

BOOK REVIEW

Emotional Intelligence review

A book by Daniel Goleman

The skills are more convincing than the theory

Evidence assessed through 15 September 2026 · 5,329 words · 18 min read

OVERALL RATING

64%

An editorial assessment, not a percentage of truth.

Scientific Accuracy	67%
Reference Accuracy	58%
Practical Value	67%

An engineer presents months of his team's work to a vice president. The response is a sarcastic question about how long he has been out of graduate school, followed by a dismissal of the specifications. The engineer leaves humiliated. For two weeks he considers quitting. When he finally asks what the executive intended, he discovers that the project was considered promising all along. The boss wanted revisions, not a resignation.¹

Goleman tells this story in his chapter on management. The executive has a legitimate criticism, but delivers it so badly that the engineer cannot tell what he is supposed to do next. Being considerate would not require lowering the standard. It would require explaining what was wrong and how to fix it.

This is where *Emotional Intelligence* is most useful: showing what people could do differently in a difficult situation. But Goleman wants to explain more than a badly handled meeting. He argues that these emotional and social skills form a largely separate kind of intelligence, one that helps determine how well we use our other abilities. Knowing how to give useful feedback supports a narrower conclusion than having a general ability that helps explain success across life. The distance between those claims is the central problem with the book.

My verdict: read it selectively for its treatment of managing emotions, parenting and education. Several important recommendations have good support. Be more cautious about the larger theory of success, the brain explanations, and how far the benefits extend into adult life and physical health.

Editorial rating	Score
Overall	64%
Scientific Accuracy	67%
Reference Accuracy	58%
Practical Value	67%

These are editorial ratings, not percentages of statements that are true. Confidence in the overall appraisal is **moderate**: several important historical statistics and source versions remain unresolved. Some communication practices are straightforward to begin; the school and clinical programs require sustained teaching, resources or professional support.

About this review. This AI-assisted review combines research findings with editorial judgment. Criticism of a source's use is not an allegation of dishonesty or misconduct. Independent human scientific review, human fact-checking sign-off and legal clearance are not documented. This is a house editorial review, not an official Red Pen Reviews review. The technical appendix contains the complete evidence audit, scoring rationale and notices.

What Goleman gets right

Emotional intelligence, often shortened to EI or called EQ, is not simply a fancy name for being pleasant. Goleman's framework covers noticing emotions, managing them, motivating yourself, recognizing other people's feelings, and handling relationships. Someone who recognizes anger but cannot keep it from becoming a personal attack has mastered only part of the job.²

There is a humane argument running through the book. A child can learn that being angry is acceptable while hitting another child is not. A manager can criticize behavior without attacking character. A person can notice a feeling without spending the afternoon rehearsing the grievance that sustains it. Goleman describes learning through practice and relationships, not transformation through one inspiring conversation.²

He is also more careful than some of the claims made in the book's name. In the tenth-anniversary introduction, he explicitly rejects a familiar piece of arithmetic: if IQ—the score from conventional intelligence testing—explains about 20 percent of success, emotional intelligence must explain the other 80 percent. The unexplained portion could include family circumstances, temperament, luck and many other influences. It does not belong to whichever theory arrives next.²

Goleman does not claim that EI beats IQ at everything. He acknowledges that demanding intellectual work requires intellectual ability and that potential differs from learned competence. The original chapter admits that EI's share in explaining success is unknown. His health chapter rejects curing serious illness through positive thinking and blaming patients for getting sick. His educational chapters discuss teacher preparation, age-appropriate teaching and the work needed to keep programs running.

He also recognizes that inherited tendencies can constrain change and some vulnerabilities persist, while allowing for learning in adulthood. The book's case for teaching does not require everyone to start with the same capacities or improve in the same way.

He even revises part of his earlier diagnosis. After describing worsening childhood emotional problems in the original book, he acknowledges later survey evidence of improvement in the anniversary introduction. His explanation involving social and technological change no longer fits neatly. That is a welcome correction, although these surveys measured reported problems and competence, not a national EQ score.³

This review assesses the 1995 text and that 2005 introduction. The 2012 ebook date does not make it a book based on 2012 research. Later studies tell us what remains dependable now; they do not tell us what Goleman could have known in 1995.

The three promises behind the review

Giving Goleman credit for those qualifications still leaves three substantial promises to test. These are reader summaries of the claims recorded before the detailed research; their full wording is preserved in the notes and appendix.²

1. Emotional intelligence helps determine how well we use our other abilities. Goleman presents emotional and social capacities as largely separate from IQ and useful across life. He argues that they can matter as much as, or more than, IQ for emotional and relationship outcomes, and for distinguishing outstanding leaders among professionals who have already met a high intellectual standard.

2. We can teach these skills, with benefits that last. Well-designed, sustained, age-appropriate learning at home and school can improve academic performance and behavior, reduce later problems, and help prepare children for adult relationships, work and citizenship. The proposed explanation is that repeated experience shapes emotional habits and developing brain circuits.

3. Better emotional care can improve physical health. Alongside medical treatment, managing chronic distress and providing emotional support can help prevent disease, aid recovery and sometimes improve survival. Goleman suggests that reducing chronically harmful emotions could potentially offer a medical payoff comparable to helping heavy smokers quit.

The evidence gives the first promise partial support and the other two substantial support with qualifications. To see why, we need to separate three questions that the book often brings together: what a test measures, what its score predicts, and what happens when we try to teach the skill.

What does an emotional-intelligence score tell you?

Imagine taking two assessments. One asks you to identify an emotion in a face or reason about how a situation might change someone's feelings. Another asks whether you consider yourself emotionally capable, adaptable and confident. Doing well on the first means answering questions well. Doing well on the second means describing yourself favorably. You could score highly on one and poorly on the other.⁴

Researchers distinguish performance tests, self-reports of emotional tendencies, and broader “mixed” measures that combine emotional skills with traits such as motivation and sociability. All can appear under the EI label. They overlap only partly with conventional cognitive ability—the reasoning and problem-solving abilities assessed by intelligence tests. O’Boyle and colleagues found more overlap for performance-based EI than for report-based measures. That gives some support to Goleman's argument that EI and IQ differ. It does not make all the EI measures interchangeable.⁴

There is also encouraging evidence that these scores tell us something useful. In 2026, Tyler Robinson and Ethan Zell combined 62 meta-analyses, each of which pooled results from multiple studies. They found favorable associations across different kinds of EI measures and outcomes such

as wellbeing and job performance.⁴

The association also survived an important check. If I rate both my emotional skill and my work performance, confidence in myself could raise both ratings. The relationship might then look stronger than it would if someone else assessed my work. Robinson and Zell still found positive associations when EI and outcomes were measured by different methods. The result cannot be dismissed as merely people giving themselves good marks twice.⁴

But does an EI score tell us something we could not already learn from other assessments? The overall synthesis does not answer that. It combines links to outcomes, such as performance, with links to neighboring characteristics, such as personality and confidence in one's abilities. Some underlying samples overlap, and checks for missing, less favorable studies leave concerns. The broad pattern is encouraging; the added value of a particular test requires a more specific comparison.⁴

Carolyn MacCann and colleagues made such a comparison for academic performance. They asked whether ability EI—the performance-test version—helped predict students' results after accounting for intelligence and five broad personality traits, known as the Big Five. It did, though the added information was modest and intelligence remained the stronger predictor. This tells us that the EI assessment can add something to a prediction. It does not tell us that taking an EI course will improve grades.⁵

Workplace findings depend heavily on what the EI measure contains and what it is compared with. One favorable synthesis found added prediction from some self-report and mixed measures after accounting for cognition and personality. Dana Joseph and colleagues later found that mixed EI added almost nothing after a wider set of characteristics was included, among them ability EI, confidence in one's abilities and self-rated performance.⁶

That difference is understandable. If a broad EI questionnaire partly measures confidence, it may help predict performance until you already have a separate measure of confidence. Its usefulness as a package and its usefulness *in addition to* other information are different questions. Neither result establishes that every EI test is useless, or that emotional skill has no value.

Goleman's argument about leadership needs a similar distinction. If a profession admits only people with high cognitive ability, the people inside it will differ less in that ability than the population does. Other differences may then help explain who excels. Think of the intellectual requirement as an entry condition: once everyone has met it, meeting it tells you little about which person will be the best manager.

That makes Goleman's argument plausible. It does not supply the missing comparison between EI and IQ within that group. His Bell Labs account describes initiative, useful networks and coordination among outstanding workers, after initial cognitive and social measures failed to distinguish them clearly. Those work strategies deserve attention. Classifying them as emotional competencies does not show how much success EI explains, or whether it predicts exceptional leadership better than IQ.⁶

For an employer, the useful question is specific: does this assessment improve prediction of independently evaluated performance in this job? A broad EI label is too little information to answer it. And findings about school and workplace performance leave the separate EI-versus-IQ comparison for emotional and relationship outcomes incompletely verified here.

What a marshmallow test can—and cannot—predict

The next step is easy to miss. Even when a childhood measure predicts later success, improving that measure may not produce the same success.

Consider the famous Stanford waiting task: a child could have a reward now or wait for a larger one. Children who waited longer later tended to achieve higher SAT scores, the tests used in US college admissions. Goleman reports a 210-point difference between the longest- and shortest-waiting thirds. The subtraction is correct, but the precise group averages and comparison with preschool IQ depend partly on private analyses that this review could not recover.⁷

The published results also depended on how the waiting task was arranged. One central SAT analysis involved just 35 children. Later work with a different group and a shorter task found much smaller associations after accounting for family circumstances and earlier abilities. Other analyses of that later group found some surviving behavioral associations, while a 2024 follow-up found little dependable prediction of most measured adult outcomes. These studies used different measures and comparisons; they are not a single decisive verdict on self-control.⁷

A child's choice in one waiting task can reflect more than self-control. As a possible example, a child who trusts adults to keep promises has a different reason to wait from a child who doubts that the promised reward will arrive. The same reliable support that makes a promise credible could also help that child with schoolwork years later. If so, making a child wait longer in a laboratory would not supply the support that helped produce both the waiting and the later academic advantage.

Broader evidence for self-control is stronger than evidence for using one waiting task to forecast a child's life. The Dunedin study combined repeated observations and reports across childhood. Children with greater self-control tended to have better adult health and finances and less offending, even after researchers accounted for IQ and social class.⁸

That gives us a reason to investigate self-control. It does not tell us what a particular lesson will accomplish. For that, we need studies in which people are taught something and the results are compared with what happens without that teaching.

Does the brain story explain the behavior?

Goleman's account of an “emotional hijacking” offers a memorable explanation for acting before thinking. The amygdala, a small structure deep in each side of the brain, features prominently. His diagram shows visual information about a snake traveling along two routes: one through the cortex,

the brain's outer layer, and a shorter one toward the amygdala. The picture suggests how a defensive response could begin before slower processing is complete.⁹

There is real science behind this account. Animal studies found that some learned defensive responses could occur without an intact auditory cortex, a region involved in processing sound. More recently, a 2025 human imaging study reconstructed nerve-fiber pathways consistent with routes carrying visual and auditory information to the amygdala beneath the cortex.⁹

But locating a possible route does not show how fast a signal travels along it, whether it is necessary for a response, or what the person consciously feels. A diagram of plausible connections therefore cannot, on its own, explain why someone snapped at a partner. The missing evidence concerns what those circuits were doing during that behavior.

Some of the book's claims go further still. Goleman writes that “all passion depends on” the amygdala. Yet in a small later study, three people with damage to the amygdala on both sides of the brain experienced panic when they inhaled air enriched with carbon dioxide. The study does not erase the amygdala's role in responding to external threats. It does show that panic can occur without an intact amygdala.⁹

Goleman's notes acknowledge remaining feelings after brain damage and reject the idea of a neatly separate emotional brain. Those qualifications matter. They also sit awkwardly beside the absolute claims that make the main story so memorable. I would keep the lesson that emotional reactions and deliberate thought interact, while treating the account of a particular neural takeover more cautiously.

The same issue appears in the anniversary introduction's discussion of PATHS, a school curriculum. Goleman describes academic gains attributable to better attention and working memory—the ability to hold and work with information briefly—as preliminary support for reshaping the brain. But the subsequently published study measured children's ability to restrain responses and generate words, along with teachers' ratings of behavior. It did not measure academic attainment, include a separate working-memory test or examine brain change. The program produced real improvements; those improvements do not verify the outcomes or explanation described in that passage. The earlier manuscript cited by Goleman remains unavailable, so an exact version comparison is unresolved.¹⁰

Teaching is where the book earns substantial credit

The case for teaching emotional skills is stronger when researchers test the teaching itself. Parent coaching provides a good example.

The early study Goleman used found that the way parents responded to emotions was linked to how their children fared. But a link can run in both directions. Supportive responses may help a child regulate emotions; a child who is easier to soothe may also make it easier for a parent to respond supportively. The original researchers later acknowledged this problem.¹¹

Subsequent trials tested the teaching itself. In a study of Tuning in to Kids, researchers randomly assigned parents to receive training or join a waiting list. Random assignment helps keep the groups comparable at the start, so a later difference is less likely to be explained by one group beginning with easier children. The program improved parents' discussion of emotions, children's emotion knowledge and some behavioral measures. A 2023 review of 15 randomized trials found benefits for parenting practices, children's emotional competence and behavioral adjustment.¹¹

That supports teaching specific ways of responding to children. The waiting-list comparison tells us less about whether one program is better than another substantial parenting program, or which part of the training helped most. Still, it moves the discussion beyond observing that some families are doing better than others.

Children can also improve without a study demonstrating that their underlying temperament changed. In a trial involving inhibited preschoolers—children inclined to withdraw from unfamiliar people or situations—a brief parent program reduced later anxiety diagnoses. The researchers did not detect a significant effect of the program on measured inhibition. Goleman's claim that temperament is not destiny is useful here: a tendency to be wary need not dictate whether a child develops an anxiety disorder.¹²

School programs have encouraging evidence too. Social and emotional learning teaches capacities such as recognizing feelings, managing responses and handling relationships. Major reviews published in 2011 and 2023 found benefits across several areas, including skills, behavior and academic outcomes, although results varied considerably between programs and settings. A closer look at academic results from 40 studies, published online in 2025, found a small average improvement. Whether that gain justifies a program depends in part on its cost and what schools would do with the time otherwise.¹³

Large review totals can mislead if we treat every study as answering the same question. The 2023 review covered 424 studies, but many did not measure achievement. The 2025 academic analysis drew from that same body of research; it did not add 40 new independent tests. Later reviews support benefits without verifying the 2005 introduction's specific figures for academic performance and behavior. Those historical calculations remain unresolved.¹³

Nor do the programs in the book's Appendix F work equally well for every purpose. PATHS has favorable evidence for emotional knowledge and some academic outcomes, but a later English trial found no positive effect on attainment. The Seattle Social Development Project has encouraging adult follow-up findings from a broader program, though its comparison groups were not randomly assigned. A school needs evidence for the outcome it wants to change, in a setting sufficiently like its own.¹⁴

What about benefits lasting into adulthood? Fast Track offers a serious test. It combined classroom teaching, parent training, home visits, peer work and academic tutoring for ten years, serving children selected for conduct problems. Schools were randomly assigned to the program or

comparison condition. At age 25, the program group had fewer psychiatric problems and benefits on some other outcomes.¹⁵

At age 31, the picture was mixed. Most overall comparisons did not show a clear difference. Full-time employment favored the comparison group. The conviction analyses yielded both favorable and unfavorable findings, depending on the part of the statistical model examined; estimates combining the frequency and severity of substance-related and violent convictions favored Fast Track. These results came from many tests without a correction for multiple comparisons. The more outcomes researchers test, the more chances they have to find an apparent difference by chance. Picking one striking result, favorable or adverse, would give a misleading impression of the whole program.¹⁵

A 2026 analysis offers another encouraging finding. In a selected subset of participants who became parents, childhood exposure to violence between a parent and that parent's partner was less strongly linked to later violence toward their own children in the Fast Track group. Researchers directly tested whether that link differed between groups. The finding suggests a possible benefit reaching into the next generation, but it concerns a particular relationship in a selected group; it does not establish every proposed adult or intergenerational benefit.¹⁵

Fast Track changed several things in children's lives at once. Tutoring, parent support and emotional lessons could each have contributed. The results support the combined program, while leaving questions about how much each part helped and whether the proposed brain explanation is correct.

This is substantial support for parts of Goleman's educational promise, including his emphasis on sustained work with children, families and schools. Adult outcomes remain uneven, and a school still needs evidence for the particular program and results it wants.

When the book's account differs from the study

Some of my reservations concern what the book says its sources found, rather than what later research overturned.

In the empathy chapter, Goleman describes spouses watching recordings of each other. He says they read their partners' feelings most accurately when their bodily responses synchronized, and concludes that empathy occurred only when their bodies were in sync. The cited Levenson and Ruef study actually involved married observers watching **strangers**. Similar physiological responses were associated with accuracy for negative emotion, not positive emotion. Heart-rate matching was not established as a requirement for empathy.¹⁶

The requirement is the consequential leap. A link between two things does not show that one cannot occur without the other. Nor did the researchers establish that observers had to feel the same emotion or that synchronization caused understanding. There is an interesting finding about bodily responses and reading distress here. It does not give us a bodily test for whether someone understands another person.

In the eating-disorder discussion, Goleman links difficulty identifying and managing feelings, together with dissatisfaction with the body, to the development of eating disorders. He presents this as a causal chain revealed by following girls over time. But the cited 1993 paper reports measurements taken at one stage of the research. Emotional and bodily-awareness difficulties were associated with eating-disorder risk at that point.¹⁷

Measuring those things together does not tell us whether the emotional difficulties came first, followed the eating problems, or shared another cause. A 1995 follow-up did report findings over time, and Goleman's interview could have drawn on emerging results. Even so, establishing that a vulnerability comes first would not, by itself, show that changing it prevents the disorder.

The famous 94-percent divorce-prediction figure has a different limitation. The number closely matches the original study: 44 of 47 couples with complete data were classified correctly. But the researchers developed their prediction using those same couples. A rule can fit the cases used to build it partly because it captures their peculiarities. To find out how well it travels, we need to apply the completed rule to new couples.¹⁸

Gottman points to later research following couples over time, and that evidence deserves consideration. Finding a similar association in another group is encouraging. Testing the original prediction rule unchanged on new couples is the more direct test of whether its accuracy carries over.

The reference record is mixed. Some memorable numbers are reported faithfully, including the cancer-survival figures discussed next. Elsewhere, the book changes who was studied, what was measured or how strong a conclusion the study allows. Those differences affect what a reader is entitled to believe.

Can emotional care improve physical health?

Goleman's health chapter makes a claim worth testing: emotional care can sometimes improve the course of illness, alongside ordinary medical treatment. Feeling less distressed, avoiding another heart attack and living longer are all worthwhile outcomes. Evidence for one, however, does not establish the others.

The common-cold study links stress with illness, but it does not test whether treating stress prevents it. Researchers measured volunteers' stress, then exposed them to respiratory viruses under controlled conditions. More stressed participants were more likely to develop clinically established colds; Goleman's approximate 27-versus-47-percent contrast is faithful. Infection and clinical signs were checked, so this was more than anxious people reporting more symptoms.¹⁹

The researchers controlled exposure to infection. They did not randomly assign stress, or ensure that stress was the only meaningful difference between participants. The study therefore could not establish that stress caused the extra colds through immune suppression, or that teaching people to

manage emotions would reverse the difference. Later research strengthened the evidence linking stress, the body's regulation of inflammation and susceptibility to infection. It still did not test the lifetime medical return from childhood EI lessons.¹⁹

Trials in heart patients ask the treatment question more directly. ENRICHED randomly assigned 2,481 patients after a heart attack to a program aimed at depression and low social support or to usual care. The program improved depression and patients' sense of social support. But it did not improve the main medical outcome: survival without another heart attack. Feeling better did not automatically mean fewer cardiac events.²⁰

Other trials were more encouraging. SUPRIM found fewer recurrent cardiovascular events after substantial stress-focused cognitive behavioral treatment. In the smaller ENHANCED trial, adding stress management to cardiac rehabilitation reduced the proportion of patients experiencing a combined medical outcome from 33 percent to 18 percent over a median 3.2 years. That outcome counted hospitalizations, procedures and deaths together; it does not show that deaths alone fell by the same amount. It was also a secondary outcome in a trial whose main test concerned stress. It supplies useful support for that particular program, with those limits.²⁰

The cancer evidence is especially interesting. In David Spiegel's original study, both groups of women received routine cancer care. Those also receiving supportive group therapy and self-hypnosis lived, on average, 37 months, compared with 19 months in the comparison group. Goleman's rounding of the original averages is accurate. Several later trials, including Spiegel's own replication, did not reproduce an overall survival advantage in metastatic breast cancer—cancer that had spread. Psychological benefits remained important.²¹

Yet a 2026 review by Asakawa-Haas and colleagues, including Spiegel, found favorable survival results across 32 randomized trials involving different cancers and psychological interventions. The main estimate suggested a roughly 20 percent lower rate of death during follow-up. This compares the pace at which deaths occurred among people still alive; it does not mean patients lived 20 percent longer.²²

That is evidence worth taking seriously. It is also an average across different settings. Results varied enough that the estimated range for another comparable setting included no benefit. There were concerns about study bias, and although many alternative analyses favored treatment, two adjustments for the possibility that less favorable studies were missing gave inconclusive estimates. Those checks raise uncertainty; they do not prove that the average benefit is false.²²

A pooled benefit across cancers and therapies also does not restore a dependable near-doubling of survival for women with metastatic breast cancer. Nor does it establish the book's claim that the original benefit exceeded anything medication could offer. The original trial did not compare the therapy with alternative medicines.

The wider medical literature supplies further support. A 2021 review of 106 randomized trials found survival benefits from psychosocial support: programs addressing patients' psychological, social or health-behavior needs. Its two analyses asked different questions. One counted how many patients were alive at a specified follow-up point; the other used the timing of deaths. They drew on partly different sets of trials.²⁴

In the first set, an unplanned comparison suggested clearer benefits for health-behavior programs than for programs focused on social or emotional support. But the second set, made up mostly of social and emotional interventions, also favored support. Selecting only the first result would understate the favorable evidence. Neither analysis tested a single, general emotional-intelligence ability.²⁴

We also need to know what a treatment actually changes. Ornish's coronary program combined diet, exercise, smoking cessation, group support and stress management. Goleman mentions the diet. Because the changes were delivered together, the observed reduction in coronary-artery narrowing cannot tell us how much relaxation contributed on its own.²³

And survival is not the only benefit that counts. Trials of better patient-clinician relationships and preparation for surgery support some clinical benefits. Writing studies show a more variable picture, including favorable recent findings for psychological wellbeing and quality of life. Those outcomes deserve attention without being recast as evidence of longer life.²⁴

The smoking comparison remains the largest leap. Goleman says better emotional management *could potentially* have a medical payoff like helping heavy smokers quit. Similar risk estimates, however, describe differences between groups; they do not tell us how much illness a particular intervention would prevent. The evidence supports meaningful medical benefits from specific programs. A comparable lifetime return from emotional education remains an untested forecast.

How I would use the book

Start with a situation you want to handle better. The executive in the opening needs to communicate a revision clearly. A parent may need to acknowledge frustration while enforcing a boundary. “I need to explain what I want changed” gives you something to practice. “I need more EQ” leaves the next action unclear.

Several of Goleman's recommendations offer a sensible beginning. His opposition to indiscriminate venting has experimental support. In one study, people who rehearsed a provocation while hitting a punching bag became angrier and responded more aggressively than a control group. A 2024 synthesis favored activities that lower arousal—the body's level of activation—over those that raise it. This supports cooling down; it does not show that all emotional expression or exercise is harmful.²⁵

His advice on criticism also has support beyond the opening story. Experiments distinguish the effects of specific, considerate feedback from personal attack. They do not show that every suggested script raises company productivity, but they give readers a reason to take the distinction seriously.¹

My suggested application is to choose one recurring difficulty and practice a relevant response. A manager might state the requested revision before expressing frustration. A couple might agree on how to pause an escalating discussion and return to it. Then look at whether the change helps in that situation. Also examine the situation itself: an impossible workload or an unsafe relationship is not simply a shortage of emotional regulation.

For a school or organization, the same principle requires a larger check. Identify the problem, then ask what the program includes, who will deliver it, what it replaces and which outcome will show whether it helped. Repeated practice and proper preparation are part of the work, as Goleman recognizes. A program's EI label cannot establish that its benefits will transfer to your setting.

I would also keep emotional skill separate from moral worth. The book sometimes connects emotional competence with character and being more fully human. That goes too far. Someone who struggles to identify feelings is not less human, and someone who reads people accurately may use that knowledge to manipulate them. The book itself recognizes manipulation and calculated aggression. Its humane lessons do not need a ranking of people's worth.²⁶

For serious symptoms, the book's exercises should not replace professional assessment or treatment. Goleman endorses professional help and rejects patient blame. He also recognizes that some children may experience emotional lessons as an intrusion on their privacy—a concern a school still has to address in practice.

Why the overall rating is 64 percent

Scientific Accuracy: 67 percent. The broad theory that EI governs the use of other abilities has partial support. Sustained emotional education and specific clinical programs have substantial support, with limits on adult outcomes, proposed brain explanations and the smoking-sized medical forecast.

Reference Accuracy: 58 percent. Many important sources are identifiable, and several striking findings are reported faithfully. Other accounts change participants, outcomes or the conclusions a study can support. Unrecovered originals limit the assessment; they are not counted as falsehoods.

Practical Value: 67 percent. Several recommendations have direct evidence of benefit, and Goleman supplies useful safeguards and a realistic account of some implementation demands. The weaker part is the move from particular programs and populations to broad, lasting benefits. This score evaluates the book's advice, not just my suggested way of using it.²⁷

The engineer in the opening story did not need a grand theory of emotional intelligence to receive better feedback. His boss needed to say what was promising, what needed changing, and why. Goleman's book is useful when it helps readers see and practice those kinds of skills. Its broader

explanations and promises deserve the same clear, demanding feedback.

Preparation, interests and editorial notice

Authorship and review status. This review was prepared for Jason Hreha / The Behavioral Scientist with AI-assisted research, drafting, source critique and editorial review. AI source readers checked the book, scientific sources and essay. The evidence appendix records the scope and limits of these checks. First-person evaluations are editorial judgments; they do not assert personal experience or original research actions. Repeated AI review does not constitute independent human scientific review. Human fact-checking sign-off and legal clearance are not documented.

Edition and material examined. The reviewed text contains the 1995 main work and the 2005 tenth-anniversary introduction. The supplied title page identifies the tenth-anniversary edition. The publisher's record dates the ebook for ISBN 9780553903201 to 11 January 2012; this does not establish newly researched 2012 content. All 16 chapters, introduction, prologue, appendices A–F, resources and supplied notes were included in the coordinated reading. The title page and scientific illustration were visually inspected. Four third-party advertisement inserts were excluded from the extraction and logged. Stable extraction paragraphs, rather than unverified print-page conversions, locate the book passages.

Research scope and limits. The three central claims and 28 purposive audit families were recorded before detailed external research, including all six Appendix F programs. The recorded formulations are preserved. The essay selects consequential findings; it is not an error-rate sample of the whole book. The evidence cutoff is 15 September 2026, without a claim of exhaustive searching across every database through that day. A synthesis's publication year does not date all of its underlying studies. Source access ranges from abstracts to selected or complete primary text and is identified in the records. No participant-level reanalysis or execution of the Robinson–Zell analysis code was performed.

Unresolved historical material. Important gaps include the 2005 668-study presentation and percentage calculations, the exact 2005 Riggs manuscript, private marshmallow comparisons, the original Metis RCCP report, several older source chapters and intervention presentations, and complete original text for the 1987 disease-prone-personality synthesis and 1989 Spiegel trial. Gaps restrict specific conclusions; they are not evidence of falsity and do not make every category unrateable. Targeted correction/retraction and author-response checks were conducted or documented for selected sources, not universally. An absent notice is not scientific clearance. No new request for comment was sent to Goleman for this review.

Relevant interests. The supplied house materials identify Hreha's commercially available behavior-change book *Real Change* and behavioral-science services. These create potentially competing intellectual and commercial interests. Other financial or personal relationships, sponsorship, funding, publisher-supplied review copies and affiliate arrangements have not been comprehensively established; this review makes no blanket declaration that none exist. Jason supplied the reviewed book. The research records and evidence bundle were produced during the AI-assisted preparation.

Responsibility and professional limits. The publisher is The Behavioral Scientist / Haystack Group LLC, as identified in the supplied house materials; Jason Hreha is the commissioning editor. This adaptation is not affiliated with or endorsed by Red Pen Reviews, and its ratings are not interchangeable with that organization's ratings. The review evaluates the identified book and evidence, not the author's character. It supplies scientific discussion and editorial judgment, not individualized medical, psychological, educational, legal or financial advice. Preparation and repeated AI review do not establish independent human scientific sign-off.

Evidence distinctions and corrections. The assessment includes both favorable and adverse Fast Track age-31 results, distinguishes study descriptions and source attributions, separates the two survival analyses in Smith et al., and aligns the scoring reasons with the unchanged central claims. The physical-health grade is 3/4 and the overall score is 64%; the reasons for both adjacent health grades remain available for scrutiny. The evidence appendix records the source checks and remaining evidence limits. Corrections and substantive responses should identify the edition, passage, supporting source and proposed change and be sent to the commissioning editor. Factual corrections, later evidence and changes of judgment are recorded separately. An author response or independent human scientific approval must not be inferred from preparation of this document.

Book and review details

The Behavioral Scientist · Book review · 15 September 2026

Reviewed edition: 1995 main text and 2005 tenth-anniversary introduction; ebook issued 11 January 2012. Evidence considered through 15 September 2026.

Evidence notes

These notes identify the principal evidence used in the essay, not every record in the 28-family audit. Book locators are stable extraction paragraphs, **not print pages**. Access labels describe what the AI-assisted preparation actually inspected. Each note links to its relevant technical appendix entry. These companion links open the published evidence record.

1. Feedback, EI-A13. Book: `text00017.html ¶1014–038`; notes `text00035.html ¶1202–205`. Goleman attributes the opening anecdote to psychologist Hendrie Weisinger in chapter 10, note 4 (`text00035.html ¶1202`); it is separate from Baron's experiments. Baron (1988), 10.1037/0021-9010.73.2.199, reports student experiments and employee ratings. Baron (1990), 10.1037/0021-9010.75.3.235, examines responses to destructive criticism. The prepared supplement inspected APA-supplied first-page previews/abstracts, not full methods. The exact shampoo-task attribution across the papers and Goleman's interview remains unresolved; this review does not declare the 1990 citation definitively wrong. **Appendix:** EI-A13.

2. Book scope and frozen claims. Selection: EI-SCOPE-20260915-v1. C1: `text00009.html ¶1008–019, ¶1048–062` and `text00003.html ¶1023–038`; C2: `text00004.html ¶1017–027, text00020.html ¶1019–022, text00022.html ¶1029–056, text00024.html ¶1124–141, text00025.html ¶1129–188`; C3: `text00018.html ¶1014–017, ¶1032–089, ¶1100–110`. The 80-percent rejection is explicit at `text00003.html ¶1024–027`. Temperament, treatment, educational implementation and anti-blame qualifications are included in the assessment. The body uses reader summaries; the exact recorded propositions remain below and in the appendix. Apostrophes are typographically normalized. **Exact claims used for scoring:** **1. Comparative importance.** Emotional intelligence is a set of largely IQ-independent emotional and social capacities that acts as a meta-ability, helping determine how effectively people use their other abilities and conferring advantages across life's domains. It can matter as much as or more than IQ, especially for emotional and relational outcomes and for distinguishing outstanding leadership among already intellectually selected professionals. **2. Teaching and development.** Emotional competencies can be taught and improved, particularly through sustained, developmentally appropriate learning at home and school. Well-designed emotional-literacy or social-and-emotional-learning education can improve academic performance and behavior, reduce a broad range of later problems, and make better adult relationships, work and citizenship more likely. Repeated experience shaping emotional habits and developing brain circuitry is the book's proposed explanation for these benefits. **3. Physical health.** Improving the management of chronic distress and the quality of emotional support can deliver medically meaningful disease-prevention, recovery and, in some settings, survival benefits alongside standard medical care—not merely greater comfort. Goleman proposes that reducing chronic harmful emotions could potentially have a medical payoff comparable to helping heavy smokers quit. **Appendix:** Section 04.

3. Childhood trends, EI-A28. Introduction `text00003.html ¶1039–044`. Achenbach and Howell (1993), 10.1097/00004583-199311000-00006, and Achenbach, Dumenci and Rescorla (2003), 10.1023/A:1021700430364: abstracts examined in the prepared audit. These compare cohort reports of problems/competence, not continuous national EI measurements. The earlier reference's date in the book is inaccurate; the author's later acknowledgment of improvement is substantively supported. **Appendix:** EI-A28.

4. Constructs and current favorable synthesis, EI-A01/A14. O'Boyle et al. (2011), 10.1002/job.714, Table 3: observed cognitive-ability correlations were .264, .070 and .050 for performance-based ability EI, self- or peer-report within an ability framework, and mixed measures; estimates corrected for measurement unreliability were .315, .083 and .060. The initial source audit covered relevant methods and Tables 5–6. Table 3 received an additional visual check, with the relevant methods and surrounding results read, during the combined critique and independent follow-up on 15 September 2026. These are reported meta-analytic estimates, not an independent reconstruction. Robinson and Zell (2026), 10.1073/pnas.2532963123. Abstract accessed directly; the prepared addendum reports complete supplementary-text/table reading, selected visual checks, screening-log/codebook inspection and selected code inspection, without executing the code. The main paywalled article was not

fully read. Overall $r=.28$, 95% CI [.25,.31]; positive cross-method estimates survive. Outcomes include neighboring constructs; sample overlap was restricted, not eliminated. The headline participant count is not certified as unique people. See the public appendix. Davis and Mahoney's separate 2026 methods review, 10.1007/s11135-026-02913-9, assessed an earlier review corpus, not Robinson and Zell; its negative quality ratings are not a false-finding rate. **Appendix:** EI-A01 · EI-A14.

5. Academic prediction. MacCann et al. (2020; online 2019), 10.1037/bul0000219: complete abstract accessed. $N=42,529$; 158 citations. Ability, self-rated and mixed EI yielded different associations and increments. The 1.7 percentage points concern additional variance accounted for, not a causal improvement. Full source-by-source methods were not independently audited for this essay. **Appendix:** EI-A01.

6. Work and leadership. O'Boyle et al. (2011), 10.1002/job.714: relevant methods/Tables 5–6 inspected in the initial research bundle; Table 6 was visually rechecked during the combined critique and independent follow-up on 15 September 2026. Joseph et al. (2015), 10.1037/a0037681: abstract accessed; expanded-control coefficient $\beta=-.02$ is specific to mixed EI. Kelley and Caplan (1993), PubMed 10127037: abstract/preview, not a complete study audit. Book: `text00017.html` ¶068–075; introduction ¶031–038. Harms and Credé (2010), 10.1177/1548051809350894, provides additional abstract-level evidence that reporting methods materially affect leadership associations. **Appendix:** EI-A01 · EI-A14.

7. Waiting, EI-A03. Shoda, Mischel and Peake (1990), 10.1037/0012-1649.26.6.978: relevant methods/results and Table 4 visually checked by the prepared audit; 94 respondents supplied SAT data, with 35 in the highlighted diagnostic condition. Book `text00012.html` ¶013–022 and notes `text00035.html` ¶103–106. Later evidence: Watts et al. (2018), 10.1177/0956797618761661; Michaelson and Munakata (2020), 10.1177/0956797619896270; Sperber et al. (2024), 10.1111/cdev.14129. Relevant primary sections were examined in the bundle. These later reports overlap in cohort and use different analyses; they are not three independent replications. Specific surviving associations preclude a blanket “nothing predicts anything” conclusion. **Appendix:** EI-A03.

8. Broader self-control. Moffitt et al. (2011), 10.1073/pnas.1010076108: abstract and relevant methods/results inspected in the prepared audit. The Dunedin analysis used multiple childhood assessments, not a single waiting task. A separate British twin comparison linked childhood self-control differences to later child outcomes. Neither analysis estimates the effect of teaching self-control. **Appendix:** EI-A03.

9. Neuroscience, EI-A06. Book `text00007.html` ¶015–016, ¶029–036, ¶050–077; qualifying notes `text00035.html` ¶023–025. LeDoux et al. (1984), 10.1523/JNEUROSCI.04-03-00683.1984, and related 1986 auditory-conditioning work provide bounded animal evidence. The 1984 author-posted abstract and author bibliography were checked; this is not a complete methods audit of those animal studies. Rafal and Koller (2025), 10.1152/jn.00033.2024, involved 19 people; abstract/indexed sections were inspected by the prepared audit. Feinstein et al. (2013), 10.1038/nn.3323: full primary text/methods read by that audit. Pessoa and Adolphs (2010), 10.1038/nrn2920, is a critical perspective on primate pathways, not an experiment proving subcortical routes absent. **Appendix:** EI-A06.

10. PATHS mechanism, EI-A27. Book `text00003.html` ¶015; note `text00035.html` ¶004 cites a 2005 under-review manuscript. Riggs et al. (2006), 10.1007/s11121-005-0022-1: abstract accessed directly; methods/results/Tables 2–3 inspected by the prepared audit. Four schools, two per condition; 318 pupils analyzed. Improved inhibitory control and verbal fluency, with teacher-rated behavioral outcomes, do not establish measured brain change or academic mediation. The original manuscript version is unavailable, so the critique is bounded to the accessible published evidence. **Appendix:** EI-A27.

11. Parenting, EI-A21. Hooven, Gottman and Katz (1995), *Parental meta-emotion structure predicts family and child outcomes*, and their 1996 reply: relevant primary sections examined in the prepared audit, including reverse-direction models. Havighurst et al. (2010), 10.1111/j.1469-7610.2010.02303.x, randomized 216 parents against waitlist. England-Mason et al. (2023), 10.1016/j.cpr.2023.102252, synthesized 15 randomized trials using untreated controls. These latter findings are abstract-level checks, not full independent audits of every trial. **Appendix:** EI-A21.

12. Temperament and clinical learning, EI-A22. Rapee et al. (2005), 10.1037/0022-006X.73.3.488, and three-year follow-up (2010), 10.1176/appi.ajp.2010.09111619: fewer anxiety diagnoses without significant measured inhibition change. Baxter et al. (1992), 10.1001/archpsyc.1992.01820090009002, concerns brain-metabolism changes among OCD treatment responders, not treatment equivalence. Hoppen et al. (2023), 10.1037/ccp0000809, synthesizes randomized PTSD treatments. Prepared-audit access for these findings was abstracts; exact mechanisms and all treatment details were not reaudited. **Appendix:** EI-A22.

13. School syntheses. Durlak et al. (2011), 10.1111/j.1467-8624.2010.01564.x: selected full methods/tables in the bundle; 213 controlled studies, only 35 contributing academic estimates. Cipriano et al. (2023), 10.1111/cdev.13968: abstract accessed; 424 reviewed studies are not 424 academic-effect estimates. Ha et al. (online 2025), 10.3102/00346543251367769: publisher methods, results, tables and limitations inspected. This secondary academic analysis of the Cipriano corpus includes 40 studies, 33,737 students and 76 effects; $g=.101$, 95% CI [.035,.168]. Included studies date from 2008–2020, and follow-up achievement effects were not analyzed. It is an outcome-specific reanalysis of overlapping evidence, not 40 additional independent trials. Taylor et al. (2017), 10.1111/cdev.12864, supplies favorable follow-up evidence, mostly much shorter than its maximum 18 years. None substitutes for the unrecovered 2005 conference calculations. **Appendix:** EI-A27.

14. All six Appendix F families remain in the research frame. text00031.html: Child Development Project ¶003–018; PATHS ¶019–046; Seattle Social Development Project ¶047–060; Yale–New Haven Social Competence Promotion ¶061–076; Resolving Conflict Creatively Program ¶077–087; Improving Social Awareness–Social Problem Solving ¶088–103. The prepared audit records both favorable results and gaps for each. Examples used here: Greenberg et al. (1995), 10.1017/S0954579400006374, and Schonfeld et al. (2015), 10.1037/spq0000099, abstract-level favorable PATHS findings; Hennessey and Humphrey (online 2019; issue 2020), 10.1007/s10212-019-00452-6, abstract/methods-level English attainment findings; Hawkins et al. (2008), *Effects of social development intervention in childhood 15 years later*, relevant full-paper sections. Original source chapters and the Metis RCCP report remain incompletely accessed; not every appendix bullet is certified. **Appendix:** EI-A26.

15. Fast Track: benefits and limits. Dodge et al. (2015), 10.1176/appi.ajp.2014.13060786, reports age-25 benefits; its authorship correction does not change findings. McCabe et al. (online 2024; issue 2025), 10.1007/s11121-024-01736-0: relevant primary methods, results and discussion, Table 2 and Supplementary Table S4 were examined. Most total estimates are nonsignificant. Adverse total contrasts concern full-time employment (OR .68 [.49,.94]) and the substance-crime structural-zero component (OR .66 [.44,.98]); favorable totals concern severity-weighted substance- and violent-crime count components (both IRR .76, with intervals [.59,.96] and [.58,.99]). The two adverse direct contrasts are OR .55 [.36,.79] and .54 [.28,.82]. These are nominal tests among many endpoints. The crime model's zero and count components answer different questions; neither is a simple observed proportion of people prevented from offending. The authors' contextual explanations remain hypotheses. Gorla, Rothenberg and Godwin (2026), 10.1111/jcpp.70133, concerns a selected parenting subset. The prepared supplement check confirms between-arm tests of the G1-partner-violence→G2-parent-to-child-violence component path, not a separate contrast of complete indirect-effect products. Mediators were not independently randomized, and current parent/child measures share a reporter. **Appendix:** EI-A25.

16. Physiological empathy, EI-A09. Book text00013.html ¶040–042; note text00035.html ¶143. Levenson and Ruef (1992), 10.1037/0022-3514.63.2.234: 31 married observers watched unfamiliar people. Relevant full methods, analyses and tables were examined and key tables visually checked by the prepared audit. Their 1997 chapter, *Physiological Aspects of Emotional Knowledge and Rapport*, was also examined in relevant sections and does not establish the book's precise spouse-viewing description or universal necessity claim. The criticism concerns this cited relationship, not invented research or author intent. **Appendix:** EI-A09.

17. Eating-disorder evidence, EI-A24. Book text00024.html ¶083–089; note text00035.html ¶1373. Leon et al. (1993), 10.1037/0021-843X.102.3.438, explicitly reports cross-sectional findings among 937 girls. Leon et al. (1995), 10.1037/0021-843X.104.1.140, reports prospective, sex-specific results with prior risk especially predictive. Complete abstracts were examined in the prepared audit. They establish the design distinction, not independent verification of every coefficient or interview statement. **Appendix:** EI-A24.

18. Marriage, EI-A12. Book text00016.html ¶025. Buehlman, Gottman and Katz (1992), 10.1037/0893-3200.5.3-4.295: methods, predictor selection and classification table examined in the bundle. Forty-four of 47 classified complete cases were correct; 52 families had follow-up. The Gottman research FAQ supplies the authors' prospective-research response. A frequently cited 29-percent positive predictive value from a different critique dataset is not a reanalysis of this study and is not used here. **Appendix:** EI-A12.

19. Stress and colds, EI-A15. Book text00018.html ¶050–053. Cohen, Tyrrell and Smith (1991), 10.1056/NEJM199108293250903: main text and relevant figure/results checked in the prepared audit; stress measured, virus exposure controlled. Cohen et al. (2012), 10.1073/pnas.1118355109: abstract-level later evidence about glucocorticoid resistance and inflammation. Original correspondence was only partly accessed; neither complete resolution of objections nor an integrity clearance is claimed. **Appendix:** EI-A15.

20. Cardiac interventions, EI-A16. ENRICH (2003), 10.1001/jama.289.23.3106, abstract: event-free survival 75.8% versus 75.9% at mean 29 months. SUPRIM (2011), 10.1001/archinternmed.2010.510, abstract: recurrent cardiovascular benefit, without established all-cause mortality benefit. Blumenthal et al. (2016), 10.1161/CIRCULATIONAHA.115.018926, abstract and relevant primary design/endpoint sections: 151 randomized patients;

stress primary, clinical composite secondary. The favorable no-rehabilitation comparison was not randomized. Richards et al. (2017), 10.1002/14651858.CD002902.pub4, and Ski et al. (2024), 10.1002/14651858.CD013508.pub3, address different review questions and leave important event uncertainty. **Appendix:** EI-A16.

21. Specific cancer trials, EI-A18. Book text00018.html ¶088–089. Spiegel et al. (1989), 10.1016/S0140-6736(89)91551-1: original abstract, not full protocol; survival figures are means. Goodwin et al. (2001), 10.1056/NEJMoa011871, and Kissane et al. (2007), 10.1002/pon.1185: abstracts and, for Goodwin, indexed primary sections. Spiegel et al. (2007), 10.1002/cncr.22890: complete abstract and selected indexed primary methods/discussion passages inspected; no complete original-paper audit is claimed. The authors acknowledge failure to replicate overall survival; their favorable receptor-negative subgroup is exploratory, and their explanations involving medical-era and population differences remain relevant rather than conclusive. **Appendix:** EI-A18.

22. Current favorable cancer synthesis. Asakawa-Haas et al. (2026), 10.1038/s44271-026-00414-x: abstract and selected methods/results/discussion, not a full supplement or original-data audit. Main death HR=.80 [.71,.90]; prediction interval [.49,1.29]. Some alternative bias adjustments cross the null; the authors question the suitability of these adjustment methods under the observed data conditions. A broader pooled benefit neither replicates the specific near-doubling nor establishes comparative efficacy against medicines. Spiegel's coauthorship is disclosed, not treated as disproof. **Appendix:** EI-A18.

23. Ornish, EI-A20. Book text00018.html ¶100. Ornish et al. (1990), 10.1016/0140-6736(90)91656-U: main text and extracted tables read by the prepared audit; 48 participants, 41 follow-up angiograms, demanding combined intervention and consent/selection limits. The outcome was coronary lumen narrowing, not direct plaque-volume measurement. The 1998 follow-up is the same trial, not an independent replication; its formal correction adds an author rather than changing cardiac results. **Appendix:** EI-A20.

24. Care and psychosocial support, EI-A17/A19. Kelley et al. (2014), 10.1371/journal.pone.0094207: selected full sections, small clinical benefit across 13 relationship-intervention trials, often with other components. Powell et al. (2016), 10.1002/14651858.CD008646.pub2: low-certainty preoperative-preparation benefits; abstract and authors' plain-language summary read, without a full-report audit. Jensen-Johansen et al. (2018), 10.1371/journal.pone.0192729, supplies a null main physical-health writing result; Han et al. (2026), 10.1188/26.ONF.e26535276, reports favorable psychosocial/quality-of-life findings in a one-page preview. A null result for physical symptoms and GP use does not directly contradict pooled benefits in anxiety or quality of life. The later review's full inclusion table was not inspected, so overlap with the earlier trial is unresolved. Neither report establishes cancer survival. Exact historical writing-marker claims remain partly untraced. Smith et al. (2021), 10.1371/journal.pmed.1003595: abstract, methods, results and selected discussion inspected. Across 87 fixed-time studies, survival OR 1.20 [1.09,1.31]; behavior-focused subgroup OR 1.35 [1.20,1.52], social/emotional-focused subgroup OR 1.01 [.87,1.16]; between-subgroup $Q=9.3$, $p=.002$. This split was requested during peer review and was not prespecified. A partly different set of 22 time-to-event studies also favored support (benefit-oriented HR 1.29 [1.12,1.49]); 18 focused on social/emotional outcomes, and no comparable program-type subgroup test was performed in that set. Three studies contributed to both sets. Neither estimate measures percentage extra lifespan. The heterogeneous clinical programs overlap other cited trial corpora; they do not test a unitary EI capacity. **Appendix:** EI-A17 · EI-A19.

25. Anger, EI-A08. Book text00011.html ¶020–051. Bushman (2002), 10.1177/0146167202289002: relevant full methods/results inspected in the bundle, 600 analyzed participants, immediate anger and laboratory aggressive responding. Kjærvik and Bushman (2024), 10.1016/j.cpr.2024.102414: abstract and indexed sections; arousal-reducing activities favored, arousal-increasing activities ineffective on average and heterogeneous. The book also recommends active exercise; that recommendation should not inherit the stronger evidence for reducing arousal. No claim that all exercise or expression is harmful follows. **Appendix:** EI-A08.

26. Practical boundaries. Book text00004.html ¶018, text00009.html ¶056–062, text00010.html ¶029–033, text00014.html ¶015, ¶046–049; professional-care qualification text00011.html ¶015–016; medical anti-blame qualification text00018.html ¶016–017; implementation/privacy discussion text00025.html ¶129–139. The distinctions between skills and moral worth, and the suggested applications in this review, are editorial judgments—not measured effects of a replacement program. Ordinary conflict advice is not a direction to manage a violent partner; the book's note text00035.html ¶186 itself distinguishes calculated from impulsive violence. **Appendix:** Section 07.

27. Scoring. Supplied house method, Edition 3.0. Scientific inputs: 2,3,3. Reference inputs: 3,2,2. Practical inputs: 3,2,3. Categories are 8/12, 7/12 and 8/12; overall 23/36. Percentages are rounded only for display. One changed input changes the overall score by approximately 2.78 percentage points. These are judgment-sensitive summaries, not statistical confidence intervals, error rates or validated estimates of reader benefit. **Appendix:** Section 09.