On Critiques of Energy-Based Size Effect Law for Beam Shear and Punching

Part I. On Muttoni and Fernández Ruiz’s Critique
Part II. On Bentz and Foster’s Critique

Abdullah Dönmez, Christian Carloni, Gianluca Cusatis
and Zdeněk P. Bažant

Report No. 19-12/mfbf

Department of Civil and Environmental Engineering
McCormick School of Engineering and Applied Science
Northwestern University
Evanston, Illinois 60208, USA

December 23, 2019
I. On Muttoni and Fernández Ruiz’s Critique of Energy-Based Size Effect Law for Beam Shear and Punching

Abdullah Dönmez¹, Christian Carloni², Gianluca Cusatis³, and Zdeněk P. Bažant⁴

1 Postdoctoral Associate, Dept. of Civil and Env. Engrg., Northwestern University, Evanston, IL
2 Chair, ACI Committee 446, Fracture Mechanics, Professor, Dept. of Civil Engrg., Case Western Reserve University, Cleveland, OH
3 Past Chair, ACI Committee 446, Fracture Mechanics, Professor, Dept. of Civil and Env. Engrg., Northwestern University, Evanston, IL
4 McCormick Institute Professor and W.P. Murphy Professor of Civil and Mechanical Engineering and Materials Science, Northwestern University, 2145 Sheridan Road, CEE/A135, Evanston, Illinois 60208; corresponding author, z-bazant@northwestern.edu.

Abstract: In a recent article in Structural Concrete (May 2019, pp. 1464-1480), Muttoni and Fernández Ruiz discussed critically the previously published (March 2019, pp. 1451-1463) Dönmez and Bažant’s justification of the energy-based size effect factor that has recently been adopted for ACI Standard 318-2019, to govern the shear strength of beams without stirrups, punching shear strength and the strength of compression struts of the strut-and-tie model. This rebuttal demonstrates that Muttoni and Fernández Ruiz’s criticism of the energy-based size effect law, and implicitly of the new ACI size effect factor, is erroneous and that the arguments for their own Critical Shear Crack Theory (CSCT) proposed for the fib Model Code and the Eurocode are baseless. In particular: a) A rigid-body kinematics of the main crack opening, with a center of rotation at the crack tip, is invalid. Considering relative rotation of line segments within regions of relatively small deformation on opposite crack sides, the center of relative rotation during failure is found to lie far very away from the beam, particularly if realistic constitutive model (such as M7), capturing the transverse expansion of the fracture process zone due to crack-parallel compression, is used. A large pre-peak rotation can be obtained based on points in the crack-tip zone, but this zone is far from rigid, deforming very nonuniformly. b) A negligible role of aggregate interlock approach to maximum load can be demonstrated by filtered subsets of a large experimental database for beams with both small and large maximum aggregate sizes, in which the variation of steel ratio, shear span, and aggregate size at increasing beams depth is filtered out. c) A negligible interlock role is also demonstrated by comparison with high-strength concrete beams in which the cracks cut through the aggregates and are so smooth that no aggregate interlock can develop. d) Believing in high aggregate lock contribution at maximum load is tantamount to denying the applicability of the strut-and-tie model, whose basic hypothesis is that, at maximum load, no stress is transferred across the shear crack running parallel to the compression strut. d) It is incorrect to write the constitutive equation as a scalar relation between the maximum principal stress and strain. Their vectors are generally non-parallel and often far from parallel, especially when shear follows normal strain. Their vectors may rotate against each other as well as against the material, especially when shear follows normal strain. And triaxial stress effects are missed. e) As clear from fracture mechanics, whether or not the structure is statically redundant can make no difference for the form and the double-log scale asymptotes of the size effect law for beam shear and, especially, redundancy of supports cannot change the large-size asymptotic slope from -1/2 to -1/3 or -1/2.5, or -1.0. The denials of the critique of the six hypotheses shown to underlie the CSCT (Dönmez, Bažant, 2019) are baseless.

Keywords: Concrete fracture, scaling of failure, energy methods, aggregate interlock, shear strength of beams, punching strength.
Recently, Muttoni and Fernández Ruiz published in Structural Concrete 2019 [1] a lengthy critique of the energy-based size effect factor which has recently been adopted for the ACI-318 code provisions for the shear strength of reinforced concrete (RC) beams without stirrups, the punching shear strength, and the strength of the compression struts of the strut-and-tie model. This size effect factor and its energy basis were previously justified in Dönmez and Bažant’s 2019 article in Structural Concrete [2], in which a comparison was also made with Muttoni and Fernández Ruiz’ Critical Shear Crack Theory (CSCT) [1] proposed for the Model Code and Eurocode. Muttoni and Fernández Ruiz’s criticism is rebutted point by point, as follows.

1) P.1469: “On the basis of the shape and kinematics assumed (with a location of the center of rotations at the tip of the crack and considering two rigid bodies separated by the critical shear crack, see Figure 6a,b and refer to the experimental results in Figure 4a, and by assuming that the relative rotation can be described as a function of the strain in the flexural reinforcement (at the location where the shear crack intercepts the flexural reinforcement, refer to Figure 6c and to Equation (1)), the complete profile of opening and sliding along the height of the crack can be calculated (Figure 6c–e). Integration of the stresses over the simplified critical shear crack (consistently with the methodology followed for the interpretation of DIC measurements), allows to eventually calculate the various shear-transfer contributions.”

To prove that this is incorrect, the microplane damage constitutive model M7 is calibrated by size effect tests of Walraven [3] (Fig. 1.1a). The finite element (FE) results obtained with the crack band model then provide, for the state at maximum load ($P_{max}$), the distribution of principal stress $\sigma_I$ along the crack line (tensile $\sigma_I$) and along its extension to the top face (compressive $\sigma_I$), shown in Fig. 1.1b. The aggregate interlock stresses $\sigma_I$ at $P_{max}$ are seen to be negligible along the crack and are dwarfed by the large compressive stresses $\sigma_I$ above the crack tip (max 0.09 MPa in tension on the crack versus -28.82 MPa in compression in front of the crack tip, while the uniaxial compression strength of the concrete is 34.4 MPa).

This is in glaring contrast to Muttoni and Ruiz’s alleged inference from their rigid-body kinematics, which is, for comparison, reproduced from [1] in Fig. 1.1c. Their alleged spike in crack-bridging stresses (attributed to aggregate interlock), which is supposed to represent the resultant of normal and shear components and is shown in Fig. 1.1c to rise up to 3 MPa and 5 MPa near the crack tip, is unrealistic and contradicts FE analysis. The calculated stress distribution in Fig. 1.1b, in fact, invalidates the stress assumption of the CSCT schematized in Fig 1.1c. In this light, it is perplexing that Muttoni and Fernández Ruiz in their Fig. 9 [1] allege the $V_a$ and $V_I$ on the crack to contribute together about 85% of the total shear strength in large beams, and almost 95% in small ones. The correct percentage at $P_{max}$ is about 30% for small beams and about 5% for large ones [2, 4, 5].

In [1] (Fig. 6c and 9b, schematically reproduced here in Fig. 1.1c), Muttoni and Fernández Ruiz introduce into their CSCT the hypothesis of a rigid-body rotation about the crack tip. The rotation angle about the alleged center of rotation at the crack tip is alleged to be controlled by the strain, $\varepsilon$, of longitudinal reinforcement at the crossing of crack mouth. To check the alleged rigid-body rotation, two initially parallel line pairs, $AB$-$CD$ and $EF$-$GH$, are anchored in zones of least non-uniform deformations (i.e., nearest to behaving as rigid), located on both sides of the diagonal crack, are shown in Fig. 1.2. The FE analysis gives, for these line segments and for $P_{max}$, the inclination angles, $\alpha$, compared to the initial state at $P = 0$ and the state at 0.95 $P_{max}$. The corresponding relative rotations, $\Delta \alpha$, of these line segment pairs are seen to be negligible. Indeed, the distances of the centers of relative rotation from the crack tip, given in Fig. 1.2, are on the order
of 100 m or more. Placing points, A, B, C, D in the zones where the FE fields exhibit very nonuniform strain, e.g., near the crack faces and close to the tip, may give a much closer location of the center of rotation. But this would not justify the “picture mechanics” of rigid-body movements. So, before \( P_{\max} \), there no overall rigid-body rotation, contrary to CSCT. Before \( P_{\max} \) (unlike after), the masses on the crack flanks move apart in almost translatory motion that is restrained by the longitudinal reinforcement at the crack mouth and is driven by transverse expansion of the compression crushing zone above the crack tip (such expansion is missed or underestimated in calculations with the common simple constitutive models for concrete, as in [6] but is captured realistically by the microplane model).

The foregoing FE results document a fundamental error in the hypothesis of large pre-peak rigid-body rotation, made in CSCT. In [1] it is claimed that this hypothesis agrees with the digital image correlation (DIC) measurements, in which the alleged rigid-body rotation was inferred from DIC measurement of crack width variation along the crack, as reported in Figs. 3 and 4 of [1]. But such inference is ambiguous since the strain field in the material near the crack is highly non-uniform and variable, especially near the crack tip, and extensive microcracking near the crack faces enhances substantially the effective opening and shear displacements (as distinct from visible crack). Muttoni and Fernández Ruiz [1] further claim that the cross-crack normal and shear stresses evaluated from the observed DIC field support their hypotheses. However, their method of calculating the stress from DIC data described in [7], is dubious.

Above Eq. 6 of [7], Cavagnis, Fernández Ruiz and Muttoni write: “the principal stress directions are assumed to be parallel to the principal strain directions and the principal stresses are directly computed from the principal strains.” This is in general incorrect. While in most of the compression strut (which is what matters for \( P_{\max} \)), the principal stress directions of stress and strain are indeed almost parallel, a scalar stress-strain relation, such as Eq. 6 of [7] based on uniaxial tests, is not correct. This is because each principal stress depends on both principal strains, and the transverse principal strain often has a strong confining or loosening effect. Furthermore, even if the principal directions are parallel, they rotate against the material [8] (Fig. 1.1e,f,g), which is not seen in uniaxial test machines but does not allow expressing here the stress-strain relation in principal tensor components only (an exception is the approximate “rotating crack model,” but that has been verified only for post-peak distributed cracking).

In the narrow zone at the end of the compression “strut” near the top face of beam, which triggers the failure of the “strut” and thus governs the shear capacity of the beam, the principal stress and strain directions are non-parallel and rotate against each other (Fig. 1.1g). This is especially marked when normal or shear stress is softening. For such behavior, in this most important zone controlling \( P_{\max} \), the scalar one-dimensional equation, such as Eq. 6 in [7], makes no sense. Rather, a general nonlinear triaxial constitutive law is required.

Further note that the stress-strain relations, including the scalar ones used in [7] and [1], depend on the base length of strain measurement (within which the strain may or may not localize), and, if no localization limiter is used, also on the structure size, as is the case in [7]. Moreover, they do not account for the layer of microcracking along the visible opened cracks. Also note that Eq. 6 in [7], an old stress-strain relation for uniaxial stress, ignores triaxiality and cannot match the crack softening tails observed in the recent fracture tests [9] of 140 beams with various sizes and notch depths.
Similar limitations apply to the biaxial strength envelope in Fig. 9b of [7] (proposed in 1969). Its shape on the compression side is unrealistic, and the envelope changes depending on the strain history. The stress-displacement relation for the crack slip in Eqs. 1 and 2 of [7] (proposed in 1981) is based on the assumption of plastic deformation of the asperities which, in reality, break off. It does not account for the crack slip with softening and ignores interaction with the softening stress-displacement response of the crack, which is not reproduced in Fig. 7e,f of [7]. The fixed compression-softening law (proposed in 1986, Eq. 9 of [7]) is based on uniaxial tests and cannot be applied to triaxial stress states, particularly not when the shear stress varies simultaneously. Simultaneous shear makes the normal stress softening much steeper. Likewise, a simultaneous increase in the normal cross-crack displacement causes softening in shear. The state-of-the-art moved way beyond these simplistic scalar stress-strain relations.

In the modern computational approach to evaluating DIC observations, evaluation of the stress field from DIC data requires an FE element program with a realistic triaxial damage constitutive model. This model must be tensorially invariant under coordinate rotations (which is not true for the models in [7]) and must use a localization limiter eliminating spurious mesh sensitivity of softening damage, compatible with the fracture energy or material characteristic length of the given concrete. The FE program is then calibrated by fitting the DIC displacement field. If the fit is satisfactory, it is then used to calculate the stress field corresponding to the DIC measurements. This is the approach used for the comparisons in [2], which were in [1] subjected to baseless criticism.

For the shear failure of RC beam, two aspects of the tensorial material constitutive relation are essential: 1) Capturing the large transverse expansion of damaging material under softening compression, and 2) the vertex effect, which means that the shear stiffness for shear strain superposed on a material softening under normal stress is far lower than the undamaged shear stiffness as if several loading surfaces in the stress space were intersecting at each current state point and thus formed a vertex on the loading surface (note that in plasticity, a stress increment parallel to the current loading surface in the stress space is always elastic). These triaxial phenomena, especially the combination of shear with cross-crack normal stress softening, intensify the reduction of crack-bridging (or interlock) stress. They also cause transverse expansion at the crack tip, which is one of the reasons for the center of relative rotation being far away. The scalar stress-strain relations and the strength expression used in [7] cannot reproduce these phenomena at all. Neither can the common triaxial Mohr-Coulomb type or Drucker-Prager type nonlinear triaxial stress-strain relations. The microplane model can.

2) P.1471: “Since the shear strength depends upon the crack opening, the model accounts directly for size and reinforcement strain effects”; and on P.1472: “The reference value of the opening of the critical shear crack is assumed to be proportional to the product of a reference strain times the effective depth of the member.”

The arguments of CSCT based on rigid-body kinematics and on the reinforcement strain run into self-contradictions. As stated above, the basis of CSCT is the equation for crack width \( w \), assumed to be \( w = \varepsilon d \) (see [6] in 1991 and subsequent paper [1, 10, 11]); here \( \varepsilon \) is called the reference strain (calculated from the steel reinforcement strain obtained from the beam bending theory), and \( d \) is the effective depth of the beam. The crack opening profile in Fig. 6c and \( d \) in [1] (or in Fig. 1.1c here), and the corresponding stress profiles are shown in Fig. 9b in [1] (or Fig. 1.1c here).
This figure and the strain assumption, unfortunately, admit self-conflicting interpretations depending on the change of the alleged rotation angle with $d$, even if a rotation centered at the crack tip is assumed. The ratio of the crack tip distance from the top face of the beam to depth $d$ does not change much with $d$. Then, since the crack mouth opening is assumed to be roughly proportional to $d$, it follows that the angle or their alleged rigid-body crack opening should be the same for all $d$ (Figs. 7b, d, and Fig. 6b in [1], or Fig. 1.1c here). But then the ratio of the distance from the crack tip of the centroid of the sketched stress profile to the beam depth $d$ would decrease in proportion to $1/d$ (Fig 1.1c, d). Such proportionality would yield a size effect asymptote of slope -1, which is not only thermodynamically impossible (Eqs. 4 – 7 in [4]) but also absurd since it would mean that, for large $d$, a doubling of $d$ would not increase $P_{max}$. This consequence exemplifies the disparity among the various ways the CSCT has been derived from the assumed kinematics. Also, this consequence is inconsistent with Eqs. 5, 8, or 9 in [1].

3) P.1466: “Failure can be triggered by merging of secondary cracks in the critical shear crack (in a stable or unstable manner) or by the development of a new inclined crack due to local engagement of aggregate interlock due to crack sliding.”

The secondary cracks merge, or localize, into the critical shear crack already before reaching $P_{max}$. The aggregate interlock is insignificant at $P_{max}$. It plays a major role only at shear crack initiation which occurs long before reaching $P_{max}$. This fact is documented by the stress field and the energy release zone at $P_{max}$ presented in [2], which shows that the failure at $P_{max}$ is triggered by compression-shear crushing of concrete above the tip of the diagonal shear crack. This trigger agrees with the failure mechanism implied by the energy-based size effect law (and thus also by the ACI size effect factor), as well as by the strut-and-tie model, but conflicts with CSCT.

The CSCT assumption that aggregate interlock stresses make a major contribution to $P_{max}$ is also rendered dubious by considering high-strength concrete (HSC), in which the crack propagates through the aggregate pieces rather than around them. The cracks in HSC are much smoother and thus can hardly provide any aggregate interlock [12, 13]. According to Muttoni and Fernández Ruiz, the shear strength would then have to be very small. Yet the HSC beams have a much higher shear strength than the normal strength beams, despite their much weaker aggregate interlock. This fact cannot be reconciled with the CSCT.

4) P.1474: “Members with higher depth present larger crack widths in the unreinforced zone for the same level of deformation in the reinforcement, thus showing lower unitary strengths (size effect).”

As a description of the source of the size effect, this is incorrect. It is true that, in a beam scaled up in size, the crack opening may become wider in some regions, but this is not the cause of size effect. Rather, the cause is that the energy released from the structure (especially from the “compression strut”) grows quadratically with structure size while the energy dissipated by FPZ grows linearly. Consequently, the material strength gets mobilized, in a larger beam, only within a shorter segment of the compression-shear crushing band above the diagonal shear crack tip (which is at the root of the “compression strut” and transfers load into it). This fact, which provides an intuitive explanation of size effect, is amply confirmed by FE simulations.

5) P.1474-1475: “In addition to the consistency for linear responses (typical behavior of statically determined beams and one-way slabs), the model also provides a suitable response when a nonlinear behavior can be expected. This topic has been extensively investigated by Fernández Ruiz and Muttoni for members with structural continuity and potential redistributions of internal
forces during the loading process (which leads to a variation of the ratio between the acting shear force and the bending moment at the control section). In these cases, the calculated size effect with the CSCT is milder (in agreement to the theoretical response of Fracture Mechanics for these cases) and can be determined for each specific case leading to typical asymptotic slopes in double-log scale of approximately -1/3.”

This argument is erroneous. The “linear responses” are not the “typical behavior of statically determinate beams and one-way slabs,” not more so than they are for redundant beams. The static determinacy or indeterminacy has nothing to do with the linearity of response. Muttoni and Fernández Ruiz mix up this concept with nonproportional evolution of internal forces, i.e., the bending moment $M$ and shear force $V$. If the crack is not growing (and if the material follows a linear stress-strain relation), the $M$ and $V$ evolve during loading proportionally, but if the crack grows or if, in some part of structure, the linear stress-strain relation is not followed, or both, $M$ and $V$ evolve nonproportionally, regardless of static determinacy or redundancy.

Muttoni and Fernández Ruiz’s [1,10] arguments about static determinacy or redundancy of “one-way” or “two-way” slabs are a misunderstanding of the mechanics of structural systems. “One-way” or “two-way” has nothing to do with static determinacy or redundancy. In [1], and in more detail in [10], it is misunderstood that cracks propagating across the thickness of a beam or slab are always “statically indeterminate”, or “redundant”, regardless of whether or not the beam or slab supports are statically determinate. The reason is that the differential equations of equilibrium, which are part of the two-dimensional elasticity solution of transverse crack in a beam or plate, must be always be complemented by compatibility conditions, which in this case are the strain compatibility conditions of continuum mechanics (rather than compatibility conditions of displacement and rotations of beam segments). The strain compatibility condition is automatically implied in a discrete manner in a two-dimensional FE analysis of the crack in the beam.

Furthermore, the static determinacy of supports or internal connections in continuous beams or frames makes no difference for the energy derivations of the energy-based size effect law (and thus also new ACI size effect factor [14]), and particularly not for the most general energy derivation in [4], which is based only on dimensionless variables and on the existence of a material characteristic length related to the aggregate size and other heterogeneity characteristics.

Adding redundant supports changes, of course, the energy release function of LEFM, and thus affects the maximum loads and the transitional size $d_0$. But the energy-based size effect law (with the new ACI size effect factor [14]) remains valid, and its final asymptotic slope remains to be -1/2. The illusion of an asymptotic slope milder than -1/2 might have arisen from experimental size effect curves for a larger transitional size $d_0$ (25 cm in ACI [14]), which end with a smaller slope. A larger $d_0$ causes, for the same depth $d$, the size effect slope to become smaller at the same $d$. So, the “calculated size effect with the CSCT is” not “milder”. The asymptotic slope remains and must remain, -1/2. Proposing “typical asymptotic slopes in a double-log scale of approximately -1/3” (or, in [11, 15] -1, -1/2.5) is baseless.

6) P.1475: “By considering a realistic nonlinear response of the structure (assuming, in any case, the member to be cracked due to external actions or eigenstresses), the asymptotic slope is again consistently milder (approximately between -1/2.5 and -1/3 in double-log scale as acknowledged by other researchers.”

This is baseless and the response to the foregoing item 4 also applies here. Further note that Muttoni and Fernández Ruiz keep using for their theory the same name, CSCT, even though
its form has been changing from paper to paper. Including the most recent equation (Eq. 9 in [1]), the CSCT equations for the size effect in shear of RC structures (beams and slabs), as presented so far, feature, in the log-log plot [of log(strength) versus log(size)], four different asymptotic slopes, which are \(-1\), \(-1/2\), \(-1/2.5\) and \(-1/3\). The first three of these slopes are implied by Eq. (LoA I) in [11], Eqs. 5, 6 in [1], and Eqs. 5, 8 in [11], respectively. The fourth slope is implied by Eq. 9 in [1]. However, an asymptotic slope in the log-log plot other than \(-1/2\) is thermodynamically impossible (Eqs. 4 – 7 in [4]). Moreover, in [1, 10, 15], the only thermodynamically correct asymptotic slope of \(-1/2\) arises in Eq. 5 of Muttoni and Fernández Ruiz as a chancy consequence of an empirical ‘derivation’ of CSCT, ignoring fracture mechanics.

In the size effect section of [11], the CSCT size effect on punching shear strength is assumed to be a function of the span alone, although what makes sense for size effect is only geometrical scaling, i.e., a simultaneous proportional change of thickness and span, in which the dimensionless ratio \(r_s/d\) in Eq. 8 of [11] is kept constant. Extending the length at constant depth represents a combination of size and shape effects, the latter being more complicated. Such manipulations then somehow resulted in an absurd asymptotic slope of \(-1\).

Moreover, contrary to Muttoni and Fernández Ruiz’s allegation in [1], Dönmez and Bažant in [16] do not acknowledge a slope milder than \(-1/2\). It is true that in the computationally filtered subsets of the ACI database comprising about 440 tests [17], the end slope of the curve of the interval means of the nominal shear strength in [16] has in the log-log plot a slope milder than \(-1/2\). But this is because the range of existing data is not broad enough. The practicable data range is insufficient to reveal the asymptotic slope. The asymptotic slope can be closely and easily approached only by tests of micro-concrete with reduced aggregate size \(d_a\), as presented in 1991 in [18] for \(d_a = 4.8\) mm (see also [19, 20, 21]), which confirmed slope \(-1/2\). Anyway, the trend of the filtered data subsets in [16] fits perfectly the energy-based size effect law (these data, from [18], are not found in the ACI-445 database because the committee, unfortunately, decided to make a cutoff at \(d_a = 5\) mm).

7) P.1475: “However, even for this latter case, the simplified closed-form expression for design (Equation (9)) provides a reasonable estimate of the complete CSCT size-effect law (Equation (8)) for the range of practical interest (see two representative cases in Figure 11b, with an effective depth varying between 55 and 5,500 mm).”

In Fig. 11, Muttoni and Fernández Ruiz show the size effect curves of their Eqs. 8 and 9, but only for a selected size range. This, and the fact that the final asymptotic slopes of Eqs. 8 and 9, \(-1/2\) and \(-1/3\), are different and mutually conflicting makes these curves suspect. In addition, parameter \(\beta\) in Eqs. 7 and 8 must be dimensionless but is not, and it is mathematically impossible to derive Eq. 9 from Eq. 8 or 5.

8) Appendix A, in which the authors deny the critique of the (considered by the writers unrealistic and unfounded) hypotheses implied in CSCT. To forgo reproducing here at length the detailed logical arguments of Dönmez and Bažant in [2], it is suggested to an interested reader to look at [2] before reading the brief rebuttals that follow. 

Hypothesis 1. Considering the beam segments on the side of the dominant crack to rotate as rigid bodies (Fig. 6c, d in [1]) around an assumed crack tip is unrealistic. The material deforms (Fig. 1.1).—Hypothesis 2. The writers should not be blamed again in [1] for not understanding Eq. A1 because it is unfounded. Its only purpose is to give a semblance of mathematical derivation. Replacing Eq. A1 by other no less plausible softening curves based on Fig. 9 [1] leads to rather different size effects. — Hypothesis 3. The
refutation offered does not remove the arbitrariness of this hypothesis. — *Hypothesis 4*. The attempted refutation obfuscates but does not justify why the size effect should be governed by the longitudinal strain calculated under the assumption of plane cross-sections for the depth of $0.6d$ and for the distance of $d/2$ from the support. It is a fiction aimed to give an appearance of logical derivation. — *Hypothesis 5*. Again, the lengthy verbiage does not justify why a linear elastic theory of beam bending with plane cross-section assumption should be used, despite out-of-plane cross-section warping and three-dimensional inelastic deformation field. — *Hypothesis 6*. Despite the extensive, albeit loose, verbiage in [1], it must be reasserted that the linear elastic beam bending theory preserving the planarity of cross-sections is blatantly insufficient to handle the triggering of shear failure, which distorts the planarity of cross-sections.

**References:**


[14] ACI Committee 318, “Building Code Requirements for Structural Concrete (ACI 318-2019) and Commentary,” American Concrete Institute, Farmington Hills, MI (2019); articles 22.5.1.3, 22.6.5.2 and 23.4.4.1.


**Fig. 1.1:** a) Calibration of M7 and SEL fit according to the data test [4], b) the principal stress profiles along the crack, c) the profiles of displacement and stress values along the diagonal crack according to CSCT (Figs 6c and 9b in [1]) and d) extrapolation of the crack width assumption of CSCT for large beams (3d).
Fig. 1.2: Parallel line pairs, located above and below the diagonal crack and corresponding angles and relative rotation ($\Delta \alpha$) calculations for the initial ($P=0$), $P_{\text{max}}$ and $\Delta P=P_{\text{max}}-0.95P_{\text{max}}$ cases. The calculated center of rotation locations based on the translated coordinate systems (note that if points A, B, C, D were placed into the near-tip zone, one could get a large rotation, but that would not be a rigid-body rotation because this zone is deforming significantly and nonuniformly).
II. On Bentz and Foster’s Criticism of Energy Analysis of Size Effect in Shear Strength of RC Beams and Slabs

Abdullah Dönmez¹, Christian Carloni², Gianluca Cusatis³ and Zdeněk P. Bažant⁴

¹ Postdoctoral Associate, Dept. of Civil and Env. Engrg., Northwestern University, Evanston, IL
² Chair, ACI Committee 446, Fracture Mechanics, Professor, Dept. of Civil Engrg., Case Western Reserve University, Cleveland, OH
³ Past Chair, ACI Committee 446, Fracture Mechanics, Professor, Dept. of Civil and Env. Engrg., Northwestern University, Evanston, IL
⁴ McCormick Institute Professor and W.P. Murphy Professor of Civil and Mechanical Engineering and Materials Science, Northwestern University, 2145 Sheridan Road, CEE/A135, Evanston, Illinois 60208; corresponding author, z-bazant@northwestern.edu.

Abstract: Bentz and Foster published in Structural Concrete (June 2019, pp. 1481-1489) severe criticism of the energy analysis and related strength calculations underlying the energy-based size effect law (SEL) recently incorporated into the new ACI Standard 318-2019 for shear and punching strength, and for strength of compression struts in the strut-and-tie model. This equation and energy analysis was previously justified in Dönmez and Bažant’s article in the same journal (March 2019, pp. 1451-1463), in which it was also critically contrasted with the Critical Shear Crack Theory (CSCT) and Modified Compression Field Theory (MCFT), theories that are now proposed for, or already used in, the Eurocode and the fib Model Code. This rebuttal shows all the critical points of Bentz and Foster to be incorrect. In particular: a) The strength of materials methods and energy methods must give, as a matter of principle, equivalent results. b) The aggregate interlock contributes negligibly to the maximum shear force, compared to compression crushing resistance. c) The belief in significant aggregate lock contribution at maximum load is tantamount to a denial of the strut-and-tie model, the basic hypothesis of which is that, at maximum load, no stress gets transferred across the shear crack paralleling the compression strut. d) In high-strength concrete, the main crack cuts through the aggregate pieces and is so smooth that no aggregate interlock stresses can be transmitted. Yet the normalized shear strength is not reduced compared to normal concrete. e) Filtered subsets of the experimental database, in which the mean aggregate size has very different, but uniform, magnitudes throughout subsequent size intervals, and the simultaneous variation of steel ratio and shear span with the beam size is suppressed, document that the aggregate size, and thus also the aggregate interlock, must contribute to the maximum shear force negligibly. f) The test of a 4 m deep beam at the University of Toronto does not prove Bentz and Foster’s claims. Its interpretation is ambiguous.

Keywords: Concrete fracture, scaling of failure, energy methods, aggregate interlock, shear strength of beams, punching strength.
The main critical statements of Bentz and Foster [1] are rebutted, point-by-point, as follows:

1) On p. 1482: “…the Dönmez and Bažant’s paper is entirely predicated on energy-based methods as governing the failure behavior of not just the listed unreinforced materials but also for failure modes in reinforced materials such as structural concrete. This is an unstated premise in their paper given without technical justification and which may or may not be consistent with observed experimental behavior.”

This opinion is incorrect. The experimental verification of the SEL (Size Effect Law) has been extensive, not only for concrete [2, 3, 4, 5, 6, 7, 8] but also many other quasibrittle materials [9, 10], and energy-based fracture analysis is as well established as Newton laws. In more detail see the next point, which is related.

2) On p. 1482, Bentz and Foster see a “dichotomy between strength-of-materials methods and energy methods.”

There is none. The former is a consequence and a special case of the latter. The equilibrium conditions follow from the vanishing of the first variation, $\delta \Pi$, of the potential energy (or free energy) of the structure, $\Pi$ [11, 12, 13]. The latter is richer because the second variation $\delta^2 \Pi$ decides stability, e.g., whether the crack grows stably at increasing load (as in RC beam shear at the outset) or becomes unstable (and thus fails dynamically) [14, Ch. 4, 5, e.g.].

3) On p. 1482: “In general, energy-based methods are important when some type of inertial effect is required to, at least temporarily, achieve dynamic equilibrium rather than static equilibrium, at least locally near a crack.”

Again, this view is incorrect. Static equilibrium represents the limit case of energy-based methods of analysis for the case of vanishing inertial force. For equilibrium, there is no difference between the static and energy methods. The energy release rate is the fundamental concept of equilibrium fracture propagation, introduced in 1921 by Griffith [15] (note that the rate is taken with respect to crack length, not with respect to time).

4) On p. 1482: “shear failures in members without stirrups are critical by an internal energy balance, that is, without testing machine rebound.”

This statement is not true. In reality, the energy release from the testing machine always matters for stability, and what decides stability is the second variation of the total energy of both the test specimen and the testing machine at controlled load (the first variation decides equilibrium).

5) On p. 1483: “energy balance methods do indeed apply to shear failures, of course, but only if all significant sources and sinks of energy are accounted for in the analysis. In particular, if an analyst significantly underestimates the energy contribution of aggregate interlock stresses moving through crack slips, the energy balance calculations can contain significant errors and can mislead the engineer.”

In truth, the “energy contribution of aggregate interlock” was not underestimated in the previous finite element analyses, e.g. [16]. Rather it was found to be negligible, for the maximum load. Bentz and Foster seem not to realize that the “energy contribution of aggregate interlock” to the ultimate strength is the work $\Delta W_{ag}$ of aggregate interlock stresses on infinitesimal incremental (rather than total) deformations at peak load. What matters are, of course, the relative magnitudes. Work $\Delta W_{ag}$ is negligible compared to 1) the work dissipated in compression-shear crushing of the
zone above the diagonal crack tip at maximum load, at the root of the compression “strut”, and to 2) the release of stored strain energy driving the failure at maximum load, as shown in [2,3] but overlooked by Bentz and Foster. Besides, at maximum load (unlike the crack initiating load), the aggregate interlock stresses contribute only a small portion of the total shear force \( V_c \), as demonstrated by finite element simulations using a realistic constitutive damage model (microplane model M7) and calibrated by optimum fitting of relevant test data [2, 3]. Also, these views of Bentz and Foster are in conflict with the well verified and widely accepted strut-and-tie model, whose basic hypothesis is that the compression strut, running parallel to the diagonal crack, is free of aggregate interlock stresses.

6) On p.1483: “For slender members without stirrups failing in shear, the critical shear crack theory (CSCT) and modified compression field theory (MCFT) both agree that the shear strength is governed by the ability to transfer shear forces across a critical shear crack”

Strangely, Bentz and Foster are not bothered by the discrepancy of the asymptotic slopes of these two theories, which are \(-1/2\) and \(-1\), respectively. Neither are they bothered by the thermodynamic impossibility of slope \(-1\) (Eqs. 4 – 7 of [5]), which also implies the absurdity that, for a large beam size, the doubling of beam size would not increase its load capacity. Further, it is interesting that, according to the derivation procedure of MCFT, the asymptotic slope could be changed from \(-1\) to, e.g., \(-1/2\) merely by changing their arbitrary formula for the decrease of crack bridging stress at increasing cross-crack displacement to another equally plausible formula. Besides, the agreement between CSCT and MCFT proves nothing because both are based on similar (and incorrect) hypotheses, ignoring compression-shear crushing above the diagonal crack tip and the release of strain energy spontaneously driving failure at the maximum load.

7) On p.1483: “The capacity to transfer stresses across this crack is largely governed by aggregate interlock, and thus the breakdown of this mechanism can be used to predict shear strength and the size effect. This means that these methods take a strength-of-materials approach to shear failure rather than an energy method...”

This view is erroneous. The “breakdown of this mechanism” (i.e., aggregate interlock) occurs already long before the peak load and, anyway, is not, by itself a predictor of peak load or size effect. Again, Bentz and Foster fail to recognize the equivalence of “strength-of-materials” and “energy” methods. These methods differ only by analysts’ convenience and simplicity. Furthermore, Bentz and Foster’s recall of aggregate interlock amounts to a denial of the strut-and-tie model.

In previous works on beam shear and punching shear [3, 5], the size range (in \( \log d \)) was subdivided into equal size intervals and a filtering program was developed to delete, one-by-one, in an unbiased way (without human intervention), the upper and lower outlier data points from each size interval, to produce a data subset with nearly uniform values of the interval means of the secondary parameters—that is, a data subset minimizing the variance of the means of the secondary variables in the individual size intervals. To be specific, denote by \( Y_{kj} \) \((j= 1, 2, .., N_k)\) all the individual data points in size interval number \( k \) \((k = 1, 2, \ldots N_{int})\). The program algorithm produces data subsets by deleting database points \( Y_{kj} \) one by one. Each point to be deleted is chosen (by the program) such that its deletion would meet the following objective – achieve near-uniformity of the secondary variables by the greatest possible reduction of the sum of squared deviations of the size interval means of the secondary variables from the overall mean. The minimization is applied simultaneously to all-important secondary variables—that is, to the overall...
combined variance. The deletions terminate when the coefficient of variation of the interval means attains a specified small enough value, such as 5%.

Fig. 2.1 shows the data subsets obtained by the filtering algorithm for four different cases in which different mean values of the maximum aggregate size are used [17]. They are 15, 20, 25 and 30 mm, while the other secondary variables (\(\rho\), and \(a/d\)) are kept the same through all the four cases. Now it should be noted in the second row of Fig. 2.1 that, contrary to Bentz and Foster’s claim, the filtered data agree with one and the same final asymptotic slope of \(-1/2\) for all the aggregate sizes considered.

Also note that in case of high-strength concrete, which must also be covered by the design code, the crack propagates through the aggregate pieces rather than around them. Thus the cracks in high strength concrete are much smoother and provide hardly any aggregate interlock [18, 19]. Yet the shear strength of high-strength concrete beams is much higher, despite much weaker aggregate interlock. How do Bentz and Foster reconcile these opposing facts? Obviously, one must conclude that, at maximum load (rather than at crack initiation load), the role of aggregate interlock is negligible, for both normal and high strength concretes.

8) On p.1483: “It appears likely from the results in their paper that Dönmez and Bažant used a numerical model that does not account for aggregate interlock well and some discussion of it is warranted”

Bentz and Foster’s claim that the microplane model M7 does not reproduce aggregate interlock is baseless. Actually, M7 represents this interlock better, in the form of a crack band (as narrow as desired). Compared to the simple stress-displacement relations for a cohesive crack, a crack band with M7 represents the aggregate interlock much more realistically because the damage law takes into account all stress tensor components and, importantly, can capture the effect of compressive strain and stress parallel to the crack, which greatly weakens the aggregate interlock but cannot be taken into account by a simple stress-displacement relation for a cohesive crack.

The importance of this effect is obvious from the fact that a large enough compressive stress can cause a splitting crack even without any transverse tension. For beam shear or slab punching, it does matter, because the crack-parallel compressive stress is large. And again, insistence on a paramount role of aggregate interlock it tantamount to a denial of the strut-and-tie model.

9) P.1484, at the end of the 2nd section: “As noted above, neither of these observations is consistent with energy methods governing failure behavior, though it remains to be seen if this type of experiment is appropriate for one-way beams in shear rather than panels as tested.”

Not correct. Bentz and Foster try to document the shear transfer by their laboratory test of pure shear (i.e., Mode II) in a pre-cracked concrete section. However, this test has hardly any relevance. The front of the beam shear crack propagates in Mode I (opening crack) and, behind the advancing crack front, transits to a mixed-mode stress transfer. Because of the asperity resistance and the resulting dilatancy, the shear slip in normal strength concrete cannot occur at constant crack opening displacement if the normal stress across a rough crack is held constant. Rather, the opening increases, which is called the dilatancy. If the opening, rather than the normal stress, were controlled, a large compressive stress across the slipping crack would develop [20], which would eliminate tensile post-peak softening.
10) P. 1484: “It is notable that in the analyses undertaken by Dönmez and Bažant, the mesh size of 12.5 mm is only slightly larger than the 10 mm size of the largest aggregate particles. It is unclear how the fracture process zone was modeled.”

It is not ‘unclear’ if one realizes that, to avoid spurious mesh sensitivity, a decrease in the finite element size (or crack band width) is, in the crack band model, compensated by an increase of the post-peak slope of stress-strain curve, such that the fracture energy $G_f$ remain unchanged [21, 22]. A smaller element size allows, of course, a more accurate resolution of the stress field.

11) At the end of p.1485: “Note that Dönmez and Bažant reported that their M7 microplane model predicted an average crack width of 1.86 mm or six times larger than that measured in the test.”

The crack width comparison cannot be made in such a simple way. Blaming the microplane modeling of the specimen for finding a crack width six-times larger than seen in the tests is unjustified. The crack width is nonuniform and varies significantly along the main diagonal crack. Thus the crack width observations depend on the location of the strain gauge, and the interpretations of the maximum crack width for beams of various sizes vary significantly in the literature. The studies in [6, 8, 23] claim that the crack widths widen as the size increases. So the aggregate interlock effect diminishes for the larger sizes. But note that the crack widths measured in [7] are inversely proportional to the beam sizes. Also, in the Toronto experiments (which they use as the reference), the maximum measured crack widths, for the beams from small to large, are 0.5, 0.25, 0.15, and 0.30 mm, respectively [24]. So, assertions based on such varied data on the crack width are dubious.

Even more importantly, one must realize that, on the surface, the crack width is visible and measurable because, due to the wall effect (increased cement mortar content and decreased aggregated content in the surface layer) the cracking localizes easily. But in the interior the crack is fuzzy and interrupted, taking the form of band of cracking.

12) On p.1487: “If aggregate interlock stresses are only taken as components of a residual direct tensile stress, as it seems was assumed by Dönmez and Bažant, unrealistic conclusions can be obtained such as a predicted average crack width that is six times too high and a shear strength governed by a diagonal strut. This is despite the fact that the model had been explicitly calibrated to match the experiment.”

As explained in detail in [5], FE calculations with M7, calibrated by tests, show that the aggregate interlock is important for the start of cracking, which occurs at about half of the maximum load, $P_{max}$. This is true not only for large beams, for which a realistic model for the softening cohesive crack indicates that, at $P_{max}$, the crack transmits only about 5% of $P_{max}$, but also for small beams, for which the crack transmits about 30% of $P_{max}$. The FE calculations with M7 show the aggregate size to affect the transitional beam size, $d_0$, in the SEL, but not the asymptotic slope of the SEL, contrary to Bentz and Foster’s claim. In particular, the calculations reveal the $d_0$ to be proportional to $\sqrt{d_a}$ where $d_a$ is the maximum aggregate size [4, 5, 10].

13) On p.1488: “In 2015 at the University of Toronto, a 4.0 m deep specimen was tested in shear by Collins et al., a member significantly deeper than any previously tested. Prior to the testing, a prediction contest was held with 66 entries from around the world and with no predictions accepted after loading had begun. The Model Code 2010 prediction gave an estimate of the machine load at failure which within 3% of the test value, amongst the best of any codes of practice.”
The Toronto test of a 4.0 m deep beam, which Bentz and Foster invoke to support their views, was certainly an excellent and important test. It demonstrated that the size effect is major and indeed extends into large sizes. However, this isolated test does not validate any theory. It cannot be interpreted by simple formulas for geometric scaling because the large self-weight causes a strong variation of the shear force, $V$ when the cross-section location is varied along the region of the dominant shear crack. This destroys the similarity of loading across sizes. When the shear force was evaluated for a certain cross-section, with a certain distance $x$ from the support, the testers found the current Model Code to give the best agreement. But if the shear force were evaluated for another $x$ within the central region of the shear crack, in which the shear force is different because of high self-weight, then the Northwestern prediction would become the best, as pointed out in ACI-445 committee.

It is important to realize that a single test result can be skewed by statistical scatter, which is high in concrete testing, and thus does not suffice to rule out any theoretical inconsistency: e.g., the fact that the beam shear model of *fib* Model Code 2010 (as well as the MCFT considered by the experimenter) has a thermodynamically impossible slope of $-1$ (see Eqs. 4 – 7 in [5]).

References


Fig. 2.1: Beam shear database [17] and filtering results based on the maximum aggregate size.