

Bridge Load Testing: Diagnostic and Proof Testing

When an analytical load rating says a bridge must be posted, but the bridge has carried heavy traffic for decades without distress, a field load test can show how the structure actually behaves. This course covers the two kinds of nondestructive load tests in Section 8 of the AASHTO Manual for Bridge Evaluation (MBE): diagnostic tests, which compare measured response with analysis and adjust the calculated rating, and proof tests, which load the bridge to a target level to confirm a lower-bound capacity. It explains the background in NCHRP Research Results Digest 234 (1998), how to plan instrumentation and loading, the MBE equations for the adjustment factor K and for the proof-test operating rating, and the safety and stop criteria that keep a test from damaging the bridge. Equations are as summarized in published DOT research reports and peer-reviewed papers that quote the MBE. Example numbers are illustrative teaching values unless a source is named. Article numbers vary by MBE edition; check the edition your agency has adopted.

Learning objectives

- Distinguish diagnostic and proof load tests
- Plan instrumentation and loading
- Update ratings from test results
- Recognize safety limits and stop criteria

Why test a bridge in the field

An analytical load rating is a model. It uses assumed material strengths, assumed support conditions, code distribution factors and code dynamic load allowance. Each assumption is chosen to be conservative, and when they stack up the result can be a rating well below what the bridge really carries. Published research has repeatedly found analytical ratings to be conservative for certain bridge types. The 2024 Illinois Center for Transportation study of reinforced concrete slab bridges (FHWA-ICT-24-022), for example, was motivated by older rural slab bridges posted to lower loads than they had shown they could carry, and it ran diagnostic and proof tests on six such bridges.

The MBE lists nondestructive load testing as one of the higher levels of evaluation available when a bridge fails a legal load rating, along with refined analysis and site-specific load factors. Load testing is not a shortcut. It costs money, needs traffic control, and has real risk if planned badly. It is most useful when:

- The rating is controlled by an uncertainty that a test can measure, such as load distribution between girders, composite action that is not shown on plans, or end restraint.
- Plans are missing, so section properties or reinforcement are uncertain.
- The bridge is posted or close to posting, and the posting has a high cost to the community or to freight.
- Deterioration makes the analytical resistance hard to estimate, and the owner wants evidence of current behavior.

The two types of test

The MBE distinguishes two kinds of static load test:

- **Diagnostic load test.** A known load, usually one or more weighed trucks, is placed at chosen positions. Strains and deflections are measured and compared with the values predicted by the rating model. The test load is normally well below the rating load. The result is used to adjust the analytical rating up or down through a factor K .

- **Proof load test.** The bridge is loaded in steps up to a target proof load chosen so that, if the bridge carries it without distress, a minimum rating is demonstrated. The result is a lower-bound operating level capacity, not a calibrated model.

Dynamic tests, in which trucks cross at speed, are also used, mainly to measure the actual dynamic amplification. Zhou and Guzda (2020) describe calculating a field impact factor by dividing the peak response from a high-speed run by the peak response from a crawl-speed run of the same truck, in the same lateral position.

Background: NCHRP Research Results Digest 234

The load testing provisions of the MBE trace back to NCHRP Research Results Digest 234, *Manual for Bridge Rating Through Load Testing*, published in 1998. As de Vries et al. (2023) note, the 1998 manual references and attaches a technical report by Lichtenstein, and the default target proof load factor of 1.4 was derived in that report from a probabilistic analysis. The MBE still refers engineers to RRD 234 for background, for example on how to choose the section factor when computing theoretical strains.

A load test does not replace inspection or engineering judgment. It answers a narrow question: how did this structure respond to this load on this day. Turning that answer into a rating that can be used for years requires deciding whether the observed behavior will still be there at higher loads and later in the bridge's life.

Diagnostic testing and the K factor

In a diagnostic test the engineer compares measured and calculated response and adjusts the analytical rating factor. The MBE expresses the adjusted rating as:

$$RF_T = RF_c \times K$$

where RF_c is the analytical rating factor and K is an adjustment factor from the test. K is:

$$K = 1 + K_a \times K_b$$

K_a accounts for the benefit (or loss) shown by the test, and K_b reflects how well the engineer understands the test results and whether the observed behavior can be relied on at higher loads.

The K_a term

The MBE's general expression for K_a is:

$$K_a = (\epsilon_c / \epsilon_T) - 1$$

where ϵ_T is the maximum member strain measured during the test and ϵ_c is the calculated strain due to the test vehicle at the position that produced ϵ_T . If the bridge strains less than predicted, ϵ_c/ϵ_T is greater than 1 and K_a is positive, so the rating goes up. If the bridge strains more than predicted, K_a is negative and the rating goes down. A diagnostic test can lower a rating, and the engineer has to report that result honestly.

The theoretical strain is calculated from the theoretical load effect L_T in the member, a section factor SF and the modulus of elasticity E :

$$\epsilon_c = L_T / (SF \times E)$$

The MBE says the theoretical strain should be calculated using a section factor that most closely approximates the member's actual resistance during the test, and refers the reader to RRD 234. This is where judgment enters. If a steel girder with no shear connectors is acting compositely with the deck during the test, using the noncomposite section modulus will make ϵ_c large and K_a large. Whether that unintended composite action will still be there at the rating load is exactly the question K_b is meant to address.

The K_b term

MBE Table 8.8.2.3.1-1 gives recommended K_b values based on two questions. First, can the member behavior seen in the test be extrapolated to 1.33W? Second, how large was the test load effect T compared with the unfactored gross rating load effect W?

Behavior can be extrapolated to 1.33W?	$T/W < 0.4$	$0.4 \leq T/W \leq 0.7$	$T/W > 0.7$
Yes	0	0.8	1.0
No	0	0	0.5

Two points stand out. A very light test (T/W below 0.4) gives $K_b = 0$, so it cannot change the rating at all; the test load is too small to say anything about behavior at rating loads. And when the engineer cannot confirm that the behavior will hold at 1.33W, the benefit is cut to zero or to half even for a heavy test.

Worked example (illustrative numbers)

The following numbers are illustrative teaching values, not from a real bridge. A simple-span steel girder has an analytical Strength I legal-load rating factor $RF_c = 0.82$ for the governing legal vehicle. A diagnostic test with a loaded three-axle truck produces a measured maximum bottom-flange strain $\epsilon_T = 150$ microstrain. The rating model, with the truck at the same position and using the section factor the engineer judges most representative, predicts $\epsilon_c = 195$ microstrain.

- $K_a = 195/150 - 1 = 0.30$
- The test load effect is $T/W = 0.55$, and the engineer has checked that the behavior (load distribution from a deck in good condition) can reasonably be extrapolated to 1.33W, so $K_b = 0.8$.
- $K = 1 + 0.30 \times 0.8 = 1.24$
- $RF_T = 0.82 \times 1.24 = 1.02$

If the engineer could not confirm that the behavior would extrapolate, K_b would be 0, $K = 1.0$ and the rating would stay at 0.82. The difference between those two answers is a decision about the structure, not arithmetic, so the test report should state the reasons for the K_b chosen.

Proof testing: target load and rating

A proof test loads the bridge to a target that, if carried safely, demonstrates a minimum capacity. The MBE target proof load is:

$$L_T = X_{pA} \times L_R \times (1 + IM)$$

where L_R is the comparable unfactored live load due to the rating vehicle for the lanes loaded, IM is the dynamic load allowance, and X_{pA} is the adjusted target live load factor.

The target live load factor

The base value is $X_p = 1.40$. It is adjusted for bridge-specific conditions using MBE Table 8.8.3.3.1-1, and when several conditions apply the percentages are summed:

$$X_{pA} = X_p \times (1 + \Sigma \% / 100)$$

The adjustments listed in the table, as reproduced by de Vries et al. (2023), are:

Consideration	Adjustment
One-lane load controls	+15%
Nonredundant structure	+10%
Fracture-critical details present	+10%
Bridges in poor condition	+10%
In-depth inspection performed	−5%
Ratable, existing $RF \geq 1.0$	−5%
$ADTT < 1000$	−10%
$ADTT < 100$	−15%

After adjustment, X_{pA} should not be less than 1.30 or more than 2.20. The one-lane increase reflects the fact that two-lane loading in the calibration includes a reduction for the low probability of two heavy trucks side by side; when one lane controls, that reduction is not available, so the factor goes up.

When the purpose of a test is only to verify a rating for a permit load, Zhou and Guzda (2020) note that the MBE allows X_p to be reduced in line with the permit load factors, which range from 1.10 to 1.40.

From the maximum load to a rating

The test is run in increments while response is watched. When it ends, the maximum live load actually applied, L_p , is known. The operating level capacity is:

$$OP = k_O \times L_p / X_{pA}$$

where $k_O = 1.00$ if the target proof load was reached without distress, and 0.88 if the test was stopped because of observed distress before the target was reached. The operating rating factor follows by dividing OP by the rating vehicle effect including dynamic load allowance, $L_R(1 + IM)$. Because the test was stopped at L_p , the result is a lower bound: the bridge may well be stronger, but the test does not show it.

Zhou and Guzda (2020) point out that a proof test should only be expected to verify a lower-bound rating at the LRFR Design Operating, Legal or Permit level, or the LFR Operating level. It does not establish an Inventory rating, and it says nothing about fatigue or long-term serviceability.

Worked example (illustrative numbers)

The following numbers are illustrative teaching values. A two-lane, multi-girder concrete bridge with redundant load paths is to be proof tested for a legal vehicle. ADTT is about 600, one lane does not control, and an in-depth inspection has just been done.

- $\Sigma\% = -10$ (ADTT < 1000) – 5 (in-depth inspection) = –15%
- $X_{pA} = 1.40 \times (1 - 0.15) = 1.19$, which is below the 1.30 floor, so use $X_{pA} = 1.30$.
- If $L_R(1 + IM)$ for the lanes loaded is 300 kip-ft of moment, the target is $L_T = 1.30 \times 300 = 390$ kip-ft.
- If the test reaches 390 kip-ft with linear, repeatable response and no distress, $k_O = 1.00$, $OP = 390/1.30 = 300$ kip-ft, and the operating rating factor is $300/300 = 1.0$.
- If the test had been stopped at 350 kip-ft because of a new crack, $k_O = 0.88$, $OP = 0.88 \times 350/1.30 = 237$ kip-ft, and $RF = 0.79$.

Planning a load test

A good load test is designed on paper before anyone drives a truck onto the bridge. The plan should answer four questions: what decision the test is meant to support, which response quantity controls that decision, what load is needed to measure it, and how the test will be stopped safely if the bridge does not behave as expected.

Start with the rating and the decision

Begin with a complete analytical rating. It tells you which member and which force effect control (positive moment in an exterior girder, shear at a dapped end, negative moment over a pier, and so on), and by how much the rating falls short. If the governing rating factor is 0.95 for a legal vehicle, a modest diagnostic gain may be enough. If it is 0.45, a diagnostic test is unlikely to close the gap on its own and a proof test, strengthening or posting may be the realistic options.

The test objective should be written down in plain words, for example: "Measure the live load distribution to the exterior girder of Span 2 and the degree of composite action, in order to update the legal-load rating of that girder." A clear objective drives sensor layout and load positions, and it keeps the test from becoming a general data-collection exercise that does not answer the rating question.

Choosing diagnostic or proof

The choice depends on what is uncertain:

- **Diagnostic** tests suit bridges where the uncertainty is in analysis: load distribution, composite action, end fixity, participation of parapets or curbs. The structure's resistance is reasonably well known, but the load effect in the member is not.
- **Proof** tests suit bridges where the resistance itself is uncertain: no plans, unknown reinforcement, unusual details, or deterioration that is hard to quantify. Zhou and Guzda (2020) describe choosing proof testing over diagnostic testing for the I-195 bridge over the Seekonk River in Rhode Island because of uncertainties in calculating the shear capacity of dapped girder ends.

Proof testing carries more risk, because the bridge is deliberately taken to a high load. Brittle failure modes, such as shear in members with little transverse reinforcement, deserve extra caution because they may give little warning before failure. Some owners will not proof test certain details for that reason.

Pre-test inspection

Before the test, the bridge should be inspected with the test in mind. The inspection records existing cracks, spalls, bearing condition, section loss and anything that might limit the test load or confuse the data. Existing cracks near sensor locations should be mapped so that crack growth during the test can be recognized. Bearings that are frozen or that have walked off their seats change the support conditions the analysis assumed, and the engineer needs to know that before interpreting strains.

Predicting the response

The engineer should run the rating model with the actual test vehicles at each planned position and predict strains and deflections at every sensor. These predictions serve three purposes:

- They let the team pick sensor locations where the signal will be large enough to measure accurately.
- They give real-time comparison values during the test, so the team can see at once whether the bridge is behaving as expected.
- For a proof test, they define the expected linear response, so the onset of nonlinearity can be spotted.

It is useful to predict response with more than one set of assumptions, for example composite and noncomposite, or pinned and partly fixed ends. The measured response can then be placed between the bounds, which helps explain what the bridge is doing.

Selecting the test load

For diagnostic tests, the test load should be large enough that the measured strains are well above the noise of the measurement system and, if a rating benefit is wanted, large enough to reach a useful T/W ratio. As the K_b table shows, a test with T/W below 0.4 cannot change the rating. Loaded dump trucks, water trucks and agency maintenance vehicles are common because they are easy to obtain and can be weighed axle by axle.

For proof tests the load must reach L_T , which may require several heavy vehicles, concrete blocks, or specially loaded trailers. Zhou and Guzda (2020) describe using two three-axle dump trucks, chosen in part for availability, with a target based on an AASHTO Type 3 equivalent. They estimated an approximate target weight of 95.52 kips for the Type 3 equivalent using the 1.40 base factor, and the final gross weight of the test vehicle at the end of the test was 101.7 kips.

Every test vehicle should be weighed, and axle spacings and wheel gauges measured, before the test. The analysis that produces ϵ_c must use the actual vehicle, not a nominal one.

Load positions and runs

Positions are chosen to maximize the force effect in the member of interest and to separate the behavior the test is meant to measure. Typical choices include:

- Trucks centered in each lane, then with the wheel line as close to the exterior girder as the barrier allows.
- Single-truck runs in each lane and side-by-side runs, so that distribution and superposition can be checked.
- Longitudinal positions that produce maximum moment, maximum shear, or maximum negative moment.
- Repeat runs at each position to confirm that the response is repeatable.

Crawl-speed runs, in which the truck creeps across while data is recorded continuously, give influence lines for each sensor. Static stops at marked positions give stable readings. High-speed runs, if traffic control allows, give dynamic amplification.

Positions should be marked on the deck with paint so drivers can stop accurately, and a spotter should confirm each position.

Traffic control and logistics

Most tests need a full closure for at least part of the work, so the plan should include a traffic control plan, a schedule, a communication plan between the data acquisition station and the truck drivers, and a decision about who has authority to stop the test. That last point matters: the person watching the data in real time must be able to halt loading at once, without having to persuade anyone.

Instrumentation and data quality

The data from a load test is only as good as the sensors, their placement and the care taken in reducing the readings. A test with a few well-placed, well-installed gauges is more valuable than a test with many gauges in places that do not answer the rating question.

Common sensors

- **Strain transducers.** Reusable clamp-on or bolt-on strain transducers are widely used on steel and concrete members because they install quickly without welding. The Illinois slab-bridge study (FHWA-ICT-24-022) used commercial strain sensors and wireless data nodes. Bonded foil gauges can also be used on steel, but they need careful surface preparation.
- **Displacement sensors.** Linear variable differential transformers (LVDTs) or string potentiometers mounted from the ground or from a reference frame measure vertical deflection. The Illinois study used LVDTs alongside strain sensors. Deflections give a global view of stiffness and distribution, but they need a stable reference, which can be difficult over water or high above ground.
- **Other sensors.** Tiltmeters at supports can show end rotation and help assess end restraint. Crack-opening gauges across existing cracks show whether cracks are working under load. Temperature sensors help separate thermal drift from load response during long tests.
- **Non-contact methods.** Survey instruments, laser devices and image-based methods are also used for deflection. They are useful where contact sensors are hard to place, but their accuracy should be checked against the expected signal.

Where to put the sensors

Sensor locations follow the test objective:

- To measure **load distribution**, place bottom-flange (or bottom-fiber) gauges on every girder at the same cross section, usually near the section of maximum moment. Each girder's share is its measured strain divided by the sum, adjusted for differences in section if the girders are not identical.
- To assess **composite action**, place gauges at two depths on the same girder, for example on the bottom flange and high on the web. The strain profile gives the location of the neutral axis, which can be compared with the composite and noncomposite values.
- To assess **end restraint**, place gauges near the ends of the span. Strains near a simply supported end should be small; significant moment near the bearing indicates restraint.
- For **shear**, rosettes or gauge arrays on webs or near dapped ends, as in the I-195 test described by Zhou and Guzda (2020), where 26 of 32 strain sensors were used to measure shear effects at the dapped ends of six drop-in girders.

Gauges should be kept away from local disturbances such as connection plates, bearing stiffeners, cover plate ends and holes, unless the purpose is to measure the local effect. A gauge a short distance from a stiffener can read very different strain from one at mid-panel.

Data acquisition practice

Good practice includes the following:

- Zero all channels with the bridge unloaded immediately before each run, and record a stretch of unloaded data so drift can be seen.
- Record continuously during crawl runs and identify truck position with a position marker, either by an operator keying the system as the front axle passes marked points or with an automatic trigger.
- Sample fast enough to capture peaks during high-speed runs.
- Return the trucks to the unloaded condition after each run and check that the readings return close to zero. A residual reading after unloading can indicate slip, a bearing that has moved, or, in a proof test, nonrecoverable deformation.
- Repeat runs. If two runs at the same position give the same peak strain within a small tolerance, the data is repeatable. If they do not, find out why before relying on either.

Converting strain to force effect

For an elastic steel member, stress is strain times E , and moment is stress times section modulus. The choice of section modulus is the same section factor question discussed for K_a . For concrete, the effective modulus and the cracked or uncracked state of the section add uncertainty, so concrete strain data needs more caution. Many engineers use concrete strain data mainly to establish distribution and neutral axis location, rather than to compute absolute moments.

Checking the data against the model

During the test, compare each reading with the prediction. Measured strains somewhat lower than predicted are common and usually reflect conservatism in the model. Measured strains far lower may mean a gauge is bad or poorly attached. Measured strains higher than predicted at one gauge, with others normal, may point to local damage, a broken connection, or a gauge error. Measured strains higher than predicted across the board are a warning that the model is unconservative, and in a proof test they are a reason to slow down or stop.

Keep a field log: run number, truck positions, truck weights, time, weather, temperature and anything unusual seen or heard. Data without a log is very hard to interpret later.

Running a proof test safely: increments and stop criteria

A proof test deliberately takes a bridge to a load well above what it normally carries. The safety of the structure, the crew and the public depends on applying the load gradually, watching the response closely, and stopping at the first sign that the structure is leaving the elastic, recoverable range.

Load in increments

The ICT slab-bridge report summarizes the MBE approach: loads are applied in stages, each stage a fraction of L_T applied incrementally, and the bridge may be loaded and unloaded several times to confirm repeatability. In practice, teams often begin with a light diagnostic-level run, then add load in steps, unloading between steps so the return-to-zero behavior can be checked. The step size usually gets smaller as the load approaches the target, because that is where the risk is highest.

At each step the team compares measured response with the linear trend from earlier steps. A plot of load against strain or deflection at the critical sensors, updated live, is the most useful single display during a proof test. While the points stay on a straight line through the origin and return to zero on unloading, the structure is behaving elastically.

Signs of distress

The test should be stopped, and the load reduced, if any of the following appear:

- **Nonlinearity.** The load-response plot starts to curve, meaning response is growing faster than load.
- **Nonrecoverable deformation.** Readings do not return close to zero after unloading.
- **Visible distress.** New cracks, growth or opening of existing cracks, spalling, buckling of plates, slip at connections, or movement at bearings.
- **Audible distress.** Cracking or popping sounds.
- **Unexpected response.** Strains or deflections at any sensor exceeding a limit set in advance from the analysis.

Limits should be written in the test plan before the test, not decided in the field. A common approach is to set a strain limit at each critical gauge from a fraction of the yield strain for steel or from a cracking or serviceability limit for concrete, with judgment for condition. The plan should name who decides and how the stop order is passed to the drivers.

What happens when a test is stopped

Stopping early is not a failure of the test. It is the system working. The MBE handles it through k_O : if distress stops the test before the target is reached, $k_O = 0.88$ and the operating capacity is based on the maximum load actually applied, L_p . The resulting rating may be lower than hoped, and the bridge may need posting or repair, but the owner now has direct evidence of behavior, plus whatever the distress revealed about the controlling weakness.

Crew and public safety

- No one should be under the bridge in the area at risk during high-load steps unless the plan has assessed and accepted the risk.
- Trucks should be placed and removed under clear instructions, with drivers ready to drive off at once.
- The bridge should be closed to public traffic during the test.
- The bridge should be inspected after the test, before it is reopened, to confirm there is no damage.

Dynamic effects in proof tests

Proof loads are applied statically, but the target L_T includes $(1 + IM)$. Measuring the actual dynamic amplification with a lighter vehicle at speed can give useful context, as in the I-195 test, where field impact was assessed by comparing high-speed and crawl-speed responses for the same truck in the same lateral position. Whether a measured impact factor may replace the code value in the rating is an agency policy question; many owners keep the code value.

Interpreting diagnostic results and updating the rating

A diagnostic test produces a large amount of data. Turning it into a rating that can be used for years means identifying what caused the difference between predicted and measured response, and deciding which of those causes can be counted on at rating loads.

Common sources of reserve

- **Better load distribution.** Code distribution factors are simplified and often conservative. Measured girder shares may be lower than the code factor for the critical girder. This benefit usually persists at higher loads as long as the deck and diaphragms stay intact.
- **Unintended composite action.** Girders built without shear connectors often act compositely at service loads because of friction and bond with the deck. At higher loads that bond may break. This is one of the most common reasons for using a reduced K_b .
- **End restraint.** Bearings that are frozen or partly fixed carry moment and reduce midspan moment. This restraint can disappear when bearings are repaired or break free, and it may induce forces the bearings and substructure were never designed for. Many engineers do not take credit for it.
- **Participation of nonstructural members.** Parapets, curbs and sidewalks stiffen exterior girders. Their contribution depends on continuity, which can be lost at joints or with damage.
- **Material properties.** Actual concrete strength and modulus are often higher than the design values.

The engineer's task is to sort these. Benefits that will reliably persist, such as distribution, can be counted. Benefits that may not persist at 1.33W or over time should be discounted, which is what K_b does in the simple K-factor method.

Using a calibrated model

Beyond the K-factor method, the measured data can be used to calibrate a refined analytical model, such as a grillage or finite element model. The model's stiffness, boundary conditions and composite assumptions are adjusted until the predicted response matches the measured response at all sensors and load positions. The calibrated model is then used to compute load effects from the rating vehicles, and the rating follows the normal MBE procedure.

Calibration should be done with discipline. Only parameters that have a physical basis should be adjusted, the adjusted values should be reasonable, and any benefit from behavior that may not persist should be removed before rating. A model that matches the data by tuning many arbitrary springs is not a better model.

Measured distribution factors

The simplest use of test data is often a measured distribution factor: the strain in the critical girder divided by the sum of girder strains at that section, multiplied by the number of loaded lanes as appropriate, and adjusted for different section properties. If the measured factor is lower than the code value, the engineer can rate with the measured value or a value between the two, according to agency policy.

When the test lowers the rating

If the measured response is higher than predicted, K_a is negative and the rating drops. The causes should be investigated: unexpected damage, missing or broken diaphragms, deteriorated connections, an error in the plans, or a bad gauge. The owner must be told promptly, and posting or other action may be needed.

Reporting

A test report should let another engineer check the result. It should include the objective, the pre-test inspection findings, the test vehicles with measured axle weights and spacings, sensor locations and types, load positions, the predicted and measured response, the analysis of differences, the K or calibrated-model approach used and why, the K_b chosen and the reasons, and the final rating factors with the governing members.

How long does a test result stay valid?

A load test describes the bridge on the day of the test. Deterioration, repairs, overlays, new barriers or bearing work can all change behavior. The rating file should say what the test relied on, so a future inspector or rating engineer can tell when a change in condition calls for a new analysis or a new test.

Published examples and lessons

Two published projects show how the MBE methods are applied in practice. Details below are taken from the cited reports.

Proof testing dapped ends: I-195 over the Seekonk River

Zhou and Guzda (2020) describe the I-195 westbound bridge over the Seekonk River in Providence, Rhode Island, built in 1967. It has 13 prestressed concrete girder spans and five steel plate girder spans. The prestressed spans use cantilevered and drop-in girder construction, and the drop-in girders are AASHTO Type IV girders with dapped ends resting on elastomeric bearings. Load ratings were governed by shear at the dapped ends, and the combination of low analytical ratings and visible shear cracks at the daps led to load testing.

Points worth noting from that project:

- Proof testing was chosen over diagnostic testing because the uncertainty was in the shear capacity of the dapped ends, which is hard to calculate reliably.
- The analysis assumed shear resistance at the dap was supplied by stirrups and draped strands, without participation of the deck. That assumption is one reason the bridge could carry more than calculated.
- They wrote the lower-bound rating from the proof test in vehicle-weight form, $RF_P = k_O(W_P/W_R)(f_V)/[\gamma_{LL}(1 + IM)]$, where W_P is the final gross weight of the test vehicle, W_R the gross weight of the rating vehicle, and f_V a vehicle adjustment factor relating the test vehicle to an equivalent rating vehicle.
- The test reached its target without signs of distress, so $k_O = 1.00$, and ratings were computed for several legal vehicles with γ_{LL} taken from the MBE for the site ADTT. They cite $\gamma_{LL} = 1.30$ for $ADTT \leq 1,000$ and 1.45 for $ADTT \geq 5,000$ in one direction, with linear interpolation between.

The vehicle-weight form and the moment-based OP form express the same idea: the load carried safely, reduced by the safety margin, gives the capacity available to the rating vehicle.

Diagnostic and proof testing of slab bridges in Illinois

The Illinois Center for Transportation study (FHWA-ICT-24-022, December 2024) tested six reinforced concrete slab bridges that had been posted. It used diagnostic and proof field tests and compared several approaches to obtaining a modified rating: the MBE diagnostic K-factor approach, the MBE proof load approach, and an effective strip width approach based on measured response. Slab bridges are good candidates for testing because the code effective strip width is simplified, and edge beams, aspect ratio, skew and railing conditions can all change how load spreads across the slab. The report notes that these factors had been identified in earlier research as important for obtaining more accurate ratings.

Research on the proof load method

De Vries et al. (2023) reviewed the background of the MBE proof load method, reproduced the Lichtenstein analysis behind the 1.4 factor, and suggested improvements. Engineers relying on the method should be aware that the target factors come

from a specific probabilistic calibration with its own assumptions, and that research continues on how well those factors cover different bridge types and failure modes.

Lessons for practice

- Use the analytical rating to decide whether a test can realistically change the outcome before spending money on one.
- Match the test type to the uncertainty: diagnostic for load-effect uncertainty, proof for resistance uncertainty.
- Be most cautious with brittle failure modes, such as shear in concrete members.
- Document every judgment, especially K_b and the reasons for counting or discounting a source of reserve.
- Record what the test relied on so future engineers know when it no longer applies.

Limits of load testing and alternatives

Load testing is a strong tool, but it has limits that the rating engineer should state plainly.

- **Fatigue.** A static test says nothing directly about fatigue life. Measured stress ranges can feed a fatigue evaluation, but the test itself does not establish fatigue resistance.
- **Time.** Behavior can change with deterioration, repair or bearing work. A test result is not permanent.
- **Brittle modes.** A proof test that does not reach the load at which a brittle mode would occur gives no warning about that mode.
- **Local effects.** A test focused on global girder behavior may miss local weaknesses such as a corroded bearing area or a cracked connection.
- **Inventory level.** A proof test gives a lower-bound operating, legal or permit level rating, not an inventory rating.

Where load testing is not suitable, the MBE's other higher-level options remain: refined analysis (for example, grillage or finite element distribution in place of code factors), site-specific live load factors based on measured truck traffic, material testing to confirm strengths, and, where the structure truly lacks capacity, strengthening, posting or replacement. In many cases the most efficient path is to combine methods, for example a refined analysis checked by a diagnostic test.

Before you seal a rating based on a load test, check that your owner agency permits it, that its procedures are followed, and that the MBE edition you use matches the agency's adopted edition. Some agencies set their own limits on K_b , on credit for end restraint or unintended composite action, and on which bridges may be proof tested.

Load test planning checklist

Use this checklist when scoping a load test. It summarizes the practices in this course; it is not an MBE or agency form, so adapt it to your owner's procedures.

1. **Rating:** analytical rating complete; controlling member, force effect and rating factor identified; shortfall quantified.
2. **Objective:** written statement of the decision the test supports and the response quantity that controls it.
3. **Test type:** diagnostic or proof, with the reason tied to where the uncertainty lies.
4. **Pre-test inspection:** cracks mapped, bearings checked, section loss recorded, existing damage photographed.
5. **Predictions:** strains and deflections predicted at each sensor for each load position, with upper and lower bounds where assumptions are uncertain.
6. **Test load:** vehicles chosen and weighed axle by axle; diagnostic T/W ratio estimated; for proof tests, X_{pA} computed with each adjustment justified and kept within 1.30 to 2.20, and L_T defined.

7. **Instrumentation:** sensor types, locations and channels listed; away from local disturbances unless intended; zeroing and drift checks planned.
8. **Load protocol:** positions marked, run sequence set, repeat runs included, increments defined for proof tests with unloading between steps.
9. **Stop criteria:** numeric limits at critical sensors, visible and audible distress triggers, and the person with authority to stop named.
10. **Safety and traffic:** closure plan, exclusion zones, communications, post-test inspection before reopening.
11. **Reporting:** report contents agreed, including the reasons for K_b or model calibration choices, and what future changes would invalidate the result.

Key takeaways

- MBE Section 8 recognizes diagnostic load tests (compare measured and predicted response to adjust a rating) and proof load tests (load to a target to show a lower-bound capacity); the background is NCHRP RRD 234 (1998).
- Diagnostic: $RF_T = RF_C \times K$, $K = 1 + K_a K_b$, $K_a = \epsilon_c / \epsilon_T - 1$. A test can lower a rating as well as raise it.
- K_b comes from MBE Table 8.8.2.3.1-1: zero when $T/W < 0.4$, and reduced to 0 or 0.5 when behavior cannot be extrapolated to $1.33W$.
- Proof: $L_T = X_{pA} L_R (1 + IM)$, base $X_p = 1.40$, adjusted percentages summed, X_{pA} limited to 1.30–2.20.
- $OP = k_O L_p / X_{pA}$, with $k_O = 1.00$ if the target was reached and 0.88 if distress stopped the test. The result is a lower-bound operating, legal or permit level rating only.
- Plan the test from a model, apply load in increments, watch response in real time, and stop on nonlinearity, nonrecoverable deformation or visible distress.

References

1. NCHRP Research Results Digest 234, Manual for Bridge Rating Through Load Testing, 1998
2. AASHTO, The Manual for Bridge Evaluation, 3rd Edition, 2018 (with interim revisions)
3. Andrawes, B., Colombani, I. A., Mirdad, M. A. H., and Chen, A. J., Load Rating of Reinforced Concrete Slab Bridges Using Field Testing, Illinois Center for Transportation, Report FHWA-ICT-24-022 (ICT-24-025), December 2024
4. Zhou, Y. E., and Guzda, M. R., Bridge Load Rating Through Proof Load Testing for Shear at Dapped Ends of Prestressed Concrete Girders, *Frontiers in Built Environment* 6:117, July 2020
5. de Vries, R., Lantsoght, E. O. L., Steenbergen, R. D. J. M., and Naaktgeboren, M., Proof Load Testing Method by the American Association of State Highway and Transportation Officials and Suggestions for Improvement, *Transportation Research Record* 2677(11), 2023

Study text only; the final quiz is taken online. Verify all references against the edition your agency has adopted.