \( d_{cr} \).
Maximum Aggregate Size $d_a$

The size of the fracture process zone (FPZ) is, in fracture mechanics (Bažant and Kim 1984; Bažant and Planas 1998), known to increase with $d_a$, but do so less than proportionally (because of the effect of $d_a$ on mix parameters, aggregate surface area and volume fraction of the interfacial transition zone). The transitional size $d_0$ is known to be roughly proportional to $d_a$ (Bažant and Planas 1998; Bažant 2002, 2004), and so $d_0$ should increase with $d_a$, though less than proportionally. Upon examining the data of Iguro et al. (1984) and of Shioiya and Akiyama (1994) [Fig. 4(d)], not included in ACI-445SF database, and also upon comparing the normal aggregate data with the reduced-scale Northwestern University data (Bažant and Kazemi 1991), for which $d_a=4.8$ mm, one concludes that, roughly, $d_0 \approx \sqrt{d_a}$ (i.e., $r_2=1/2$).

Longitudinal Steel Ratio $\rho_w$

An increase of $\rho_w$ must cause an increase of the inclined compression force $F$ transmitted through the ligament above the tip of the main diagonal crack. This must induce formation of a deeper ligament, stronger for transmitting $V$ (because no vertical shear stress can be transmitted near the top surface). The dependence of $v_c$ on $\rho_w$ is verified by Kung’s (1985) tests, in which $\rho_w$ varied from 0.5 to 1.82%; see Fig. 4(c) which shows the curve $v_c \propto \rho_w^{0.8}$ to fit well ($p_w^{0.6}$ gives here the best fit but $p_w^{0.8}$ is the best for the entire database).

Statistical Calibration by Least-Square Fitting of Database

The undetermined coefficients in Eq. (1) for $v_0$ and $d_0$ have been calibrated by least-square optimum fitting of the ACI-445SF database, consisting of 398 data points (see Fig. 2 of Part I and the dimensionless plots in Fig. 5). The subcommittee assembling this database decided to restrict the data to three-point loading (even though the design code must apply to all kinds of loading). Unfortunately, this restriction excluded the valuable Japanese tests of record-size beams (Fig. 2), because they were made under distributed loading. The reduced-scale three-point loaded tests at Northwestern University, with scaled deformed bars, were also excluded from this database, based on two arguments: (1) that the aggregate, of maximum size 4.8 mm, was supposedly too small (but this was advantageous for achieving the highest brittleness number $B$ so far among all tests); and (2) that the beams were supposedly too narrow (but the beam width was eight times the aggregate size, which is clearly sufficient, and if width of these tests were enlarged to 50 mm, the lower limit adopted by the subcommittee, the results would likely be about the same because it is generally agreed that beam width has no significant effect on $v_c$). In spite of these objections, the database adopted by ACI-445SF is used for the present calibrations.

Another problem with using the ACI-445SF database is that the maximum aggregate size $d_a$ was not reported for 57 data points (or 14%) among its 398 data points, although it is widely thought to have a significant effect on $d_a$ (this is best demonstrated by comparing the Northwestern and Toronto tests). Therefore, in addition to using for calibration the complete ACI-445SF database, also a modified database, consisting only of those points of the ACI-445SF database for which $d_a$ is known (341 points), and of the Northwestern and Japanese tests, has been used for an additional comparison.

The data fitting is a nonlinear statistical regression problem, and the choice of approach calls for some discussion. Let us denote by $\hat{v}_i$ (i=1, 2, ..., n; n=398) the measured data points in ACI database, and by $v_i$ the corresponding values of $v_c$ calculated from the formula. It appears that the best approach (Appendix) is not to minimize the sum of squared errors (or residuals), $\Sigma(v_i-\hat{v}_i)^2$, because the variance of data (or the scatter band width) decreases with increasing size (i.e., the data are heteroskedastic). To minimize statistical bias, the statistical variable $v_c$ should be transformed so as to make the variance approximately uniform (i.e., make the data approximately homoskedastic, independent of $d$). This may be achieved by least-square fitting in the logarithmic scale of $v_c$, i.e., by minimizing the square of the standard error of regression, $s_L$. Its unbiased definition is

$$s_L^2 = \frac{1}{n-n_p} \sum_{i=1}^{n} \left(\ln \frac{v_i}{\hat{v}_i} \right)^2$$

where $n_p$=number of free parameters in data fitting; in our problem, $n_p$ includes at least 5 (parameters $c_0$, $k_0$, $k_1$, $r_1$, $r_2$), and better 6 (with $r_2$ counted). The reason for subtracting $n_p$ is that $n_p$ data points can always be fitted perfectly. Because $\langle \ln v_i \rangle^2 = \langle d v_i^2 \rangle / v_c^2$, the scale transformation from linear to logarithmic has a similar effect as applying weights proportional to $1/v_c^2$ (except for the fact that the implied error distribution is Gaussian in the scale of $\ln v_c$ rather than $v_c$). The minimization was accomplished by a standard library subroutine for the Levenberg–Marquardt nonlinear optimization algorithm, which reduces the problem to a sequence of linear regressions. Since $s_L$, defined in the scale of natural (not decad) logarithm of $v_c$, is dimensionless, it may at the same time be regarded as the coefficient of variation. The reverse transformation to the linear scale of $v_c$ gives the following coefficient of variation of regression, characterizing the ratio of standard error in the linear scale to the mean of all data $\hat{v}_i$:

$$\omega = \left( e^{s_L} - e^{-s_L} \right) / 2$$

Note that $\omega$ and $s_L$ are almost equal because usually $s_L < 0.2$ (in which case their difference is of the order of $s_L^2/3$, i.e., $\ll 0.008$, as can be checked by Taylor series expansions). Also note that, whereas the approximate equality of $s_L$ and $\omega$ requires fitting in the scale of natural logarithm, $\ln v_c$. Figs. 5 are plotted, for convenience, in terms of the decadic logarithm, $\log v_c$.

The beam sizes in the database were chosen by various experimenters according to funding limitations and other subjective considerations. Unfortunately, the data points are crowded in the range of small sizes [d = 20 in. (510 mm), Part I, Figs. 1 and 2]. This distorts the resulting fit, giving insufficient weight to large-beam data which are the most important for extrapolation to still larger sizes that could be used in practice.

Ideally, the histogram of a database should be a horizontal line. Therefore, the range of $d$ was divided into constant intervals, of 10 in. (254 mm) width, and the resulting histogram, plotted in Fig. 5(a), was approximated by the smooth curve plotted. Each point in the database was then assigned a weight inversely proportional to the smooth histogram (because this is what would be needed to make the weighted histogram a horizontal line in the scale of $d$).

Because it is debatable whether it might be better to use a histogram in which the intervals are constant in the scale of $\log d$ rather than $d$, another calibration was conducted using weights inversely proportional to this kind of histogram. However, the results did not differ appreciably.
Fig. 5. (a) Histogram of existing data and its smoothing as a function of beam depth $d$, used to assign data weights in inverse proportion; (b–d) comparison of proposed Eq. (4) to weighted data, for (b) only those data for which $d_0$ is known; (c) complete ACI-445F database; (d) extended ACI database; (e) comparison of proposed Eqs. (9) and (10) to complete ACI-445F database; and (f) comparison of Eq. (4) to complete ACI-445F database in a linear regression plot terms of steel ratio $\rho_w$.

To provide additional safety margin, the code-making committee has customarily passed the curve of a design formula near the lower fringe, rather than the middle, of the scatter band. However, it has not been mentioned in the code that such practice implies a covert understrength factor (Bažant and Yu 2003) which implies an additional, hidden safety margin. This factor, which accounts for the uncertainty of design formula, is covertly imposed in addition to the safety margins provided by (1) the load factors, (2) the capacity reduction factor, and (3) by the fact that $f'_c$ is defined as a significantly smaller value than the mean compressive strength from testing. Properly, to avoid arbitrary bias and make any estimates of failure probability feasible, this kind of covert understrength factor must be obtained as a specified probability cutoff according to the least-square method—by taking the prediction formula obtained by least-square regression (solid curves in Figs. 5) and reducing it by the standard error of regression multiplied by a factor corresponding to the chosen probability cutoff (which can be taken from the Gaussian distribution table); this yields the dashed curves in Figs. 5 [besides, to make probabilistic estimates of failure probability feasible, this multiplying factor, the corresponding cutoff probability and the coefficient of variation of errors should all be stated in the design code; Bažant and Yu (2003)].

After determining the form of the dependence of $v_c$ on $d/d_0$,
\( a/d \), and \( \rho_w \) from selected relevant test data and deciding the statistical approach, the following design formula (in which \( v_c \) and \( f'_c \) are in psi, 1 psi = 6,895 Pa) has been identified by weighted least-square regression of the database and is now proposed for the design code:

\[
v_c = \mu \rho_w^{3/4} \left( 1 + \frac{d}{a} \right) \sqrt{\frac{f'_c}{1 + d/d_0}} \quad d_0 = \kappa f'_c^{-2/3} \tag{4}
\]

where \( \kappa = 3,800 \sqrt{d_a} \) if \( d_a \) is known, \( \kappa = 3,330 \) if not \tag{5}

\[ \mu = 13.3 \] for best fit, \[ \mu = 10 \] for design \tag{6}

The value \( \mu = 13.3 \) gives the solid curve shown in Fig. 5(c), representing the least-square fit (mean fit) of the ACI-445F database. For the aforementioned modified database, including the Japanese and Northwestern data, the same formula (with the same \( \mu \)) is shown in Fig. 5(b). The errors of the foregoing formula are characterized by

\[
\omega = 15.0\% \quad \text{for ASI-445F database}
\]

\[
\omega = 14.2\% \quad \text{for modified database}
\]

(\( \omega \) is the coefficient of variation). The optimum values for both databases were found to differ only negligibly. The optimized parameter values are rounded off to the extent that \( \omega \) would not be appreciably affected. Admittedly, the coefficients of variation of errors are quite large, but this is due to the enormous scatter in the complete database. When the contamination by non-size parameters is eliminated, the scatter is far less; see Figs. 1–3. Fig. 5(f) shows that the effect of steel ratio in the database is captured well.

In keeping with the previous practice, a reduced value, \( \mu = 10 \), giving the dashed curve in Fig. 5, is proposed for design code [to provide an extra safety margin; see Bažant and Yu (2003)]. This value is obtained as the mean \( \mu \) minus 1.65\( s_\mu \), which corresponds to a 5% probability cutoff based on Gaussian distribution of error scatter across the data band. Thus the probability of \( v_c \) lying below the dashed curve is 5%.

The reason for giving in Eq. (5) two options for calculating \( d_0 \) is that sometimes the design needs to be made before the maximum aggregate size \( d_a \) has been decided. Both expressions for \( d_0 \) give the same value when \( d_a = 0.77 \) in. (20 mm), which is almost equal to the mean \( d_a \) in the database, which is 0.7 in. (17.8 mm).

As seen in Fig. 6, for very small \( d \), the \( v_c \) value according to the proposed formula (4) is greater than predicted by the current ACI formula:

\[
v_c = 2 \sqrt[3]{f'_c} \tag{8}
\]

So, this classical formula may continue to be used safely within a certain range. The permissible safe range for Eq. (8) is \( d = 6 \) in. (150 mm). This is ascertained from the (ESDB) database plotted in Fig. 6, which reveals that, for \( d = 6 \) in. (150 mm), no beam test gave a shear strength less than value given by Eq. (8).

As a simple and safe (albeit often uneconomical) alternative [Fig. 5(e)], the simple formulas

\[
\text{for } d \leq 6 \text{ in.:} \quad V_c = 2b_w \sqrt{f'_c} \tag{9}
\]

\[
\text{for } d > 6 \text{ in.:} \quad V_c = 5b_w \sqrt{f'_c} \tag{10}
\]

can be used instead of Eq. (4).

\[ \begin{align*}
\text{Fig. 6. Left: Shear strength compared to the portion of ACI-445F database for } d \leq 6 \text{ in. (150 mm). Right: Sensitivity of standard deviation of regression errors to various parameters influencing } s_t. \end{align*} \]

Is it not necessary that the design satisfy both Eqs. (4) and (9), thus preventing \( v_c \) from ever being higher than according to the current code? No, because no experimental or theoretical evidence calls for that.

For the size intervals delimited by \( d = 0, 10, 20, 30, 40, 50, \) and \( 80 \text{ in. (0, 254, 510, 760, 1,020, 1,270, and 2,030 mm), which contain 164, 182, 20, 22, 7, and 3 data points of ACI-445F database, respectively, the coefficients of variation of formula errors are found to be } \omega = 18.9, 15.6, 11.7, 15.3, 14.5, \text{ and } 15.8\%, \text{ respectively. Note that these } \omega \text{ values are approximately uniform.} \]

This means that the criterion of proper statistical approach is met. The uniformity of \( \omega \) has been achieved by transformation of the regression variable from \( v_c \) to \( \ln v_c \).

The sensitivity of the optimized parameters of the formulas is documented in Fig. 6 (on the right), which shows the dependence of \( \omega \) [of the errors of formula (4) compared to the ACI-445F database] on various parameters when only one parameter varied (the others being fixed).

In the proposed formula (4), like in the current formula (8), \( \iota \) is taken as proportional to \( (f'_c)^{1/2} \). Varying the exponent of \( f'_c \) between 1/3 and 2/3 is found to have little effect on the standard error of regression. Exponent 1/2 is adopted because it is in the middle of this range, and because it agrees with the current practice in ACI, historically motivated by Pauw’s (1960) observation that the tensile strength of concrete is approximately proportional to \( \sqrt{f'_c} \).

Eq. (4) gives for the shear force the expression

\[
V_c = v_c b_w d = \mu \rho_w^{3/4} b_w \left( 1 + \frac{d}{a} \right) \sqrt{\frac{f'_c d_0 d}{1 + d/d_0}} \tag{11}
\]

This makes conspicuous that \( V_c \) increases with \( d \), though less than proportionally.

**Other Formulas for Beam Shear and Finite Element Simulations**

The close agreement with the broad-range Northwestern tests and Toronto tests (Fig. 1), and with the record-size Japanese tests (Fig. 2), together with the lack of disagreement with the classical size-effect tests (Fig. 3) of limited size ranges or small maximum sizes, provides strong experimental support for using a formula based on fracture mechanics. Let us now discuss other available formulas. The CEB-FIP (1991) formula (Fig. 2 in Part I), and the power law of the type \( v_c \propto d^{-1/3} \) (see the curves in Figs. 1 and 2),
are purely empirical and thus there is nothing to discuss beyond statistical comparisons with individual size-effect test series. Theoretical arguments, however, have been proposed for three other formulas, and so it needs to be explained why these arguments are not adopted here.

Fractal Characteristics of Fracture and Carpinteri’s Multifractal Scaling Law

The role of the fractal characteristics of fracture has been debated for more than a decade, and recently a detailed critical analysis of the fractal viewpoint has been given by Bažant and Yavari (2005); also, see RILEM (2004). Although some questions remain unresolved, two salient points crystallized:

1. The so-called “multifractal scaling law” (MFSL) proposed by Carpinteri (1994) cannot be applied to failures occurring after large stable crack growth, which is the case of diagonal shear cracks in beams, the main reason being that the stress redistribution due to large cracks and the associated energy release are not taken into account in the existing form of the fractal theories.

2. The MFSL is identical to a special case of a more general size effect formula for crack initiation derived from (non-fractal) fracture mechanics (Bažant and Li 1995, 1996) and refined in Bažant (1998) [also see Bažant (1991, 2002)]. So there is no longer any disagreement about the formula itself, but only about the physical justification of the formula (Bažant and Yavari 2005; RILEM 2004). Because MFSL is intended only for failures at crack initiation (e.g., modulus of rupture), there is no need to dwell on the difficult fractal questions any more. They are not relevant to our problem.

Formula of Japan Society of Civil Engineers

As a consequence of a pioneering proposal of Okamura and Higai made in 1980 (several years before the onset of the energetic fracture-based theory), and of a later recalibration by Niwa et al. (1986, 1987), a power law of the type \( v_c \propto d^{-1/4} \) [proposed on an empirical basis already by Kani (1967)] was adopted for the concrete design code in Japan. In 1980, Okamura and Higai’s proposal was a breakthrough and it appeared logical to underpin it by Weibull’s (1939) statistical theory, the only size effect theory available at that time. That theory (Bažant and Planas 1998) indicates a power law size effect with exponent \(-3/m\) for three-dimensional similarity and \(-2/m\) for two-dimensional similarity, where \( m \) is the empirical Weibull modulus (shape parameter), the value of which is normally determined from the coefficient of variation of tensile strength of many identical specimens. Based on the classical work of Zech and Wittmann (1977), Okamura and Higai assumed that \( m = 12 \). Further assuming three-dimensional similarity to apply, they came up with the exponent \(-3/m = -1/4\).

Recent in-depth studies (Bažant and Novák 2000a,b,c), however, showed that the apparent value of \( m \) for concrete increases markedly with structure size and that, after separation of deterministic nonlocal effects, the correct value of Weibull modulus for concrete is around \( m \approx 24 \). This gives, for three-dimensional similarity, the exponent \(-3/m = -1/8\). Furthermore, it transpired that not only the exponent value but also the assumption of three-dimensional similarity of beam shear failures needs to be revisited. The reason is that material failure at one point in three dimensions does not suffice to ruin the beam. Rather, the beam must fail simultaneously over its whole width, which means that the localization of the failure initiation point could be random only in two dimensions (in the length and depth coordinates, but not in the width coordinate). Therefore, if the Weibull theory were applicable, the size effect exponent would have to be \(-2/m = -1/12\). This means that if Weibull theory applied, the slope in the bi-logarithmic plot would have to be \(-1/12\) rather than \(-1/4\). But such a weak size effect blatantly disagrees with all the test data. Aside from that, fitting of the data for various sizes does not give the same values of Weibull modulus \( m \), as it should, and the measured values do not satisfy Weibull relation between \( m \) and \( \sigma_0 \).

Furthermore, as gradually established, the assumption that the shear failure of reinforced concrete beams is governed by Weibull statistical theory is itself fundamentally unacceptable. That theory is predicated on the hypothesis that the failure occurs as soon as a macroscopic crack initiates from one microscopic flaw, before any significant stress redistribution in the structure is caused by the fracture process. This is true for fatigue fracture of metals or fine-grained ceramics, but not concrete (except perhaps on the scale of large dams, the cross section of which is far larger than the characteristic length \( l_0 \) of the fracture process zone in concrete, which is, typically, 0.5 m). The beam does not fail at the initiation of diagonal shear crack, as would be required by Weibull theory, but only after this crack has propagated in a stable manner through most of the cross section. The growth of the main diagonal shear crack is governed primarily by redistribution of the mean (deterministic) stress field. The randomness of the local strength of concrete at points located far away from the crack path dictated by fracture mechanics cannot significantly deflect that path. Therefore, the mean size effect observed in beam shear failure is caused predominantly by deterministic stress redistribution and the associated energy release prior to failure. The statistical contribution described by Weibull’s statistical theory is negligibly small for the mean response (Bažant and Novák 2000a,b,c; 2001; Bažant 2001, 2002), although it doubtless influences the scatter.

All of this confirms that the Weibull theory cannot be applied to the mean size effect in the beam shear problem. However, if extended to a nonlocal form, it nevertheless appears relevant for characterizing the random scatter (Bažant 2004). Besides, Weibull’s statistical size effect, which is much milder than the energetic size effect associated with stress redistribution, is applicable to the first cracking load. That this size effect exists, and that it is indeed much weaker than the size effect on nominal strength, is confirmed, e.g., by the test data in Fig. 2 of Shioya and Akiyama (1994).

Crack Spacing Hypothesis and Role of Aggregate Interlock

The maximum crack opening width \( w \) is roughly proportional to effective crack spacing \( s_e \), while \( s_e \) is roughly proportional to \( d \). Hence, the deeper the beam, the larger is \( w \) at maximum load. The larger the crack width, the smaller are the cohesive normal and shear stresses \( \sigma_i \) transmitted across the diagonal shear crack. These facts are indisputable. However, it was proposed (Vecchio and Collins 1986; Collins and Mitchell 1991) that the reduction of the cohesive stresses with increasing \( d \) should be the physical source of size effect. This proposition, which they incorporated into their modified compression field theory (MCFT), is untenable.

In MCFT, it is assumed that \( v_c \) decreases with \( w \) as

\[
v_c = \text{constant}/(1 + c_1 w)
\]

where \( c_1 \) = parameter depending on the maximum aggregate size. Assuming \( w \) to be approximately proportional to \( s_e \), one gets the relation

\[
v_c = v_0/(1 + c_0 s_e)
\]

where \( v_0, c_0 \) are...
are certain constants. Considering beams with no stirrups and no horizontal steel bars other than those at the bottom, and assuming $s_x$ to be proportional to $d$ (Collins and Kuchma 1999), one gets, for the MCFT crack spacing theory of size effect, a formula of the type:

$$v_c = \frac{v_0}{1 + d/d_0} \quad (12)$$

where $d_0$ and $v_0$ are parameters independent of structure size. For $d \gg d_0$, this formula approaches the final asymptotic size effect $v_c \propto d^{-1}$ (which is sketched by the dashed lines in Part I, Fig. 2). Although the coefficients of the formula are set up so that the test data would lie in the transitional range between the horizontal and inclined asymptotes in the log-log plot, in which the formula gives a much milder slope than $-1$ and can match some of the existing data, the asymptotic size effect $d^{-1}$ (proposed by Leonardo da Vinci in 1500s and disputed by Galileo in 1638) is theoretically objectionable and in fact thermodynamically impossible (exponent $-1/2$ is the strongest size effect possible). This observation suffices to conclude that the crack spacing theory does not have a sound theoretical basis. But there are more practical objections.

Finite element simulation of propagation of a diagonal shear crack according to the cohesive [or Hillerborg's, fictitious (Hillerborg et al. 1976)] crack model shows that, if the concrete is assumed to have an unlimited compression strength, the load-deflection diagram is always rising, i.e., has no peak; see Fig. 7(e). Therefore, the stresses transmitted across the diagonal shear crack cannot control the maximum load, contrary to what is assumed in the crack spacing theory. Rather, the maximum load must be controlled by the shear-compression fracture of the ligament above the tip of the diagonal shear crack [Fig. 7(e)]. This conclusion is supported by the fact that the compression stresses in the ligament attain the compression strength of concrete (which in turn implies that, in the strut-and-tie model, the failure must be decided by the crushing of the compression strut). So the size effect in beam shear physically represents the size effect of shear-compression fracture (Bažant and Yu 2004).

Further support of the fact that the cohesive stresses across a diagonal shear crack cannot have a significant influence on the maximum load can be obtained from the experimental observations at the University of Toronto. Figs. 7(a and b) shows the major cracks observed near the maximum load, and also the distribution of the measured steel strain [Fig. 7(c)]. Using the steel strain measured at the cross section passing through the tip of the main diagonal shear crack, one can calculate the axial force in the steel, and knowing this force and the bending moment, one can find the precise line of the compression force resultant $F$ in this cross section of the beam tested [Fig. 7(b)]. Because resultant $F$ must also pass through the point of intersection of steel bars and the vertical resultant at the support, one readily can identify the line of $F$, which is drawn in Fig. 7(b). Now note that resultant $F$ passes above the observed main diagonal crack and runs parallel to the top segment of this crack, not intersecting the crack. This implies that shear stresses due to aggregate interlock cannot have an important effect on the maximum load. In the zone just above the tip of the main crack, the compressive stresses $\sigma_y$ parallel to crack must be reaching $f_y'$ (from the magnitude of resultant $F$ one can further estimate that, to transmit average compressive stress $\sigma_y=f_y'$, the localization zone above the tip of the main crack must be about $3d_y$ deep, which coincides with the typical width of localization band). Because the tensile cohesive stresses $\sigma_t$ in the diagonal crack at maximum load must be much less than the tensile strength $f_y'$, and because $f_y' \approx 0.1 f_{t0}$, it follows that $\sigma_y$ must be less than $\sigma_y/10$, and probably much less. So it must be concluded that the contribution of the tensile cohesive stresses $\sigma_t$ in the diagonal crack at maximum load must be negligible compared to the contribution of the compressive stresses $\sigma_y$ parallel to the crack [Fig. 7(d)], and thus cannot control the maximum load (this conclusion puts again in question the physical explanation of size effect by the crack spacing theory).

Also note that, in the strut-and-tie (or truss) model, the straight line connecting the support and the applied load (of slope $a/d$) would normally be assumed as the axis of the imagined "compression strut." The fact that the Toronto measurements imply a much steeper slope of the compression resultant means that the simplistic version of the strut-and-tie model (or truss model) is invalid. While the fracturing strut-and-tie model (Bažant 1997) can explain the size effect in beam shear, uncertainty about the correct slope of the compression strut in beam shear remains to be one obstacle to using the strut-and-tie model as a predictive tool.

**Finite Element Studies and Crack Spacing Hypothesis**

To clarify the role of crack-bridging stresses and aggregate interlock, finite element simulations were conducted. Since regular commercial codes cannot capture cracking localization and frac-
ture propagation, a special code with the crack band model and microplane constitutive law M4 was used (it was the commercial code ATENA, by Červenka Company, Prague, and also an adaptation of FEAP by R. Taylor, Berkeley). All the measured data (including the load-deflection history, evolution of crack pattern, and strain distribution along the steel bars) were faithfully reproduced by the simulations.

Fig. 8(a) shows the results of simulating a beam 1.89 m deep, tested at the University of Toronto (Angelakos 2001). The shear stress components σs (Fig. 8(a)) due to aggregate interlock were calculated, and their vertical resultant, plus stresses due to dowel action of the steel bars, was found to represent only 17% of the total shear force V. The normal cohesive stresses σn in concrete, acting at maximum load at various points of the crack face [Fig. 8(a)], were also calculated, and their vertical resultant was found to represent only 6% of V.

Therefore, the crack-bridging stresses, and thus also the crack spacing, have a negligible effect on the maximum load of a large beam (although they surely have a large effect on the first cracking load). Thus they cannot be the cause of size effect (Bazant and Yu 2004). It follows that, at maximum load, a major part of the shear force in a large beam (77% for the Toronto beam, 1.89 m deep) is transmitted by inclined compressive stresses through the ligament above the tip of the main diagonal crack. Further simulations reveal that the contribution of the crack-bridging stresses decreases as the size increases. In an extremely large beam, 6 m deep, only 9% of the total shear force is transmitted by crack-bridging stresses, while in a small beam, 0.3 m deep, 40% is.

It is interesting to compute how the distribution of the horizontal compressive stress components σs across this ligament evolves as the maximum load is approached [Figs. 8(c and d)]. To avoid any bias due to spurious limitations on the localization of cracking, the finite element sizes h in the ligament region of the small and large beams are chosen to be the same and equal to the effective crack band width w = Gc/Af (w = 0.03 m; Gc = fracture energy of concrete, assumed for these computations as 65 J/m²; and Af = area under the softening stress-strain curve of microplane model M4). At maximum load, the concrete is getting crushed in compression just above the tip of the main crack. Now note that, in the small beam, the σs distribution at maximum load is close to uniform [Fig. 8(c)], while in the large beam it is highly localized, with the point of peak stress fs propagating downward as the load is increasing [Fig. 8(d)]. Consequently, the average stress σt in the ligament is much smaller in the large beam than it is in the small beam. This and the localization of stress profile explain the size effect because, by equilibrium of the end beam segment, v = σt(c/d)(r/a) where r = distance from beam top to the tip of main crack, and c/d and r/a are, in similar beams, almost independent of beam size.

**Conclusions**

Based on both parts I and II, the following conclusions may be drawn.

1. Because the size range of main practical interest lies outside the range of the available test data, the size effect law cannot be set up purely empirically. A realistic theoretical foundation is inevitable.

2. Since concrete is just one of many quasibrittle materials failing due to quasibrittle fracture or softening damage, it would be illogical to expect different laws to govern failure. Laws that are common to all these materials are a more logical choice, especially since the experimental evidence is not completely unambiguous, due to high random scatter and limited size range (dictated by the cost of testing very large beams).

3. The hypothesis that the maximum load in shear failure is controlled by propagation of cohesive fracture or softening damage leads to the same size effect law as established for other quasibrittle materials. The experimental evidence can be matched with this law as closely as one could desire in view of the inevitable experimental scatter.

4. Although a specialized fracture-based model, such as the fracturing truss model, yields a realistic form of the size effect formula and intuitively explains the mechanism of size effect, this model is insufficient for capturing the composite beam action and for predicting the dependence of size effect law coefficients on the shear span, reinforcement ratio, aggregate size, material strength, etc.

5. In the asymptotic situations of infinitely small and infinitely large structures, the analysis of failure and size effect becomes far simpler and clearer than in the practical size range. Therefore, it is appropriate to derive a theoretical formula by asymptotic matching, a technique that "interpolates" between the known asymptotic behaviors at the opposite infinities.

6. Assuming the failure load to be controlled by cohesive fracture parameters (material strength σo and fracture energy Gc), and exploiting the known first two asymptotic terms of the large- and small-size asymptotic expansions of the cohesive crack model (or the nonlocal damage models), one can easily deduce a simple transitional size effect law by means of dimensional analysis.


8. The size effect formula developed agrees with all of the existing test series in which the beam depth was varied significantly, the number of which currently is 11. Especially, this
size effect law agrees with the tests conducted at Northwestern University and at University of Toronto, which represent the only data with a broad size range and almost perfect geometrical scaling.

9. Most importantly, the Northwestern and Toronto test series, and the Japanese tests of Shioya et al. (which are the only scaled tests of a significant size range) confirm that the large-size asymptotic slope of size effect is \(-1/2\). This in turn proves that the explanation of failure lies in fracture mechanics.

10. Finite element simulations of the data measured on large beams (based on crack band model and microplane model M4) show that, in large beams (more than about 1 m deep), the crack-bridging stresses normal to crack face, as well as the tangential crack-bridging stresses (due to aggregate interlock), contribute only a small part of shear force capacity (although they are important at the first cracking load). The larger the beam, the smaller is this contribution. The same applies to dowel action. On the other hand, for small beams (less than about 0.3 m deep), the crack bridging stresses, aggregate interlock and dowel action are important, contributing about 40% of $V$.

11. A major part of the shear force at maximum load is, in large beams, transmitted by inclined compressive stresses through a small zone above the tip of the main diagonal crack. The effective depth of this zone, stressed to compressive strength limit, is roughly two to four maximum aggregate sizes. The profile of the compressive stresses over the ligament above the main crack tip is close to uniform in small beams but highly localized in large beams. This provides the most plausible explanation of the size effect.

12. The coefficients of the size effect law are calibrated by optimum fitting of the ACI-445F database (which includes 398 data) and also a modified database (consisting of 341 data for which the aggregate size was reported, and enhanced by large-scale Japanese tests and reduced-scale tests). However, since these large databases are inevitably contaminated by random variation of factors other than the size, the differences in the coefficients of variation of errors are, in comparison to other proposed formulas, relatively insignificant.

13. The size effect law in Eq. (4) provides a good and simple representation of the trend of both databases. The coefficient of variation of regression errors, which is about 15%, allows determining the 5% probability cut off which is suitable for a design code formula. The proposed formula also agrees with test data on the effects of shear span, reinforcement ratio and aggregate size.

14. Comparisons with the required asymptotic behavior and with the test data pertinent to the size effect in beam shear reveal serious contradictions with the old formulations [including the JSCE (1991) formula based on Weibull statistical theory, the crack-spacing enhancement of MCFT, the empirical CEB-FIP formula, and the application of MFSL to diagonal shear failure]. They also reveal significant disagreement with the empirical power law of exponent \(-1/3\).

### Acknowledgments

Financial support by the Infrastructure Technology Institute of Northwestern University is gratefully acknowledged. Professor M. T. Kazemi of Sharif University of Technology, Tehran, is thanked for valuable comments.

### Appendix. Questions of Statistical Evaluation and Bias

Although formula identification from beam shear data is a problem of statistical regression, in one recent investigation it has been tried to reduce comparisons of various proposed formulas for beam shear to elementary population (ensemble) statistics of the ratios $\gamma_i$ of the measured and calculated values. In that investigation, the coefficient of variation of errors has been defined as follows:

$$\text{CoV}^* = \frac{1}{\bar{\gamma}} \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (\gamma_i - \bar{\gamma})^2} \quad \text{with} \quad \bar{\gamma} = \frac{1}{n} \sum_{i=1}^{n} \gamma_i, \quad \gamma_i = \frac{\hat{\gamma}_i}{u_i}$$

(13)

where $\gamma_i$ are called the “model safety factors.” However, such an approach is not correct, for four reasons. First, if the parameter identification is not based on least-square regression, the results cannot be unbiased, i.e., the resulting mean and variance of formula parameters cannot be the mean and variance of the statistical distribution of these parameters. Second, if the minimized expression is not a sum of squares of errors, the tangential linearizations made in any data fitting algorithm can lead to numerical instability. Third, fitting of the ratio $\hat{\gamma}_i/u_i$ implies weighting of the data $\gamma_i$ as a function of the unknown values of $u_i$ to be solved, which is inappropriate. Fourth, if all $u_i$ in Eq. (13) are replaced by $k_0 u_i$, $k$ being any constant between $-\infty$ and $\infty$, then $\gamma_i$ and $\bar{\gamma}$ are replaced by $\gamma_i/k$ and $\bar{\gamma}/k$, and so it is found that the value of CoV does not change. Such a definition of the coefficient of variation, which is insensitive to multiplying the formula for $u_i$ by any number, makes no sense at all. Obviously, minimization of $(\text{CoV}^*)^2$ cannot be used to calibrate formula parameters (Bažant 2004; Yu 2004).

As a workable alternative to the use of ratios $\gamma_i = \hat{\gamma}_i/u_i$, it has been tried to identify formula parameters by minimizing not $(\text{CoV}^*)^2$ but the sum $\Phi = \sum (\gamma_i - 1)^2$. This sum may be rewritten as:

$$\Phi = \sum_{i=1}^{n} \left( \frac{\hat{\gamma}_i}{u_i} - 1 \right)^2 = \sum_{i=1}^{n} w_i (\hat{\gamma}_i - u_i)^2$$

(14)

where $y_i = 1/u_i$, $\hat{\gamma}_i = 1/\hat{\gamma}_i$ and $w_i = (\hat{\gamma}_i)^2$. So, the minimization of $\Phi$ represents simply a weighted least-square regression of data $\hat{\gamma}_i$, minimizing the sum of squared errors in $y_i$ with weights $w_i$, proportional to the squares of the measured values $\hat{\gamma}_i$.

Aside from a misguided desire to avoid regression statistics, the motivation for use of the ratios $\gamma_i = \hat{\gamma}_i/u_i$ has been to increase the weight of smaller $u_i$ values. But there is a problem with this motivation. As is well known (Lehmann 1959; Plackett 1960; Ang and Tang 1976; Beck and Arnold 1977; Draper and Smith 1981; Mandel 1984; Fox 1997), to minimize bias, the regression should be conducted in such variables in which the variance $\text{Var}(u_i | d)$, in our case, is as uniform as possible, and nonuniform weights should not be used unless the variance varies by an order of magnitude or more (i.e., if the data are heteroskedastic). For beam shear size effect data (as well as most size effect data), the variance appears to be the least nonuniform in the plots of $\log \nu_c$ versus $\log d$, and so these are the preferable coordinates for statistical regression.

The implication of unjustified shear strength dependence of the weights in Eq. (13) is that the minimization of Eq. (14) is not unbiased, i.e., the estimates of the mean and the standard deviation (or CoV) of the optimized parameters are not the actual mean
and standard deviation of the statistical distribution of these parameters (in other words, are not the "maximum likelihood" estimates). One consequence is that the mean minus 1.65 times standard deviation does not give the correct (unbiased) value for the 5th percentile, needed for setting up the design formula, and generally all the statistical estimates are not correct (note that data weighting to compensate for the crowding of data points into the small-size range is an entirely different matter). On the other hand, the absence of shear-strength dependent weighting from the least-square fitting of ln $v_s$ ensures that the optimized parameters are unbiased. Because $|d[ln(u_{s})]_f|=|d(u_s)_{f}|$, the transformation of scale from linear to logarithmic has a similar effect as the weighting of the data in proportion to $1/(u_s)^2$, but without causing any statistical bias.

References


strength of reinforced concrete beams without web reinforcement.”
Proc., Japanese Society of Civil Engineers, Vol. 300, Japanese Society
of Civil Engineers, Tokyo, 131–141.
Gloucestershire, U.K.
“Shear database for reinforced concrete members without shear
reinforcement.” ACI Struct. J., 100(2), 240–249; with discussions by
Zürich, 33(7), 59–61.
in reinforced concrete structures.” Size Effect in Concrete Structures,
Proc., Japan Concrete Institute Int. Workshop, Sendai, H. Mihashi, H.
theory for reinforced concrete elements subjected to shear.” ACI J.,
83(2), 219–231.
members without shear reinforcement’ by K.-H. Reineck, D.A.
assessment of the containment of a PWR, Part II. Probabilistic ap-
proach to describe the behavior of materials.” Trans. 4th Int. Conf. on
Structural Mechanics in Reactor Technology, Vol. H, Sec. J1/11,
T. A. Jaeger and B. A. Boley, eds., European Communities, Brussels,
Belgium, 1–14.