TECHNICAL APPENDIX

The evidence behind Grit

Claims, source checks, scoring, and disclosures

Angela Duckworth · Evidence cutoff: 13 September 2026

Report ID: TBS-GRIT-APPENDIX

79 min read

This is the supporting record for the reader review. It retains the detailed analysis rather than requiring every reader to work through it. Use the claim register to find a finding, then follow its sources and qualifications. The common tables standardize navigation; the original audit designs and identifiers remain distinct.

Overall 64% · Scientific Accuracy 58% · Reference Accuracy 83% · Practical Value 50%.

About This Review

This review combines factual reporting with editorial judgments about the identified book and its evidence. Ratings apply our published rubric; they are not percentages of truth or judgments about an author's honesty. Criticism of support does not itself show that a claim is false. Evidence is assessed through 13 September 2026, with source-access limits disclosed. Preparation used AI, and independent human fact-checking or legal clearance is not documented. Read the methodology, book-specific disclosures, and full editorial notice with the review. Corrections and substantive responses are welcome.

Find the supporting evidence

EVIDENCE MAP

Claim register

This common register makes the detailed evidence easier to navigate. Each row identifies a scoped proposition or source relationship, a verdict, and the treatment it warrants. A label applies to that row—not to the whole book or its author. The full wording, sources, qualifications, and original audit identifiers remain in the linked exhibits.

> **Verdicts:** Supported; Partly supported; Overstated; Incorrect; Unresolved; Practical judgment. "Incorrect" requires an identified factual or measurement problem; "unresolved" is an access or verification limit, not a failed claim. "Practical judgment" identifies advice or a safeguard rather than a tested effect. These routing labels add no numerical score. Definitions.

ID	Claim or focal issue	Verdict	Treatment and supporting detail
C01	Grit predicts persistence in demanding settings	**Supported**	Use persistence measures as possible context-specific predictors—not as verdicts on future excellence or character. Full assessment and sources
C02	Grit generally matters more than talent	**Overstated**	Ask which abilities and behaviors predict the specific outcome, in the specific population, at the specific stage. Full assessment and sources
C03	Grit is a substantially distinctive psychological explanation	**Overstated**	Treat grit as a useful label or candidate facet until it adds reproducible value beyond well-measured existing traits. Full assessment and sources
C04	Passion is adequately captured by stable interests	**Partly supported**	Measure sustained effort, interest stability, enjoyment, goal importance, and actual commitment separately when the distinction matters. Full assessment and sources

ID	Claim or focal issue	Verdict	Treatment and supporting detail
C05	A brief grit score can tell readers how gritty they are	**Partly supported**	Use the items to prompt reflection. For hiring, admissions, or evaluation, require purpose-specific validation and behavioral corroboration. Full assessment and sources
C06	Effort counts twice	**Overstated**	Effort can contribute at multiple stages. Its marginal value must be learned from the task, not inferred from a slogan. Full assessment and sources
C07	Deliberate practice is the unmatched route to improvement across domains	**Overstated**	Use targeted practice where the skill and feedback are valid. Combine it with exploration, knowledge acquisition, and real-world outcome testing. Full assessment and sources
C08	Practice explains why gritty spellers perform better	**Partly supported**	The findings suggest a plausible route from sustained effort to better preparation; they do not independently prove the entire causal chain. Full assessment and sources
C09	Age trends and heritability establish that grit can be deliberately grown	**Overstated**	Separate evidence that grit differs, evidence that it changes within people, and evidence that a program causes beneficial change. Full assessment and sources
C10	Developing interests improves satisfaction and performance	**Partly supported**	Explore promising activities, then evaluate the actual tasks, learning trajectory, constraints, and opportunities—not “passion” alone. Full assessment and sources
C11	Purpose sustains grit and energizes performance	**Partly supported**	Connect difficult tasks to a purpose the person actually endorses; do not substitute purpose language for autonomy, resources, or fair treatment. Full assessment and sources
C12	Growth mindset reliably produces grit and achievement	**Overstated**	Treat growth-oriented feedback as a potentially useful support for effective learning, not a substitute for effective learning. Full assessment and sources
C13	Mastering adversity can inoculate people for life	**Overstated**	Provide supported challenges and real routes to influence the outcome. Do not infer a license to manufacture trauma. Full assessment and sources
C14	Cognitive behavioral therapy has longer-lasting effects than antidepressants	**Overstated**	Psychological treatment can have durable benefits, but comparisons depend on condition, treatment quality, continuation, and the outcome being measured. Full assessment and sources
C15	Warmth plus high standards produces gritty children	**Partly supported**	Combine warmth, meaningful expectations, skill support, and age-appropriate agency. Do not generalize a celebrated person’s harsh-family anecdote into a rule. Full assessment and sources
C16	Extracurricular follow-through develops transferable grit	**Partly supported**	Offer good activities for their direct learning and social value. Treat generalized character transfer as a hypothesis, not the guaranteed return. Full assessment and sources
C17	The Hard Thing Rule is an evidence-based grit intervention	**Practical judgment**	Treat commitment periods as revisable experiments with clear review points, not as a scientifically established dose of character development. Full assessment and sources
C18	Joining a gritty culture makes a person gritty	**Partly supported**	Improve the setting’s feedback, support, standards, and incentives. Evaluate the resulting behavior rather than assuming a personality transformation. Full assessment and sources

ID	Claim or focal issue	Verdict	Treatment and supporting detail
C19	More grit is generally compatible with well-being and unlikely to have costs	**Overstated**	Assess the full outcome: progress, health, relationships, ethics, and opportunity costs—not perseverance alone. Full assessment and sources
C20	Stable top-level goals should anchor flexible lower-level tactics	**Practical judgment**	Be more stable about values than about forecasts. Reconsider even important goals when their reasons no longer hold. Full assessment and sources

01 The verdict in context

The case for sustained, well-directed effort is stronger than the case for grit as a uniquely powerful, trainable general trait. The book provides valuable distinctions about deliberate practice, interest development, goals, and persistence. Its larger rhetoric outruns the evidence when selected prediction studies become a general explanation of achievement.

Three distinctions carry the review. Perseverance and consistency of interests do not contribute equally. Prediction is not proof that raising a score will improve an outcome. And staying in a demanding program is not the same outcome as performing exceptionally within it. Overlap with conscientiousness also narrows the claim of a wholly distinctive construct.

The complete endnotes improve the assessment of the book's sourcing. They acknowledge important qualifications, some unpublished evidence, and the deliberate-practice debate. This edition must not be criticized as though those notes were absent. The ten-reference sample includes two inaccessible details; these were not replaced with easier sources. Claim and endnote audits.

Best use: improve the quality of practice and the fit of goals while preserving an explicit right to change strategy or stop. Do not use a popular grit score as a diagnostic test of character, destiny, or suitability for consequential selection decisions.

> **Recommendation:** Read selectively for practice and persistence; do not treat grit as a dominant universal cause of success.

Confidence: Moderate confidence in the overall appraisal. The two unresolved sampled source passages remain unresolved; they are not counted as errors.

Ease of application: Requires sustained effort, useful feedback, and often outside support.

Best use: Read selectively for better practice and goal management, not a universal character prescription.

Rating profile: Overall 64% · Scientific Accuracy 58% · Reference Accuracy 83% · Practical Value 50%. Full rationale.

02 Scope and method

Question: How well do the reviewed book's important scientific claims hold up, how faithfully does it use sources, and what can readers responsibly do with its advice? The intended audience is an interested general reader, with enough detail for researchers and editors to audit the reasoning.

Edition and scope. 2016 Scribner ebook; ISBN 9781501111129; complete endnotes. Coverage comprises 20 substantive claims; 10 seeded selections from a restricted 158-anchor frame; 20 additional targeted endnote checks; all 13 chapters. Book citations identify chapters and named sections of the reviewed edition. The source copies have not been authenticated against publisher-controlled masters.

Preparation. The source basis includes the complete endnotes, chapter assessments, rating rationales and dated primary-source checks. The reader review and appendix were read in full, and the reference/claim inventory, relevant later evidence and publication notices were assessed. All nine scoring inputs were evaluated under the stated method. This is a targeted critical evidence review, not a preregistered systematic review or independent raw-data replication. The supplied critique and preparation limits are described in Section 12; the shared methodology states the scoring rules.

Claim selection. Three central propositions summarize the book's organizing argument, its proposed mechanism or intervention, and important practical extensions. Selection is purposive and retrospective, not random or preregistered. All chapters remain covered to expose peripheral but consequential claims. The three proposition grades are not a sample-based estimate of the truth of the whole book. Narrow existence claims are easier to support than universal superiority claims; the stated proposition must always accompany the grade.

Evidence rules. The analysis distinguishes source identity, what the cited study actually found, the strength of its design, independent corroboration, and transfer to the book's practical claim. A controlled component study does not validate a branded package; an association does not identify an intervention effect; failure to locate evidence is not proof of its absence. Null findings, attrition, selection, selective reporting, measurement limitations, and competing explanations matter where they change an inference. Evidence available at publication and later updates are identified separately.

Ratings. The three categories are Scientific Accuracy, Reference Accuracy and Practical Value. Each uses three explicitly anchored judgments. Category percentages are calculated from the original inputs; the overall rating gives each category equal weight. Final displays use whole-number percentages, with all rounding performed after calculation. Ease of application and consequential cautions are reported separately. Section 09 and the series methodology provide the exact inputs, rules and limits. These are editorial judgments, not probabilities, estimates of the percentage of true content, or validated psychometric measurements.

Reference checks. Targeted source audits are stress tests, not prevalence estimates. The number of discrepancies in a purposive sample cannot estimate the proportion of a book's references that are wrong. The Grit report additionally preserves an explicitly restricted-

frame seeded selection, with its missing-source cases and inherited inclusion coding disclosed. It is not the same audit design as the other books.

Relationship to Red Pen Reviews. The series adapts the separation of central scientific claims, reference support and practical consequences from the public Red Pen Reviews process and method, version 2.2. It uses a behavioral-science Practical Value category instead of nutrition-specific Healthfulness. Reference Accuracy is a common three-criterion assessment, not Red Pen Reviews' ten-reference random-sample score. The three category percentages are averaged, but there is no claim of an independent second expert reviewer. These departures mean the scores are not directly interchangeable with official Red Pen Reviews scores. This review is not affiliated with or endorsed by Red Pen Reviews.

Research cutoff: 13 September 2026. The assessment includes relevant corrections, original-source checks and later evidence available by this evidence date. It does not warrant a claim of exhaustive database or notice coverage. All nine numerical inputs were evaluated against the stated anchors; the rubric and item grades are in Section 09. This is an evidence date, not a publication date.

Research search record

The passage below describes the documented research. Not every listed search, source read, or assessment was independently repeated; dated retrieval and access limits are listed in Section 10.

Search and adjudication

Targeted searches and source checks were conducted on 8 September 2026. The assessment considers the book's actual endnotes. Searches covered grit measurement, conscientiousness, retention, academic meta-analyses, randomized educational programs, deliberate practice, vocational interests, purpose, feedback, costly persistence, and the book's parenting and clinical extrapolations. New checks followed the endnotes to original articles, author-hosted manuscripts, institutional descriptions, and bibliographic records.

This is a critical evidence review, **not a preregistered systematic review, new meta-analysis, or raw-data replication**. The reference frame was coded after reading the book and the available review material, and was not externally preregistered. The deterministic pseudo-random draw is reproducible but not an independently witnessed randomization. There was no exhaustive database export, duplicate independent screening, or comprehensive correction/retraction census. Accessible original text takes precedence over search summaries, and inaccessible details remain explicitly unverified.

The unit of inference matters. An observational study can support prediction without establishing an intervention effect. A randomized curriculum can improve grades without identifying a general personality trait as the cause. A laboratory task can test momentary behavior without validating a claim about decades of persistence. The report distinguishes the book's propositions, external findings, methodological inferences, and editorial recommendations.

03 The strongest fair reading

The strongest fair reading of the book

Duckworth's argument is more sophisticated than "work hard and anything is possible." She separates talent from acquired skill and skill from productive accomplishment; describes grit as sustained passion plus perseverance; and proposes four developable assets—interest, practice, purpose, and hope. The final section adds parents, extracurricular settings, mentors, and organizational culture. That sequence is important: the book is not simply a lecture on willpower. [B, ch. 3; ch. 5; chs. 10–12]

It also contains qualifications that a fair review must preserve. Opportunity may matter more than individual psychology; talent differences are real; interests need exploration; lower-level goals can be replaced; rest belongs in intensive practice; hardship can damage people; and grit is not the most important moral quality. The conclusion recognizes that effects on spouses and families are not established. [B, chs. 2–4; ch. 6; ch. 7; ch. 9; ch. 13]

These admissions rule out several easy criticisms. The book does not literally prescribe uninterrupted work, uniformly harsh parenting, persistence in every failed project, or a fixed 10,000-hour law. Its problem is subtler: **the strength of the motivating story often exceeds the specificity of its evidence, even after the qualifications are counted.** A disclaimer that a model is incomplete does not establish the size or causal importance of the factors it retains. Acknowledging some reasons to quit does not demonstrate that stable top-level goals are generally optimal.

Six constructs that must not be collapsed

Construct	Meaning in this review
Grit as measured	A self-report composite of perseverance of effort and consistency of interests; the book displays ten items, unlike the original twelve-item and short eight-item research forms. [B, ch. 4][2][3]
Perseverance of effort	Reported diligence and continued effort despite obstacles. It is not the same as observed hours, skill, or strategic effectiveness.
Consistency of interests	Reported stability of interests and pursuits. Duckworth expressly uses "passion" in this sense, rather than simply emotional intensity. [B, ch. 4]
Deliberate practice	Structured attempts to improve a specific skill, with focused effort, informative feedback, and revision. A behavior, not a personality score. [B, ch. 7]
Achievement versus retention	Doing better or producing valuable work versus remaining in a course, job, or training program. Staying may enable achievement, but is not itself proof of superior performance.
Grit-building intervention	A particular educational or behavioral package. Its label does not establish what changed, why it changed, or whether effects transfer beyond the measured setting.

The steelman is compelling: many valuable achievements require repeated, well-directed effort, and people can sometimes change the practices or circumstances that sustain it. The stronger proposition—that the grit construct uniquely explains exceptional success and that the book teaches readers to reliably increase it—is the one that needs more evidence.

The complete notes add an important strength: **evidential candor.** Duckworth flags adapted measurement, unpublished studies, competing findings, and active-control limits. A fair critical review should not flatten those disclosures into a uniformly overconfident book. The relevant issue is whether the main argument and recommendations remain proportionate after those caveats are taken seriously. [N04-01, Notes to ch. 4]; [N03-15, Notes to ch. 3]; [N09-40, Notes to ch. 9]; [N13-01, Notes to ch. 13]

04 Three central scientific claims

These are the three rated propositions, stated at the level assessed—not three verbatim quotations or an exhaustive list of the book's claims.

CLAIM 1 · 50%

Grit is a powerful general predictor of achievement, sometimes more important than talent.

There is predictive signal, but it depends on the outcome and comparison. Retention, performance, and elite achievement are not interchangeable; a grit score is not a clean causal estimate.

CLAIM 2 · 75%

Sustained, high-quality effort builds skill and helps turn skill into accomplishment.

Purposeful practice and persistence have a meaningful evidential basis. The book's equations are explanatory metaphors, not estimated production functions proving that effort has exactly twice the importance of talent.

CLAIM 3 · 50%

The recommended assets and environments can build grit and improve long-term success.

Some educational interventions help in particular contexts. That does not validate a general recipe for changing a unitary, durable trait across domains. The complete notes acknowledge more uncertainty than a simple dismissal would suggest.

Three central scientific theses

Thesis A · Grit is a powerful general predictor of achievement, sometimes more important than talent

50% · Weak support for the broad framing [B, Preface; chs. 1–3]

The narrow proposition that a persistence-related measure predicts some consequential outcomes is credible. The broader proposition conflates different outcomes, selected samples, and comparator measures. West Point admission scores are not pure innate ability; surviving an initiation period is not the same outcome as academic performance or military excellence. The book itself notices that distinction, but its rhetoric encourages a larger conclusion. [B, ch. 1]

The later synthesis of academic studies and the direct comparison with established personality measures do not support a dominant general predictor. The large subsequent West Point investigation instead supports outcome-specific contributions. The score recognizes genuine predictive signal; it withholds credit for generalized superiority and for an unestablished causal ranking of effort against talent. [5][6][7]

Thesis B · Sustained, high-quality effort builds skill and converts skill into accomplishment

75% · Moderately supported, not a quantitative law [B, ch. 3; ch. 7]

This is the strongest of the central theses. The book's distinction between merely accumulating experience and intentionally improving a weakness is valuable. Evidence on practice and trials that encourage better study behavior supports important parts of the argument. It does not establish that one practice method has no rival in every domain, that amount of practice explains nearly all differences, or that effort has exactly twice the causal importance of talent. [13][14][16]

The two equations in ch. 3 express an explanatory metaphor. Algebraically substituting one into the other produces a squared effort term only because the same symbol is placed in both equations. That is not an estimated production function. Training time, performance time,

practice quality, prior skill, changing returns, and opportunity are not separately identified. The underlying two-stage insight survives; the numerical slogan does not become a measurement result.

The complete endnotes acknowledge the major deliberate-practice critique and supply the author's preferred response. The 75% judgment is about the strength of the proposition, not an allegation that contrary research was uncited. [N03-15, Notes to ch. 3]

Thesis C · The recommended psychological assets and environments can build grit and improve long-term success

50% · Promising components; unvalidated general package [B, ch. 5; chs. 7–9; ch. 13]

It would be wrong to say there is no intervention evidence. Some randomized educational programs produce meaningful improvements. But effects vary, packages differ, and a better test score does not necessarily mean that long-term interest consistency or a durable general personality trait increased. Later large-scale findings include positive, mixed, and null achievement results. [10][11][12]

The conclusion warranted is conditional: particular practices and curricula can help some people in particular settings. A reliable recipe for growing a domain-general grit trait, with durable benefits across work, school, relationships, and elite accomplishment, has not been established by the evidence examined here. Parenting and extracurricular claims especially rely on a chain of plausible but incompletely tested links.

> **Why the score is neither a dismissal nor an endorsement**
> The book earns credit for a real behavioral problem, useful distinctions, some meaningful prospective findings, and several plausible or experimentally supported practices. It loses credit when those findings are expanded into claims about a unified trait, causal primacy, or general trainability.

05 Chapter-by-chapter assessment

Chapter-by-chapter assessment

This chapter map preserves the book's organization. Detailed evidence and methodological arguments are in Sections 04–08 and the linked claim audit; the judgments below are an editorial synthesis, not an independent numerical scoring of each chapter.

1 / Showing Up

A good research question; too sweeping a rhetorical answer. [B, ch. 1]

The chapter turns an interesting puzzle—why apparently qualified people leave demanding settings—into a research program. Measuring grit before later outcomes is a real methodological strength. It also honestly notes that admission scores predict later performance even when they do not predict early attrition. The weakness is the transition from that particular contrast to the impression that grit is what fundamentally matters for success. Different outcomes require different models. Read the examples as demonstrations of potentially useful incremental information, not as a contest that intelligence has lost.

2 / Distracted by Talent

A valuable corrective to fatalism, not evidence for a causal ranking. [B, ch. 2]

The personal histories illustrate how an early test, label, educational placement, or first impression can miss later potential. They do not establish that talent is generally secondary. A success story selected after the outcome cannot tell us how many similarly persistent people did not succeed or which alternative explanation best fits. The naturalness-bias experiments are conceptually separate: a bias in how observers evaluate an identical performance is not evidence about what actually produces superior performance. The chapter is strongest when opposing premature ceilings and weakest when illustrative biographies carry the explanatory burden.

3 / Effort Counts Twice

A memorable conceptual distinction dressed in overly suggestive mathematics. [B, ch. 3]

The separation of talent, acquired skill, and productive accomplishment is useful. The explicit admission that opportunity may matter more than psychology is also substantial, not a token disclaimer. But the equations do not estimate causal contributions. The treadmill, potter, writer, and performer narratives illuminate persistence without identifying how much effort matters relative to ability, circumstances, or strategy. The chapter's actionable insight is to work both on improving capability and on using it. Readers do not need to accept a literal squared-effort law to benefit from that distinction.

4 / How Gritty Are You?

The most useful strategic distinction; unresolved measurement issues. [B, ch. 4]

The hierarchy of goals allows flexibility without abandoning an important purpose. This is much better advice than literal stubbornness. At the same time, the questionnaire mixes generalized self-descriptions with an interpretation about sustained devotion that deserves more validation. The score table looks precise, but cannot establish a person's potential or prescribe how long to continue a venture. The chapter's historical estimates of eminent people's childhood IQ should not be read as actual contemporaneous tests. Use the goal hierarchy as a planning tool, and the scale as a conversation starter rather than a credential.

Endnote assessment: the ten-item adaptation and norm source are traceable. That improves documentation but does not make the score a validated credential. [N04-01, Notes to ch. 4]; [N04-02, Notes to ch. 4]

5 / Grit Grows

Good caveats about nature and nurture; a gap between change and intervention. [B, ch. 5]

The text explains why genetic influence need not make a trait unchangeable and explicitly distinguishes cross-sectional age differences from longitudinal growth. Those are strengths. Yet the final confidence about cultivating grit outruns what those designs can show. Twin residuals are not simply a collection of usable parenting levers; cohort differences are not a treatment effect. The four psychological assets organize the rest of the book attractively, but their classification does not itself validate a developmental sequence or a joint intervention. The decisive evidence must come from well-measured change and credible comparisons.

Endnote assessment: the dated Plomin communication limits what can fairly be inferred about the author's access to later twin findings. The separate count of 697 "genes" should be corrected to variants. [N05-16, Notes to ch. 5]; [N05-18, Notes to ch. 5][42]

6 / Interest

One of the strongest practical chapters, with some exaggerated magnitude language. [B, ch. 6]

The idea of fostering rather than discovering a predestined passion reduces an unhelpful all-or-nothing search. Exploration, repeated contact, encouraging people, and growing expertise are more actionable than waiting for certainty. The text also allows changing course and acknowledges economic constraints. Its scientific weakness is moving from associations involving interests and fit to very strong statements about satisfaction or vocational destiny. A reader should preserve the chapter's experimental spirit: investigate the daily activity, not merely its title; allow interest to develop, but do not treat indefinite commitment as the price of legitimacy.

7 / Practice

Useful skill-development advice; strongest claims need boundaries. [B, ch. 7]

The distinction between time served and deliberate improvement is excellent. So are the practical requirements for a targeted goal, feedback, reflection, and repetition. Importantly, the book calls the 10,000-hour figure a rough average and includes rest and uncertainty about flow; criticizing it for denying those points would be unfair. The problems are universal language and the inference from observed practice histories to causal sufficiency. The classroom intervention description is more persuasive than the celebrity examples. Implement a good learning loop without assuming that one practice taxonomy explains exceptional performance everywhere.

Endnote assessment: the practice debate and the value of quizzing are explicitly acknowledged. The remaining criticism is overgeneralization, not the absence of those qualifications. [N03-15, Notes to ch. 3]; [N07-24, Notes to ch. 7]

8 / Purpose

A plausible source of motivation, not an obligatory moral ingredient. [B, ch. 8]

Connecting work to other people can organize effort, and some of the described academic interventions go beyond correlation. But the large purpose–grit survey cannot show that purpose creates grit, while cases of devoted workers may reflect selection and retrospective meaning-making. The book recognizes self-interest, ordinary paid work, and even harmful grit; these qualifications matter. The job-crafting case is corroborated through researcher accounts and a later final report. Its field assignment was workshop-level, with mixed short- and long-term outcomes rather than sustained performance gains. The best takeaway is to look for genuine meaning and useful contributions, while remembering that “calling” is not evidence that unhealthy sacrifice or poor working conditions are acceptable.

Endnote assessment: the purpose result is verified narrowly as time per review question; the final job-crafting report specifies workshop allocation and outcome-specific benefits. [17] [43] [44] [45] [54]

9 / Hope

An effective account of agency that bundles distinct evidence bases. [B, ch. 9]

The chapter combines learned helplessness, explanatory style, therapy, growth mindset, brain plasticity, and resilience. These bodies of work cannot simply certify each other. A brain that changes with learning does not demonstrate that a particular message raises general intelligence, and a rodent follow-up does not show lifelong human protection. The text is commendably explicit about harm from uncontrollable adversity and the value of professional help. Preserve its practical emphasis on trying a new strategy, seeking support, and interpreting a setback accurately; do not elevate optimism into a substitute for evidence about the situation.

10 / Parenting for Grit

Thoughtful guidance presented with some appropriate uncertainty. [B, ch. 10]

The chapter's admission that parenting-specific grit evidence is incomplete is important. Its warmer, autonomy-supporting interpretation is more defensible than the severe anecdotes sometimes extracted from it. The key scientific question is what the parenting behavior causes, not whether an accomplished adult remembers it favorably. Child effects on parents and shared environments complicate the story. The feedback research is relevant but narrower than a whole parenting philosophy. The most reasonable use is to combine expectations with credible support and the child's agency, while evaluating the actual response rather than enforcing a branded formula.

11 / The Playing Fields of Grit

A strong case for opportunity; an incomplete case for generalized transfer. [B, ch. 11]

The chapter recognizes both selection and cultivation, as well as unequal access to good activities. Those admissions should prevent a causal reading of simple participation–success associations. Past follow-through may contain useful information, but the Grit Grid also incorporates achievement and opportunity. The warning about faking self-reports in high-stakes settings is especially valuable. The household Hard Thing Rule is an interesting practical proposal, not a proven intervention. Offer activities because their direct benefits and quality warrant them; look for transferable habits, but do not promise that attendance will manufacture a general personality asset.

12 / A Culture of Grit

Environment matters; the chapter cannot isolate "culture of grit" as the cause. [B, ch. 12]

The sports, military, and business examples contain many simultaneous ingredients. The reader cannot separate recruitment, coaching, money, incentives, trust, and persistence from narrative alone. The West Point account of reduced attrition despite stable incoming grit scores is a particularly valuable counterweight to individual-focused explanation. Norms that encourage continued learning may help, but norms against complaint may also make useful negative information harder to express. The central managerial task is to create conditions for effective work and honest feedback, not to make employees perform an identity of relentless toughness.

13 / Conclusion

The ethical qualifications are stronger than the universal-growth promise. [B, ch. 13]

The ending deserves credit for recognizing opportunity, other virtues, reasons to quit, and uncertainty about family costs. It explicitly does not make grit the highest moral quality. However, wanting more grit is not evidence that increasing it will benefit almost everyone, and cross-sectional satisfaction data cannot establish safety. The final invitation to develop interest, practice, purpose, and hope is best read as an agenda for action and testing. It should not be mistaken for a completed causal account of exceptional achievement or a validated intervention package for all readers.

06 Claim, number, and source audit

Quantitative evidence and its limits

These are selected quantitative anchors, not a new pooled estimate. The studies use different samples, outcomes, instruments, and designs; some reviews include overlapping primary studies. Their sample sizes must not be added together as though they were independent replications.

Evidence	Verified anchor	Interpretation
Original grit studies, 2007 [2]	The paper describes an average of about 4% of outcome variance across its studies.	Evidence of an association, not 4% of every person's success or a universal causal contribution.
Credé et al. synthesis [4]	88 independent samples; N = 66,807. Grit–conscientiousness: observed r = .66; corrected ρ = .84.	A large overlap. The corrected estimate is not a raw individual-level correlation; neither statistic alone proves complete equivalence.
Rimfeld et al. twin study [5]	4,642 adolescents / 2,321 twin pairs. Big Five traits explained 5.6% of GCSE variance; the grit facets added 0.5 percentage points.	Incremental prediction was small in this setting. Phenotypic analyses used one twin per pair, so 4,642 is not the independent regression sample size.
Lam & Zhou synthesis [7]	137 studies, 156 dependent samples; N = 285,331. Weighted r: total grit .19; perseverance .21; consistency .08.	Modest average academic associations; effort perseverance carried more signal. Abstract-level extraction; not an audit of all included studies.
Alan et al. classroom experiments [10]	About 0.2 SD on a standardized math test at a 2.5-year follow-up.	A genuine favorable experimental result for a particular program. Not an estimated effect of reading Grit or of changing its whole scale by one point.
North Macedonia manuscript [11]	35,340 students in 352 schools. Small average later GPA effects; up to .28 SD for the disadvantaged Roma subgroup.	The subgroup result is not the population average. The manuscript also reports divergent movement of the grit facets.

Evidence	Verified anchor	Interpretation
England working paper, 2026 [12]	100 schools. Mindset beliefs +.417 SD (95% CI .27 to .57); no detected effect on national achievement tests.	A belief shift did not guarantee a test-score benefit. Working-paper evidence; “2026” is the paper’s publication year, not an assertion about when the trial was conducted.

Two additional checks on the book’s supporting theories

Research question	Quantitative result	Boundary
Practice and performance across domains [13] [48]	Read with its 7 May 2018 publisher corrigendum, the 2014 synthesis reports variance accounted for of 24% in games, 23% in music, 20% in sports, 5% in education (r = .22), and 1% in professions; the corrected overall figure is 14%. The professions estimate remains nonsignificant.	These are associations under the review’s practice definitions, not causal shares of achievement. The corrigendum corrects dependent-effect sample-size handling and leaves the substantive conclusions unchanged. The unexplained remainder is not automatically talent, and a small association does not mean practice can be removed without consequence.
Mindset interventions: broad and quality-restricted estimates [20]	Macnamara and Burgoyne: overall d = .05 (95% CI .02 to .09); selected higher-quality subset d = .02 (95% CI −.06 to .10).	Quality restrictions and bias analyses affect the answer. These estimates do not establish a zero effect in every context.
Mindset interventions: targeted, high-fidelity subset [21]	Burnette et al.: academic d = .14 (95% CI .06 to .22); 95% prediction interval −.08 to .35.	This is a conditional subset, not the same estimand as an overall mean. The prediction interval concerns heterogeneity across settings; it is not a confidence interval for the mean.

Calculation 1 · Correlation is not a causal percentage

Squaring .19 gives .0361, or 3.61%. In a simple single-predictor linear model, r^2 is the proportion of outcome variance accounted for statistically. Here it is only an illustration based on a reported meta-analytic correlation—not a newly estimated pooled R^2, an incremental effect after other predictors, or an intervention effect. Squaring .21 and .08 similarly gives 4.41% and 0.64%; those facet figures cannot be added because the facets overlap. [7]

Small between-person associations can coexist with an important behavior. A behavior everyone performs can be necessary yet explain little variation. Conversely, a predictor can correlate with success because successful people acquire it, because both share causes, or because the measure partly redescribes the outcome. The appropriate response is not to equate small r with uselessness, but to ask a more specific causal and practical question.

Calculation 2 · Odds are not probabilities

Hypothetical example—not a reanalysis of any book study: Suppose 95% of a comparison group completes a program, and a predictor has an odds ratio of 1.50. Applying that odds ratio gives $p_1 = (1.50 \times .95) / (.05 + 1.50 \times .95) = .9661$, or 96.61%. That is a 1.61-percentage-point difference, not a 50-percentage-point difference and not 50% more people completing the program.

Such a difference can still matter at scale. But a statistically reliable association in a high-retention sample is not a magic classifier of who will quit. An organization needs discrimination, calibration, false-positive costs, and incremental value—not just a

significant odds ratio. This report does not attribute the hypothetical numbers above to Duckworth.

Calculation 3 · Overlap is not an identity test

The observed grit–conscientiousness correlation .66 implies 43.56% shared variance in that simple bivariate representation. The corrected .84 implies 70.56% at the disattenuated level under its assumptions. Neither percentage means that the remaining part has been proven new, causally important, or useful. Compare equally reliable facet measures and test incremental prediction on new data before claiming a distinctive instrument. [4]

Calculation 4 · A task ratio is not general behavioral change

The purpose experiment's raw time-per-question ratio is 49 / 25 = 1.96. That calculation supports "roughly twice" for the measured task. It does not tell us that students studied twice as much outside the experiment, learned twice as much, or became twice as gritty. The unit of measurement must travel with the headline number. [17]

Twenty-claim evidence audit

The following judgments concern the stated scope of each claim. "Supported" does not imply causation; "not established" does not imply false. Confidence describes confidence in the review's assessment, not a probability assigned to the author's proposition.

Grit predicts persistence in demanding settings

Supported, with restricted scope. Confidence: High. [B, ch. 1]

The book's opening cases ask whether people stay through difficult training, remain employed, finish school, or advance in a spelling competition. These are meaningful outcomes, and prospective measurement is stronger evidence than retrospective stories about famous people. The military, sales, and school research provides credible evidence of predictive signal. [2][27]

The claim becomes weaker when "predicts retention" becomes "explains success." Leaving can reflect poor fit, changed opportunities, health, finances, or an informed choice. Staying can be necessary for one institution's goal without being optimal for the person. A retention predictor must therefore be evaluated against both performance and the value of remaining.

More defensible formulation: Use persistence measures as possible context-specific predictors—not as verdicts on future excellence or character.

Grit generally matters more than talent

Not established as a general ranking. Confidence: High. [B, Preface; ch. 1; ch. 3]

The comparison changes with the outcome. The book acknowledges that West Point's admission composite predicts subsequent performance even when it does not distinguish early leavers. The later large West Point study likewise distinguishes completion from performance. [6]

An admissions-selected sample cannot establish that ability is unimportant in the population. Selection compresses the range of ability and may induce associations among the characteristics used to select people. Moreover, an admission score blends past achievement, opportunity, and assessed attributes; it is not a clean measurement of innate talent. "Not sufficient" never implies "less important."

More defensible formulation: Ask which abilities and behaviors predict the specific outcome, in the specific population, at the specific stage.

Grit is a substantially distinctive psychological explanation

Distinctiveness is overstated in the practical framing. Confidence: High. [B, ch. 1; ch. 4]

The relevant comparison is not grit versus "nothing," but grit versus established constructs such as conscientiousness and its achievement-oriented facets. The overlap and limited incremental prediction summarized in Section 06 create a serious novelty problem. [4][5]

This is not a claim that Duckworth invented the persistence idea or explicitly positioned grit as a sixth Big Five trait; the original scientific work situates it near conscientiousness. Nor must every useful scale measure a wholly new trait. A narrow measure can be convenient. But convenience, branding, theoretical novelty, and incremental utility are different achievements.

More defensible formulation: Treat grit as a useful label or candidate facet until it adds reproducible value beyond well-measured existing traits.

Passion is adequately captured by stable interests

Measurement does not fully match the richer narrative. Confidence: High. [B, ch. 4]

Duckworth explicitly defines the scale's passion component in terms of consistency, not intensity. That makes the usage transparent, but it does not establish that avoiding changes of interest measures deep enjoyment, purpose, or devotion to a particular goal. Those meanings are frequently adjacent in the narrative.

Jachimowicz and colleagues proposed a separate passion-attainment measure as a moderator of persistence-related performance. That is a candidate extension, not retrospective validation of every original scale interpretation. Duckworth and colleagues' separate 2021 commentary acknowledges a mistaken inference from the earlier factor model and calls for improved measurement. [8] [9]

The 2019 primary exchange makes the passion-attainment interpretation contested. Credé questioned describing the whole Grit–S score as perseverance; he did not have the Study 2 data and did not recover the proposed perseverance interaction in Study 3. Guo, Tang and Xu obtained the raw data for Studies 2 and 3 and modeled perseverance and interest consistency separately. They found no perseverance-by-passion interaction; an interest-consistency-by-passion interaction remained for Study 3 grades. These are different analytical scopes, not one identical failed participant replication. [49] [50]

Jachimowicz and colleagues replied that the apparent subfactors reflected positive versus negative item wording and defended interpreting the whole scale as perseverance rather than passion. They also reported new factor-structure analyses of revised items: an all-positive version in 958 participants and an all-negative version in 781. These are new measurement samples, not a new test of the original performance interaction. The reply supplies a counterargument, not a settled resolution. This report did not execute either side's participant data or code. The initial reply-access limit and the further full-text check are recorded at reference 51. The disagreement warrants separating effort, consistency, and attained passion; it is not a retraction or proof that every original association is false. [51]

More defensible formulation: Measure sustained effort, interest stability, enjoyment, goal importance, and actual commitment separately when the distinction matters.

A brief grit score can tell readers how gritty they are

Reasonable reflection tool; weak basis for consequential classification. Confidence: High. [B, ch. 4; ch. 11]

The questionnaire has a research history, including reliability and predictive-validity work. But the ten-item book instrument is not identical to either the original twelve-item form or the eight-item Grit–S. Evidence for one version should not automatically be transferred to every altered version, norm table, population, or use. [2][3]

No score in the book is a diagnosis or a direct observation of years of behavior. Self-presentation, interpretation, comparison standards, and response style can matter. Duckworth is notably candid later in the book that high-stakes self-report is easy to fake; her separate measurement paper also warns against consequential uses of inadequate measures. [26]

A later large empirical test supports the comparison-standard caution. Lira and colleagues studied 229,685 adolescents across three studies of self-regulation: Studies 1 and 3 included grit measures, while Study 2 measured conscientiousness. In Study 3, grit self-reports predicted college graduation positively within high schools but negatively across schools. Different local standards can change how people judge themselves. That is a specific aggregation and policy-use problem, not evidence that all within-setting grit prediction is invalid. The underlying models were not independently executed here. [57]

The complete notes remove a provenance uncertainty: Duckworth explains that the book's ten-item form shortens the original twelve-item scale, reports a .99 correlation between them, and identifies a wording expansion. The norms point to the original 2007 adult study. That documentation deserves credit. The .99 figure is a reported validation claim, not a raw-data result independently reproduced here; it does not establish current norms, clinical thresholds, or validity in consequential selection. [N04-01, Notes to ch. 4]; [N04-02, Notes to ch. 4]

More defensible formulation: Use the items to prompt reflection. For hiring, admissions, or evaluation, require purpose-specific validation and behavioral corroboration.

Effort counts twice

Useful metaphor; quantitative interpretation unsupported. Confidence: High. [B, ch. 3]

The two-stage distinction is sensible: work can improve capability, and further activity can put capability to productive use. The equations in chapter 3 make this memorable. They are not estimates obtained by fitting data, and the book provides no scale on which one unit of talent can be exchanged for one unit of effort.

Combining the equations assumes that the same effort term can stand in for two potentially different activities. It also leaves out how returns change with fatigue, technique, prior knowledge, and time. The equation cannot tell a reader to double effort rather than change strategy, improve selection, or buy better instruction. This is a logical assessment of the displayed model, not an empirical claim that effort has no effect.

More defensible formulation: Effort can contribute at multiple stages. Its marginal value must be learned from the task, not inferred from a slogan.

Deliberate practice is the unmatched route to improvement across domains

Useful principles; universal superiority not demonstrated. Confidence: Moderate to high. [B, ch. 7]

The book correctly distinguishes purposeful refinement from repeating familiar work. However, observational practice–performance research cannot establish universal dominance over alternative learning methods. The cross-domain synthesis and violin replication constrain strong explanatory claims without showing that practice is dispensable. [13][14]

For a well-defined skill, a coach can often identify an error and supply fast feedback. In entrepreneurship or open-ended research, the target and feedback may themselves be uncertain. A "practice more" prescription can then optimize the wrong routine. Distinguishing the amount, design, and quality of practice is essential; a small between-person practice correlation is not evidence that training has little causal value.

The endnotes explicitly cite the 2014 Macnamara–Hambrick–Oswald meta-analysis and direct readers to Ericsson's response. The critique is therefore not absent from the book's documentation. The remaining objection is the strength of the main-text synthesis: acknowledging a dispute in a note does not resolve it or establish universal superiority. [N03-15, Notes to ch. 3][13]

More defensible formulation: Use targeted practice where the skill and feedback are valid. Combine it with exploration, knowledge acquisition, and real-world outcome testing.

Practice explains why gritty spellers perform better

Consistent with mediation; causal mechanism not isolated. Confidence: High. [B, ch. 7]

The spelling-bee research links grit, preparation, and later performance; solitary study was more prognostic than the other preparation categories examined. This is a useful process hypothesis and more informative than an anecdote. [15]

But a statistical mediation pattern does not randomize grit or practice. Prior skill, support, resources, and motivation could influence both. The book deserves credit for considering alternative explanations of why gritty spellers enjoy practice more and for recognizing quizzing as a way to identify weaknesses. That nuance should remain part of the lesson.

More defensible formulation: The findings suggest a plausible route from sustained effort to better preparation; they do not independently prove the entire causal chain.

Age trends and heritability establish that grit can be deliberately grown

Malleability plausible; these particular inferences are insufficient. Confidence: High. [B, ch. 5]

The twin estimates describe variation within a population under the study's conditions, not an individual's fixed genetic percentage. The book is right that heritability does not imply immutability. However, residual variance also includes measurement error; calling everything not genetic "experience" is too neat. [5]

The age graph is a cross-sectional snapshot. Duckworth explicitly says it cannot distinguish maturation from generational differences, which is good scientific communication. The later transition to practical malleability is less secure. A trait varying across ages or contexts does not establish that a specified intervention will move it durably in the intended direction.

The source trail also changes the fairness assessment. The two grit heritabilities are attributed to a personal communication from Robert Plomin dated 21 June 2015. The completed Rimfeld paper is used here to assess the broader theory, but the notes do not establish that the author possessed its final results when drafting this passage. Deliberate omission of that paper's less favorable findings should not be inferred. [N05-16, Notes to ch. 5][5]

More defensible formulation: Separate evidence that grit differs, evidence that it changes within people, and evidence that a program causes beneficial change.

Developing interests improves satisfaction and performance

Direction supported; magnitude and guarantees overstated. Confidence: Moderate to high. [B, ch. 6]

Interest–environment fit is relevant to performance and persistence. Later satisfaction evidence also supports an association, but not the enormous difference suggested by the book's strongest wording. The numerical estimate appears in the targeted audit. [22][23]

The chapter's best practical correction is to foster interests through contact and repeated engagement rather than wait for a predestined calling. Nevertheless, liking a topic, liking the daily work, earning a living, and having comparative advantage are separate considerations. A person can test these rather than committing to an identity on the strength of initial enthusiasm.

The actual satisfaction reference is Mark Allen Morris's 2003 dissertation. Its underlying results were not retrieved, so the current meta-analytic estimate cannot certify what Morris originally found. A separate life-satisfaction statement cites work described as a manuscript in preparation. These are distinct evidential statuses. [N06-06, Notes to ch. 6]; [N06-07, Notes to ch. 6][31][23]

More defensible formulation: Explore promising activities, then evaluate the actual tasks, learning trajectory, constraints, and opportunities—not "passion" alone.

Purpose sustains grit and energizes performance

Some experimental support; necessity and long-term generality unproved. Confidence: Moderate. [B, ch. 8]

The large self-report survey connects grit with a prosocial orientation, but cannot establish which causes which. Separate purpose interventions provide stronger evidence for some academic behaviors and outcomes. These are not interchangeable forms of support. [17]

The book recognizes that gritty villains exist and that prosocial purpose is not an absolute requirement. That qualification matters because a moralized account could otherwise reclassify ordinary paid work as deficient or normalize excessive sacrifice. Finding meaning can be helpful without making a person's value dependent on devotion to work.

The book's study-time example has a precise source basis. In the original purpose experiment, the relevant outcome was time per exam-review question: 49 versus 25 seconds in a small undergraduate experiment. This corroborates a near-doubling on that task, not a doubling of general study effort or a durable increase in grit. [N08-44, Notes to ch. 8][17]

The job-crafting example has an accessible final original report, published online in 2022 and in a 2023 issue. It improves the source and design assessment but reports outcome- and duration-specific benefits, not a purpose-to-durable-grit effect. The workshop allocation, separate randomized study, attrition, and distinct data-access conditions are detailed in targeted check 11. [54]

More defensible formulation: Connect difficult tasks to a purpose the person actually endorses; do not substitute purpose language for autonomy, resources, or fair treatment.

Growth mindset reliably produces grit and achievement

Effects are conditional; the broad causal sequence is not established. Confidence: High on limits; moderate on boundary conditions. [B, ch. 9]

The chapter links beliefs, interpretations of setbacks, and persistence. That is a plausible model, not a fully identified pathway. Later research includes a well-designed national experiment with benefits for selected outcomes and groups, as well as competing meta-analytic interpretations. [19][20][21]

The disagreement is partly about study quality, aggregation, target populations, and implementation—not simply whether believing in change is good. Effects on achievement are generally much smaller and more conditional than a motivational reading of the chapter implies. Changing an answer on a mindset questionnaire is also not the same endpoint as acquiring skill.

The notes on the Penn Resiliency Program explicitly acknowledge the absence of an advantage over active comparison conditions in a cited meta-analysis. The review must credit that disclosure rather than treating the documentation as uniformly positive. It does not independently establish that the entire hope-to-grit-to-success pathway has been tested. [N09-40, Notes to ch. 9]

More defensible formulation: Treat growth-oriented feedback as a potentially useful support for effective learning, not a substitute for effective learning.

Mastering adversity can inoculate people for life

Bounded mechanism plausible; lifespan and human extrapolation overreach. Confidence: High on the extrapolation problem. [B, ch. 9]

The book describes adolescent rats exposed to controllable stress and tested again in adulthood, roughly five weeks later. The complete official NIH accepted manuscript of Kubala and colleagues' 2012 paper was subsequently read, including methods, results, discussion and captions. Male rats received controllable or yoked uncontrollable tail shocks at postnatal day 35; later tests examined social exploration and response to stress, not a human grit trait or

the animals' entire lifespan. This resolves the earlier method-access limit but not the temporal or human extrapolation. The manuscript was accepted and unedited, not independently line-compared with the final typeset version. [N09-34, Notes to ch. 9] [41]

Nor would a rodent stress paradigm establish that deliberately exposing children to hardship increases general grit. The book expressly warns that uncontrollable adversity can weaken people. Its human advice is most defensible when the emphasis is on support and attainable mastery, not adversity for its own sake.

More defensible formulation: Provide supported challenges and real routes to influence the outcome. Do not infer a license to manufacture trauma.

Cognitive behavioral therapy has longer-lasting effects than antidepressants

A defensible narrower comparison is presented too broadly. Confidence: Moderate to high. [B, ch. 9]

An original randomized-treatment follow-up supports enduring effects of cognitive therapy relative to withdrawal from medication among treatment responders. It did not establish a universal advantage over continued medication. The comparator and treatment stage are central, not technical footnotes. [25]

This report evaluates the wording, not an individual's treatment choices. It does not infer a clinical recommendation to stop or change medication. Duckworth's recommendation to seek professional help rather than expect a short chapter to reverse entrenched problems is an important safeguard.

The endnote itself supplies both the acute-treatment comparison and the relapse follow-up. That is better documentation than an undifferentiated appeal to cognitive therapy; the reader still needs to distinguish medication discontinuation from continued treatment. [N09-14, Notes to ch. 9][25] [46]

Favorable later evidence should also count. Voderholzer and colleagues' 2024 synthesis of 19 randomized follow-up trials supports enduring psychotherapy benefits after treatment termination, with limited studies, heterogeneity, and comparator-dependent conclusions. That is consistent with the narrower enduring-effect proposition; it does not establish universal superiority over continuing medication. The primary abstract and indexed comparator/quality discussion were read, not the original clinical datasets or a re-estimation of the synthesis. [60]

More defensible formulation: Psychological treatment can have durable benefits, but comparisons depend on condition, treatment quality, continuation, and the outcome being measured.

Warmth plus high standards produces gritty children

Plausible and often sensible; direct grit-specific causation not shown. Confidence: Moderate. [B, ch. 10]

The chapter explicitly acknowledges that direct research on parenting for grit was not yet sufficient and identifies some recommendations as hunches. That restraint deserves credit. Its narratives cannot, however, isolate parenting effects from children's predispositions, reciprocal influence, shared circumstances, or selection into the examples.

The wise-feedback experiments support a specific combination of substantive criticism, credible standards, and confidence in a student's capacity under particular conditions. They are not a randomized validation of an entire parenting style or a proof of increased general grit. [18]

A later randomized undergraduate test by Troy, Moua and Van Boekel found no wise-feedback trust advantage among 94 completers when both conditions began and remained high in trust. The population, baseline trust, and precision differ from the original school experiments about racial mistrust. This is a contextual null result, not conclusive disproof of those original effects, and neither study is a parenting-for-grit trial. Access was to the complete primary abstract and data-availability statement; the article's full methods and requested data were not obtained. [53]

The wider parenting evidence is not uniformly adverse: Pinquart and Kauser's synthesis of 428 studies found generally favorable associations for authoritative parenting, with cultural and ethnic variation. Those associations support the plausibility of warmth, standards, and autonomy; they do not isolate a parenting intervention that causes general grit. Current access was to the primary abstract, not complete study-level coding or a new causal analysis. [63]

More defensible formulation: Combine warmth, meaningful expectations, skill support, and age-appropriate agency. Do not generalize a celebrated person's harsh-family anecdote into a rule.

Extracurricular follow-through develops transferable grit

Predictive evidence stronger than causal transfer evidence. Confidence: High. [B, ch. 11]

The chapter repeatedly notes the selection problem: activities may cultivate perseverance, but already persistent children may also stay in them. Its follow-through and Grit Grid examples are useful indicators of past behavior, not clean experiments on the effect of participation. The grid also rewards achievement, making it more than a pure persistence measure.

A credible causal test would distinguish participation, coaching, social connection, prior traits, and access to resources; then measure transfer to new tasks and later outcomes. A child continuing an activity because it is enjoyable or well-supported has not thereby demonstrated general resilience across life domains.

The Grit Grid's college-persistence data are cited to a manuscript described as in preparation. This identifies the source of the book's numbers, but is not equivalent to an accessible final article with methods and estimates open to inspection. The notes also acknowledge the strong predictive value of standardized tests; they do not simply dismiss those measures. [N11-12, Notes to ch. 11]; [N11-07, Notes to ch. 11]

More defensible formulation: Offer good activities for their direct learning and social value. Treat generalized character transfer as a hypothesis, not the guaranteed return.

The Hard Thing Rule is an evidence-based grit intervention

Reasonable family practice; the particular package is unvalidated. Confidence: High about the evidence limit. [B, ch. 11]

The rule combines a demanding activity, a commitment to a natural stopping point, and personal choice; later, the commitment period lengthens. It is not a literal prohibition on quitting. Those details make it more reasonable than a blanket persistence mandate.

But there is no controlled evaluation of this household rule in the material examined here. A fixed duration might support follow-through in one case and obstruct adaptive change in another. Reasonable safeguards include exceptions for harm, poor instruction, financial strain, and a seriously mistaken fit. These safeguards are this review's judgment, not a tested modification of Duckworth's program.

More defensible formulation: Treat commitment periods as revisable experiments with clear review points, not as a scientifically established dose of character development.

Joining a gritty culture makes a person gritty

Plausible contextual influence; selection and mechanisms unresolved. Confidence: Moderate. [B, ch. 12]

Teams and organizations can alter what behavior is rewarded, modeled, and made feasible. Yet stories about successful cultures combine recruitment, resources, coaching, incentives, leadership, and persistence. They do not identify a single "grit culture" treatment.

The chapter's own West Point account is particularly revealing: attrition fell substantially across cultural changes while incoming self-reported grit stayed approximately unchanged. This is a narrated historical comparison, not a controlled estimate. Still, it weakens an individual-only explanation and highlights conditions, not just character.

More defensible formulation: Improve the setting's feedback, support, standards, and incentives. Evaluate the resulting behavior rather than assuming a personality transformation.

More grit is generally compatible with well-being and unlikely to have costs

Positive associations do not establish universal safety. Confidence: High. [B, ch. 13]

The conclusion reports associations with life satisfaction and acknowledges that effects on family members have not been adequately studied. Self-selected, cross-sectional reports cannot show that inducing greater persistence improves well-being or that exceptional work intensity is harmless.

Laboratory research also supplies examples of grit-associated persistence with costs. That is a warning against unconditional praise, not proof that gritty people generally make worse decisions. Because the trait was not randomly assigned, even those findings require causal restraint. [24]

The positive-emotion and life-satisfaction evidence is traced in the notes to a 2015 conference poster; the survey in which no respondent wanted less grit is described as unpublished data. These disclosures are relevant to confidence, not evidence that the results are fabricated. Neither a preference for more grit nor a favorable cross-sectional association establishes the safety of increasing it. [N13-01, Notes to ch. 13]; [N13-03, Notes to ch. 13]

More defensible formulation: Assess the full outcome: progress, health, relationships, ethics, and opportunity costs—not perseverance alone.

Stable top-level goals should anchor flexible lower-level tactics

A useful heuristic, not a demonstrated universal optimum. Confidence: Moderate. [B, ch. 4; ch. 6; ch. 13]

This is among the book's strongest distinctions. Abandoning an ineffective tactic need not mean abandoning the underlying purpose. But the level at which a goal is described is flexible: the same switch can look like quitting a company, sustaining entrepreneurship, or pursuing family security.

That flexibility can make the theory difficult to falsify. Persistent success counts as grit; successful switching can be reclassified as commitment at a higher level. A decision rule needs prospective criteria for what will remain fixed, what can change, and which evidence would justify abandoning even the high-level aim. A goal can become unethical, obsolete, infeasible, or no longer worth its cost.

More defensible formulation: Be more stable about values than about forecasts. Reconsider even important goals when their reasons no longer hold.

Reproducible sampled reference audit

Endnote source basis. The complete endnotes permit checks of the actual claim/reference links. Two selected original source passages were not retrieved. They are shown as U, retained in the sample, and not quietly replaced or assigned a failure score.

The ten entries below were selected by sorting SHA-256 hashes of a fixed seed and each note ID, then taking the first ten without replacement from the 158-entry frame. The full protocol appears in Section 10 and the audit supplement. This restricted-frame exercise is adapted from the reference-support principle, not represented as an official all-reference Red Pen Reviews audit. [1]

> **Reference-sample scope.** Ten anchors were selected from the 158-entry restricted frame; eight could be assessed and two source passages remained unresolved. A sample percentage would not be comparable with other books' purposive audits and is not used as the headline Reference Accuracy score. Each entry's evidence and limitation are stated below. The draw was reproduced; the frame was not independently recoded.

Read each assessment against the exact claim. Describing an observed age pattern can earn a strong source-support assessment even when a causal interpretation of age differences would not. Similarly, citing a theory accurately does not show that the theory's strongest practical prescriptions have been experimentally established.

01 / Genetic influence on mathematical ability **Mostly faithful; qualification needed**

N05-13 · [B, ch. 5] · [N05-13, Notes to ch. 5] · Access: A [32]

Book proposition: Some genetic differences make mathematical learning easier, illustrated by solving a quadratic equation.

What the cited source supports: Oliver and colleagues studied teacher-rated mathematics in seven-year-old twins. The original abstract supports substantial genetic influence on assessed mathematics performance; it did not directly measure the rate of learning quadratic equations.

Judgment: Moderate support for the broad genetic-influence point, with a measurement and age extrapolation. It is not evidence that an individual's future skill is fixed or that teaching cannot change it.

02 / Positive fantasies versus effective goal pursuit Mostly faithful; qualification needed

N04-18 · [B, ch. 4] · [N04-18, Notes to ch. 4] · Access: A [33]

Book proposition: Pleasant fantasies about future success, without confronting obstacles, can undermine later effort and attainment.

What the cited source supports: Oettingen's review makes this distinction and contrasts fantasies with expectations and mental contrasting. Its abstract also describes adaptive goal disengagement when success expectations are low.

Judgment: Moderate support at the stated scope. The claim should not be enlarged into "all visualization is harmful." The source's conditional endorsement of disengagement also resists a simple always-persist reading.

03 / Attributional style measurement Mostly faithful; qualification needed

N09-05 · [B, ch. 9] · [N09-05, Notes to ch. 9] · Access: A plus creator description [34]

Book proposition: A questionnaire distinguishes explanatory styles by asking respondents to explain hypothetical events and evaluate the causes.

What the cited source supports: The original article introduces the Attributional Style Questionnaire; the creator's institutional description specifies twelve hypothetical events and self-ratings on internal–external, stable–unstable, and global–specific dimensions. The book foregrounds the latter two.[47]

Judgment: Moderate support for the measurement description, not a diagnostic separation into two immutable types of person. The book's page range ends at 300; the publisher records 299. That small bibliographic error does not prevent identification.

04 / The original treadmill duration statistic U / unscored

N03-38 · [B, ch. 3] · [N03-38, Notes to ch. 3] · Access: A; needed detail unavailable [35]

Book proposition: The treadmill test's average running time was about four minutes.

What the cited source supports: Phillips, Vaillant, and Schnurr's original abstract confirms a long-term treadmill–mental-health association and statistical adjustment for fitness. It does not supply the test-duration distribution needed to verify this selected claim.

Judgment: Unscored. A general association does not verify a specific mean. The paper's overall sample of 188 and the book's test-group count of 130 are not declared contradictory: a treadmill subset could explain the difference.

---

05 / Deliberate practice is highly effortful Mostly faithful; qualification needed

N07-27 · [B, ch. 7] · [N07-27, Notes to ch. 7] · Access: T, relevant sections [36]

Book proposition: Deliberate practice is experienced as highly effortful, motivating limits and breaks.

What the cited source supports: Ericsson, Krampe, and Tesch-Römer explicitly discuss effort constraints and evidence for limited effective practice durations. This is directly relevant to the experience of demanding practice.

Judgment: Moderate support. The source does not turn the book's nearby one-hour-session and three-to-five-hour-day generalizations into invariant human physiological ceilings. Practice schedules and fatigue depend on the activity and learner.

---

06 / Naturalness bias in music judgments Faithful at the stated scope

N02-14 · [B, ch. 2] · [N02-14, Notes to ch. 2] · Access: T, original article text [37]

Book proposition: Musical experts judged an equally accomplished performer described as a natural more favorably than one described as a striver.

What the cited source supports: Tsay and Banaji's experiments manipulate achievement descriptions while controlling the underlying performer. The reported expert preference and contrast with stated beliefs support the narrated result.

Judgment: Strong support for the bounded experimental claim. This does not show that talent matters less to actual performance, nor that the effect is universal across observers or real hiring decisions.

---

07 / The age gradient in self-reported grit Faithful at the stated scope

N05-23 · [B, ch. 5] · [N05-23, Notes to ch. 5] · Access: T [2]

Book proposition: In the author's adult sample, older age groups had higher mean grit scores.

What the cited source supports: The cited 2007 paper reports the relevant age-group differences. The book calls its graph a cross-sectional snapshot and explicitly distinguishes cohort differences from maturation.

Judgment: Strong support for the descriptive sample-specific claim. This assessment deliberately does not certify the later inference that a particular intervention will make the same people grittier.

---

08 / Goals organized at different levels **Mostly faithful; qualification needed**

N04-12 · [B, ch. 4] · [N04-12, Notes to ch. 4] · Access: T [38]

Book proposition: A hierarchy of lower-order means and higher-order ends is a useful way to understand grit and self-control.

What the cited source supports: Duckworth and Gross's article presents this framework, distinguishes lower-level actions from superordinate goals, allows multiple hierarchies, and proposes questions for future testing. The note also identifies broader goal-theory sources.

Judgment: Moderate support for a clearly framed conceptual account. Citation fidelity is not experimental proof that everyone should have exactly one top-level professional goal or never revise it.

---

09 / Adversity, grit, and engagement among detectives **Partial or overextended support**

N08-39 · [B, ch. 8] · [N08-39, Notes to ch. 8] · Access: T, Study 1 methods/results [39]

Book proposition: A note gives detectives' prior victimization as an example linking personal adversity, grit, and work engagement.

What the cited source supports: The original Study 1 is a small, self-selected survey of 86 respondents, with self-reported violent victimization before entering the occupation, grit, and work engagement. A mediation model is consistent with the proposed sequence.

Judgment: Weak support for the suggested developmental pathway. Neither victimization nor grit was randomized; selection, shared causes, and reporting can account for the association. The finding does not justify intentionally exposing people to adversity.

10 / The treadmill “punishment” quotation U / unscored

N03-40 · [B, ch. 3] · [N03-40, Notes to ch. 3] · Access: M; cited page unavailable [40]

Book proposition: A quotation characterizes willingness to continue or quit the treadmill as the punishment becomes severe.

What the cited source supports: The printed “Ibid., 74” correctly resolves through the previous note to Clark W. Heath’s 1945 book. A catalog confirms the work, but the cited page was not retrieved.

Judgment: Unscored. Bibliographic existence is not quotation authentication. The note is not marked false or repaired by substituting a different publication.

> **No inflation from easy references**
> Private interviews and unpublished claims were largely outside the restricted sampling frame. That selected-subset assessment cannot establish that the entire documentation base is published, accessible, or independently reproducible. The targeted checks below include several less accessible but consequential claims.

Twenty targeted endnote checks

These twenty checks are purposive, separate from the reproducible sample, and not pooled into it. They follow the actual endnotes, include the ten narrative source checks, and examine consequential qualifications and source problems. A confirmed source identity, faithful summary, convincing study, and established causal theory are four different achievements.

01 / West Point prediction, not general superiority

[B, ch. 1] · [N01-06, Notes to ch. 1]; [N01-07, Notes to ch. 1] [2] [6]

The notes confirm the original grit papers rather than leaving the source match to inference. Prospective prediction of summer retention is supported. The larger later study distinguishes academic performance from completion. Neither the original comparison nor the later investigation makes grit a universally dominant cause of success.

02 / Sales, school, and military retention

[B, ch. 1] · [N01-12, Notes to ch. 1]; [N01-16, Notes to ch. 1]; [N01-17, Notes to ch. 1] [27]

The actual notes identify the multi-setting retention paper. Its original abstract supports the broad direction of the book's predictive summary. Exact coefficients and differences among covariate sets were not all re-extracted here. Staying employed or enrolled is a measured outcome, not automatically the best personal decision in every case.

03 / Ten-item scale and the norm table

[B, ch. 4] · [N04-01, Notes to ch. 4]; [N04-02, Notes to ch. 4] [2] [3] [26]

The note explains the twelve-to-ten-item adaptation, reports a .99 relationship, explains a wording expansion, and points to the 2007 adult data for norms. It also directs readers to measurement cautions. The first report's provenance concern is narrowed: documentation exists. Current population norms and consequential uses remain separate validation questions.

04 / The practice controversy is cited

[B, ch. 3] · [N03-15, Notes to ch. 3] [13] [14]

The notes include the major 2014 meta-analysis that limits strong deliberate-practice claims and point to Ericsson's response. This is important counterevidence to a claim of simple citation omission. The substantive disagreement about operationalization and explanatory reach survives; it belongs in the main assessment, not in an allegation that the author never acknowledged a challenge.

05 / Spelling practice and the role of quizzing

[B, ch. 7] · [N07-23, Notes to ch. 7]; [N07-24, Notes to ch. 7] [15]

The source mapping is confirmed. The paper's abstract supports an observational mediation result involving solitary preparation. The note also recognizes the diagnostic role of being quizzed and cites retrieval-practice research. It is unfair to portray the book as claiming that quizzing cannot help learning. Neither the note nor the association identifies the entire grit-to-practice-to-success causal chain.

Favorable learning evidence remains relevant. The original note cites Roediger and Karpicke's review of testing memory, not the different same-year experimental paper. Agarwal, Nunes and Blunt's later classroom synthesis reports learning benefits across tested settings; only a small fraction of its experiments were outside Western, educated, industrialized, rich,

democratic (WEIRD) settings. Current access was to the primary abstracts and the later source's data-availability page, not complete experiment-level recoding or execution of the open files. This supports retrieval as a learning technique, not the full grit-to-practice mechanism. [61] [62]

---

06 / Brief deliberate-practice interventions

[B, ch. 7] · [N07-50, Notes to ch. 7] [16]

The notes identify the intervention research rather than an unspecified study. The original experiments provide evidence for some short-term learning outcomes, especially among lower achievers. The book's subgroup qualification deserves credit. A practice-oriented intervention is not by itself a controlled test of a durable, general grit trait.

Later boundary evidence: Balan and Sjöwall's Swedish seventh-grade program combined deliberate-practice and growth-mindset instruction across eight sessions over 14 weeks. It found more deliberate-practice behavior on a test but no improvement in the measured attitudes or mathematics performance. Students belonged to unsystematically composed mentor groups, and teachers randomized those groups—not individual students—to intervention or active control. Delivery also differed from the earlier experiments. The initial check accessed the primary abstract and relevant sample/allocation passages. A further check read the publisher-formatted online article through an institutional repository, including its methods, results and limitations. This is mixed conceptual evidence, not an exact replication that disproves the original learning results. [52]

---

07 / Twin heritability: correct numbers, important timing

[B, ch. 5] · [N05-16, Notes to ch. 5] [5]

The 37% and 20% estimates match the later twin paper. The endnote, however, attributes the information to Plomin on 21 June 2015. The report withdraws any implication that the author necessarily had the completed 2016 paper's less favorable predictive findings in hand. That paper is still relevant evidence against a strong novelty or general-success claim.

---

08 / Height: variants are not genes

[B, ch. 5] · [N05-18, Notes to ch. 5] [42]

The book turns 697 into a count of genes. The cited Wood study instead reports 697 genome-wide significant variants clustered in 423 loci. This is a genuine terminology error, not a result that merely became outdated. The central claim that height is influenced by many genetic factors is unaffected; the specific enumeration needs correction.

09 / Interest fit, job satisfaction, and life satisfaction

[B, ch. 6] · [N06-06, Notes to ch. 6]; [N06-07, Notes to ch. 6]; [N06-08, Notes to ch. 6] [31] [22] [23]

The satisfaction meta-analysis is Morris's 2003 dissertation; the life-satisfaction extension is a manuscript described as in preparation; performance is supported by Nye and colleagues' published synthesis. These are three distinct sources, not one settled package. Morris's underlying estimates were not retrieved. The later interest-fit estimate of corrected $\rho = .19$ challenges "enormously" as current magnitude language without retroactively substituting for Morris.

10 / Purpose: what actually doubled

[B, ch. 8] · [N08-44, Notes to ch. 8] [17]

The original Study 3 had 71 undergraduates providing outcome data and measured average seconds per exam-review question. The raw means were 49 in the purpose condition and 25 in control, a ratio of 1.96; the reported transformed-outcome test gave $p = .038$. Number of questions completed was not significantly different. This verifies a narrow near-doubling, not general study time, enduring grit, or a universal effect.

The note labels the journal "Attitudes and Social Cognition," which is a section label rather than the journal's name. The correct venue is Journal of Personality and Social Psychology; the title, volume, and pages still identify the article.

11 / Google job crafting: final report found, allocation and durable outcomes bounded

[B, ch. 8] · [N08-45, Notes to ch. 8] [43] [44] [45]

The note cites a 2001 theoretical article, a book account, and 2015 private correspondence. The 2013 researcher-authored chapter describes a 2012 field quasi-experiment; the 2016 Berg interview names Google. Berg and colleagues' final report explicitly identifies an early 2012 Academy of Management version. Its analyzed 149 employees plus three without observer ratings match the 152-person account, making it a strong source match. The final paper leaves the corporation unnamed, and the private correspondence was not authenticated. This is later publication of earlier work, not a new independent replication. [43] [45] [54]

Study 1 calls itself quasi-experimental: employees volunteered and entered workshops according to availability, while workshop conditions were randomized after arrival. That is not individual random assignment to the full program. The job-only condition showed a six-week increase in observer-rated happiness (unadjusted $p=.042$, $d=.29$), but Table 3's Bonferroni adjustment gives $p=.085$. This changes the strength of statistical support, not the estimated increase. The six-month job-only contrast was not significant ($p=.58$); it does not

establish that no lasting effect is possible. The dual self-and-job approach showed a later happiness benefit that retained statistical support after adjustment (six months, adjusted p=.007). The exploratory performance result was confined to the job-only six-week comparison, not a sustained performance gain. There was no untreated comparison group that could identify the full package or its interaction against doing nothing. [54]

Study 2 analyzed 398 online workers across three individually randomized active conditions, after excluding 16 of 414 survey completers for meaningless responses. At six months, 165 completed follow-up, about 41% of the analytic sample. The dual approach again supported some later happiness benefits; the short-term job-only pattern did not recur. Job-crafting intentions were not directly observed job changes, and mediation estimates do not establish a causal mechanism. These favorable and mixed findings support conditional use rather than a general durable-performance promise. [54]

The field study's participant data are protected by the firm and unavailable; its analysis code is public. Study 2 data and code are public at OSF. Neither was independently executed here. The paper discloses the authors' financial/copyright interests in the Job Crafting Exercise. The current Michigan page is a tool-provider description that cites the final report, not independent validation or the source for the allocation details. [44] [54]

---

12 / Wise feedback: adjusted rates, not raw universal doubling

[B, ch. 10] · [N10-34, Notes to ch. 10] [18]

The note identifies the essay-feedback experiments. The book's approximate 40% and 80% are consistent with averages of the two groups' adjusted values in the source figure: 39.5% and 79.5%. They are not reconstructed pooled raw proportions. The effect is about credible feedback in a specific educational and social context, not a proven general parenting formula.

Later context: the undergraduate wise-feedback experiment found no trust advantage at high baseline trust. Its different sample and trust conditions limit transfer; the result does not negate the original adjusted essay-revision finding or validate a whole parenting style. Full later-article methods were not accessed. [53]

---

13 / Cognitive therapy: acute and relapse evidence both cited

[B, ch. 9] · [N09-14, Notes to ch. 9] [25] [46]

The complete note supplies both the acute-treatment study and the follow-up on relapse. Among responders, the follow-up supports enduring benefits relative to medication withdrawal; it did not establish a statistically significant advantage over continuation medication. Better documentation does not remove the need to specify the comparator. No treatment change is recommended by this review.

14 / Rat controllability: accepted manuscript read, lifespan extrapolation unsupported

[B, ch. 9] · [N09-34, Notes to ch. 9] [41]

The endnote identifies Kubala and colleagues' adolescent-rat study. The complete official NIH accepted manuscript supplies the methods, results, discussion and captions. It examines male rats' social exploration and stress responses about 35 days after adolescent controllable stress, not lifelong human character development. This is access to the original report, not a newly published study or a successful independent reproduction. [41]

The original paper itself gives important limits: increased baseline exploration complicates interpreting protection, the surgery experiment did not reproduce that baseline increase, and an analogous adult long-delay experiment was unsuccessful and reported as data not shown. The authors treat lifespan immunization as a hypothesis. These source-specific qualifications preserve the case for bounded controllability effects while withholding a lifelong-human claim. No rat-level data were analyzed, and manuscript reporting details were not certified against the typeset version. [41]

15 / Hope and the Penn Resiliency Program

[B, ch. 9] · [N09-40, Notes to ch. 9]

The endnote explicitly acknowledges that the cited meta-analysis did not find an advantage over active control conditions. That is a substantive qualification and improves the assessment of balance. The identifiable Brunwasser, Gillham and Kim synthesis includes 17 evaluations and 2,498 participants, with small benefits against no intervention but no active-comparator advantage. Sensitivity and publication-bias limits matter; the results were not re-estimated here. This is confirmation of a caveat already disclosed by the book, not a newly discovered omission. [55]

The original prevention findings are more specific than a universal resilience recipe. Gillham and colleagues' 1995 follow-up reported favorable symptom outcomes for at-risk schoolchildren against matched no-treatment controls, not verified individual random assignment. Seligman and colleagues' 1999 university trial randomized an intervention against assessment only and found benefits for some outcomes, not uniformly for severe depression or clinician ratings. Current access was to primary abstracts; neither is a direct grit intervention. [58] [59]

Tak and colleagues' adapted Dutch program, delivered universally and randomized across nine schools, found no depressive-symptom benefit at one or two years versus school as usual. It changed the population and delivery relative to targeted original prevention studies. Its primary abstract and introductory design context were accessed, not a full sample-flow or raw-data audit. This contextual null result does not prove all PRP findings false or establish that all forms of hope are ineffective. [56]

16 / Parenting advice: conjecture is labeled

[B, ch. 10] · [N10-20, Notes to ch. 10]; [N10-21, Notes to ch. 10]; [N10-22, Notes to ch. 10]

The text openly distinguishes grit-specific hunches from the wider parenting literature. Its notes identify supporting parenting studies and reviews. That supports a characterization of informed extrapolation rather than unsupported invention. It still does not isolate how much an intervention on parenting changes a child's general grit, separate from child effects, genes, opportunities, and other family conditions.

17 / Grit Grid and extracurricular transfer

[B, ch. 11] · [N11-07, Notes to ch. 11]; [N11-12, Notes to ch. 11]; [N11-13, Notes to ch. 11]

The college-persistence grid is tied to a manuscript in preparation, while other follow-through and teacher findings have different sources. The notes identify the trail and also acknowledge the validity of standardized tests. The exact unpublished grid estimates are not independently reproducible from the material retrieved. Selection into activities, prior achievement, and resources remain plausible explanations alongside cultivation.

18 / Hard Thing Rule: the package is still a family rule

[B, ch. 11] · [N11-25, Notes to ch. 11]

The surrounding chapter cites learned-industriousness research, but the endnotes do not supply a controlled trial of the named household rule, its natural stopping points, or its later two-year requirement. Related learning experiments do not establish the rule's optimal duration or broad transfer. The rule can be sensible while remaining an unvalidated specific intervention.

19 / An internal attribution problem in the culture chapter

[B, ch. 12] · [N12-26, Notes to ch. 12]

The main text attributes the claim about language to coach Anson Dorrance. The corresponding note says "Dimon, interview." This is an identifiable internal source-label mismatch. The review does not guess which interview the author intended or infer that the underlying quotation is fabricated. The correction needed is to reconcile the source label with the speaker.

20 / Well-being and demand for more grit

[B, ch. 13] · [N13-01, Notes to ch. 13]; [N13-03, Notes to ch. 13]

The notes identify a conference poster for the well-being results and unpublished data for the three-hundred-person preference survey. These are candid disclosures, not peer-reviewed confirmation. Wanting more grit is not evidence that more would benefit the person, and a favorable association does not establish effects on partners, children, colleagues, or people induced to persist.

07 Practical use, limitations, and safeguards

Practical Value · 50%

The intended uses assessed are individual achievement, education and parenting. Practice and support receive credit independently of the stronger claim that a single measurable trait can be raised to produce broad long-term success. Selection applications require their own validation.

Practical Value is 50%: there is real value in selected practices, but generalized grit-building benefits, cross-context transfer, and a sufficiently explicit persist-or-stop decision rule remain incompletely established. Criterion-by-criterion rationale.

Ease of application — not included in the score: Requires sustained effort, useful feedback, and often outside support.

The implementation suggestions that follow are the reviewer's synthesis, not evidence that the book already supplies every safeguard or that this review has validated a replacement program. The score evaluates the book itself, not the reviewer's improved version of its advice.

The recurring inferential problems

1. The intervention gap: a predictor is not automatically a lever

Suppose a score forecasts who completes a difficult course. That does not establish that raising the score will raise completion. The score might partly reflect preparation, stable opportunities, self-knowledge about likely success, or willingness to report socially desirable traits. Even when changing a behavior would help, changing the questionnaire response is not necessarily changing the behavior.

The relevant chain is: **intervention → change in the intended construct → durable behavior → worthwhile outcome**. Evidence can fail at any link. A randomized program may improve a grade while leaving the presumed mediator unchanged, or while changing several mediators

together. Conversely, a measured belief may improve while achievement does not. The intervention literature examined here demonstrates why these endpoints should be separated. [11][12][16]

2. The selection problem: survivors explain themselves

An accomplished person's account can reveal a strategy worth testing. It cannot identify the success rate of that strategy. The relevant denominator includes people who worked similarly hard but encountered different constraints, pursued poorer opportunities, or failed. The narrative also usually lacks the counterfactual: what would the same person have achieved after changing course?

Selection operates inside the quantitative evidence too. A highly screened military academy, national spelling final, or selective university is not the general population. Range restriction can attenuate associations. Conditioning on admission or high performance can also create misleading relationships among the ingredients of selection. Neither issue means that a study is invalid; it means the inference belongs to a population and selection process.

3. The construct problem: a label can hide multiple mechanisms

The narrative "gritty person" may combine conscientiousness, interest, identity, purpose, skill, tolerance of discomfort, and access to reinforcement. The scale samples a narrower set of self-descriptions. Evidence for a broad narrative should not be credited to a narrower score by association. This is the central measurement concern, not a complaint that ordinary language lacks perfect definitions.

A high reliability coefficient does not establish uniqueness, a single latent entity, measurement invariance, or suitability for consequential decisions. The 2021 author commentary is important because it corrects an inference about a higher-order factor without abandoning a conceptual case for the compound. That is a scientific refinement—not proof that persistence is imaginary. [8]

4. The incremental-value problem: compare against the right alternative

Predicting better than chance is not enough to justify a new assessment. The practical question is whether the measure improves decisions beyond information already available, at acceptable cost. A broad conscientiousness scale, a carefully matched industriousness facet, prior performance, work samples, and references are different comparators. Beating one weak or noisy comparator does not establish superiority to the whole set.

Conversely, a measure need not be wholly distinct to be useful. A short, inexpensive proxy might be worthwhile even if theoretically redundant. To demonstrate that, evaluate out-of-sample prediction, calibration, incremental utility, faking susceptibility, and subgroup performance. These are proposed validation requirements, not claims that the book's instrument has already satisfied them.

5. The substitution problem: retention, excellence, and well-being are not interchangeable

A program might retain more people without improving their learning; a person might perform better while damaging health or relationships; a strategic departure might produce a better outcome than staying. The evaluation target should be specified before choosing the measure. Otherwise, "success" can move between whatever outcome happened to support the theory.

This also changes institutional ethics. An employer benefits from retention, but employees may benefit from better outside opportunities. A school can encourage persistence while also changing the conditions that make learning difficult. The book's own environmental and opportunity discussions support resisting a purely individual attribution of failure. [B, ch. 3; ch. 11; ch. 12]

6. The falsifiability problem: the goal hierarchy can be moved after the result

Goal hierarchies are useful for planning but permissive for storytelling. A scientist who switches fields can be called inconsistent, or deeply consistent in pursuing truth. A founder who closes a company can be called a quitter, or persistent in pursuing a broader mission. Both interpretations may be reasonable, but an explanation that can absorb every outcome needs prospective definitions to make testable predictions.

Before acting, state the goal's level and timescale, the evidence that would count as progress, and the conditions under which even the top-level goal should be revised. Otherwise, grit can become a flattering after-the-fact description rather than an informative cause.

A necessary restraint on this review's own criticism

These problems do not imply that small associations are worthless, that observational research has no role, or that only a perfectly isolated trait intervention can ever be useful. They imply that the strength and scope of the conclusion must match the design. A helpful book need not invent a new trait. But a scientific explanation of exceptional achievement must do more than package useful existing practices under a compelling name.

What to keep, modify, and reject

The recommendations below are the review's decision-oriented synthesis. They are not a newly validated intervention, and they do not claim a known average effect from following this report. They preserve the book's strongest distinctions while adding safeguards against its weakest inferences.

Keep	Modify	Do not infer
Practice a defined weakness and use informative feedback.	Check whether the feedback predicts the real outcome and whether the skill is the present bottleneck.	More hours necessarily create more value.
Develop interests through real experience.	Test fit, opportunity, daily tasks, and economic feasibility alongside enjoyment.	Everyone has one predetermined calling or must never change fields.

Keep	Modify	Do not infer
Maintain an important purpose while changing tactics.	Review the purpose itself when circumstances or values change.	Persistence is intrinsically superior to exit.
Combine high standards with credible support.	Make support material: instruction, feedback, resources, and room to report problems.	A slogan or stern anecdote is a validated parenting or management program.
Use reflection to notice patterns of avoidance or inconsistency.	Corroborate self-description with behavior and context.	A grit percentile measures moral worth or is ready for high-stakes screening.

For an individual: identify the bottleneck before prescribing persistence

Begin with one concrete outcome: finish a thesis chapter, improve a particular sales skill, train a specific movement, or complete an application. Then ask what currently prevents progress. Is it insufficient skill, inconsistent initiation, poor feedback, low expected value, fear of evaluation, lack of resources, or a goal that no longer fits? "Be grittier" is not a diagnosis that distinguishes these mechanisms.

Choose a small intervention matched to the bottleneck. For a skill gap, obtain better instruction and targeted repetitions. For unreliable initiation, specify a start condition and a manageable next action. For poor feedback, shorten the cycle and improve its validity. For a low-value opportunity, compare alternatives. Track the outcome and the cost, not only hours worked or resolve felt. These are reasoned applications of the review, not proof of a general grit-building mechanism.

For parents and educators: direct benefits first, transferable grit second

Choose activities with good instruction, safety, genuine engagement, and a feasible commitment. Give children a voice in selection and an agreed review point. A hard activity can be valuable without conferring a proven general character advantage. Monitor whether the child is acquiring skill and becoming more willing to engage—not merely complying longer.

Make feedback concrete and actionable. High expectations without a credible route to meet them are different from supportive challenge. When a child struggles, inspect the task and setting as well as effort. Avoid using a questionnaire score to label a child as inherently lacking persistence. The book's own cautions and the measurement literature warrant this restraint. [B, ch. 10; ch. 11][26]

For founders and managers: persistence is not a substitute for strategy

Define what the team should persist with and what it should rapidly change. A mission may endure while a product, channel, or operational process fails. Reward accurate reporting and correction, not just stamina. A culture that punishes negative information can keep a team committed to an error. This is a decision principle; the report does not claim that a particular management package has been experimentally established by grit research.

Do not deploy a grit cutoff because the construct is associated with retention elsewhere. Test whether an assessment improves the actual decision beyond existing evidence, whether it predicts the relevant outcome, and whether applicants can readily alter their answers. Keep

the distinction between selecting people and improving their working conditions explicit.

A prospective persist–pivot–stop rule

> **Persist** when the goal remains worthwhile, the next increment of effort has a credible expected return, and you are learning or improving.
>
> **Pivot** when the goal remains worthwhile but the current method, market, teacher, environment, or allocation of effort has repeatedly failed a fair test.
>
> **Stop or replace the goal** when updated expected benefits no longer justify future costs, a better use of resources dominates, or continued pursuit breaches a health, safety, relational, or ethical constraint. Past investment alone is not a reason to continue.

Before the next commitment period, write down a time or resource budget, a progress signal, a review date, and a stopping condition. Do not demand certainty: some projects need a long runway before their value is visible. Instead, specify why the runway is justified and what evidence would change the decision. That converts perseverance from an identity claim into a governable strategy.

What evidence would change this review's verdict?

A stronger trait account would need improved measurement that captures enduring commitment without penalizing sensible adaptation, demonstrates validity across intended populations, and adds useful prediction over matched personality facets and prior performance. A stronger intervention account would need well-powered randomized comparisons, durable behavioral outcomes beyond the trained task, evidence on mechanisms, and explicit measurement of costs and adverse effects. Evidence that a whole-book program improves important outcomes would also deserve direct credit.

08 What changes with time—and what should change in the book

What changed after the original book?

Two questions should remain separate: **Was the original presentation appropriately calibrated to evidence available at the time?** And **how should a reader assess it now?** The rating here addresses the latter. Later studies are not treated as information the author should somehow have possessed before publication.

Concerns were not all hindsight

Before the book appeared, the practice literature already included a substantial quantitative challenge to unusually strong practice-based explanations, and laboratory work had examined costly persistence. The book itself knew that parenting effects, extracurricular transfer, age-related growth, and some costs were not established. It also referred to the twin work discussed here. Those circumstances warranted caution even without later publications. [13][24][B, ch. 5; ch. 10; ch. 11; ch. 13]

Later evidence narrowed the trait claims

The large academic synthesis and the authors' own factor-structure commentary make it harder to defend a simple story in which a well-isolated, uniformly powerful grit trait has already been measured. The appropriate update is not to erase the early predictive findings; it is to give greater weight to facet differences, overlap with established personality, and the outcome being predicted. [7][8]

Later intervention evidence cuts in both directions

The classroom experiments in Turkey are an important positive result, not an inconvenient exception to discard. The North Macedonia manuscript is more mixed: it reports improvements in some behaviors and outcomes alongside lower consistency-of-interest scores, potentially reflecting a wording-related measurement problem rather than genuine loss of long-term commitment. The English working paper shows that an intervention can alter beliefs without a detected achievement benefit. None of these programs is simply "reading Grit." [10][11][12]

The mindset debate requires more than choosing the preferred meta-analysis

Macnamara and Burgoyne report a small overall academic effect and a near-zero estimate in their selected higher-quality subset. Burnette and colleagues report more favorable effects for targeted groups under high-fidelity implementation. Their estimates answer different conditioning questions. Tipton and colleagues argue that modeling heterogeneity differently changes the conclusions. The defensible synthesis is a conditional, contested intervention literature—not universal effectiveness and not a settled zero effect for every population. [20][21][30]

> **Current verdict**
> Evidence for some useful practices has strengthened. Evidence for the strongest, unified "grit explains and can transform success" interpretation has not. Those updates can coexist without contradiction.

What the complete endnotes establish

This report brings together the claim, chapter and source assessments. Its judgments reflect the documentation and evidence identified in the accompanying analysis.

Topic	Finding and consequence
Edition and source access	The complete 2016 edition, including its notes and copyright information, is the basis for book citations.
Reference audit	The sample uses a disclosed 158-anchor frame and a fixed ten-entry draw, with eight assessed cases and two unscored access cases. No unverifiable single ten-pair score is assigned.
Questionnaire	The note explains ten-item provenance, a reported .99 correspondence, wording changes, norms, and measurement caveats. The criticism concerns use-specific validation.
Practice	The book cites the major practice critique and a response. This documents the debate without resolving the scientific reservations.
Genetics	The endnote cites a dated 2015 communication, not the final 2016 paper. No deliberate omission is alleged. A separate variants-versus-genes error is identified.
Purpose	The near-doubling is verified as 49 versus 25 seconds per exam-review question, not generic study time.
Job crafting	The Google case is corroborated through researcher accounts. The final original report identifies workshop allocation, mixed duration-specific outcomes, and separate field/online-study data access. These support a strong but qualified match to the book's account.
Hope	The endnote's explicit active-control null qualification receives credit.
Publication status	Unpublished manuscripts, a conference poster, private correspondence, and historical-source gaps have distinct, stated access limits.
Citation quality	The wrong journal label for the purpose article, a small ASQ page-range error, and the Dorrance/Dimon internal attribution mismatch are recorded without alleging fabrication.
Central scientific rating	The three central scientific input grades are 2, 3 and 2. Their exact normalized result is **58% Scientific Accuracy**. The complete notes improve fairness and traceability without establishing the stronger general claims. The separate Reference Accuracy and Practical Value categories have their own stated criteria.

What the endnotes contribute: a clearer basis for assessing the author's documentation and its relationship to the evidence. The notes frequently support a narrower, more careful account than a popular summary of the book. Conversely, a source's appearance in the notes is not itself independent confirmation of its findings.

What would change the remaining verdict: stronger evidence that the full grit construct adds useful prediction beyond well-measured existing traits; replicated interventions that change durable, behaviorally verified grit and improve valued outcomes; and direct tests of when persistence should yield to changing goals. Better evidence for one component should update that component, not automatically validate the entire framework.

Scope of the 2019 West Point correction

The indexed primary text of the 2019 West Point correction establishes that it concerns procedures for requesting the data and additional variables. No dataset reanalysis is claimed. Correction text.

09 How the ratings were determined

The headline scores summarize three different editorial judgments. **Scientific Accuracy** evaluates three central propositions; **Reference Accuracy** evaluates traceability, description and inference in the examined evidence; **Practical Value** evaluates likely benefits, applicability and net value. Ease of application is reported separately and is not included in the average.

The source audits are not identical probability samples. Reference Accuracy therefore uses the same three anchored criteria in every review—not a percentage of references that passed an audit. Both faithful and problematic source use inform the assessment, and inaccessible passages are not counted as errors. See Sections 02, 06 and 10 for scope and coverage.

Overall rating: 64%. The three categories receive equal weight. This is an editorial convention, not an empirically validated estimate of overall truth or benefit. The percentage does not override the recommendation or the consequential caution on the cover.

Score rationale. All nine inputs are evaluated under the stated anchors. Scientific Accuracy is [2,3,2]: narrower prediction and useful practice receive credit, but general superiority and durable trait-building remain unestablished. Reference Accuracy is [4,3,3]: the complete notes are traceable and often faithful, with important description/inference limits; the appendix uses the corrected research figures without treating a later research correction as a book-reference defect. Practical Value is [2,2,2]: favorable component evidence coexists with uncertain transfer, duration, and costs. The 64% overall rating is not scientific certification.

Scientific Accuracy · 58%

Do the book's main claims hold up?

Grit is a powerful general predictor of achievement, sometimes more important than talent.

There is predictive signal, but it depends on the outcome and comparison. Retention, performance, and elite achievement are not interchangeable; a grit score is not a clean causal estimate. [Section 04]

Sustained, high-quality effort builds skill and helps turn skill into accomplishment.

Purposeful practice and persistence have a meaningful evidential basis. The book's equations are explanatory metaphors, not estimated production functions proving that effort has exactly twice the importance of talent. [Section 04]

The recommended assets and environments can build grit and improve long-term success.

Some educational interventions help in particular contexts. That does not validate a general recipe for changing a unitary, durable trait across domains. The complete notes acknowledge more uncertainty than a simple dismissal would suggest. [Section 04]

Reference Accuracy · 83%

Do the cited sources support what the book says?

Traceability

The complete anchored endnotes identify sources, private communications, and some qualifications clearly. Two inaccessible original passages are access limitations, not demonstrated source failures. [Section 06]

Accurate description

The checked reports are often faithful at a bounded scope; some numerical, construct, and measurement descriptions need qualification. [Section 06]

Appropriate inference

Important caveats appear in the notes, but the main narrative sometimes extends prediction or selected cases into broad developmental prescriptions. [Section 06]

Practical Value · 50%

Is the advice likely to be worthwhile for its intended reader?

The intended uses assessed are individual achievement, education and parenting. Practice and support receive credit independently of the stronger claim that a single measurable trait can be raised to produce broad long-term success. Selection applications require their own validation.

Evidence of intended benefit

Practice, feedback, interest development and some educational interventions have relevant support. That earns substantive credit. Much of the central trait evidence remains predictive rather than intervention evidence, however, and results from a structured classroom curriculum do not establish the effects of the book's full grit-building advice for adults or families. [10; 13; 16; 22]

Applicability and durability

The book distinguishes goals from tactics and gives useful contextual qualifications, including in its complete endnotes. Transfer across school, work, parenting and elite performance nevertheless remains uncertain. Specific positive trials coexist with weaker or null results in other programs; changing a general, durable trait cannot be assumed from improvement on a trained task. [10; 11; 12; 19]

Benefits relative to burdens and risks

Well-directed effort can be worthwhile, and the book credits supportive environments and evolving interests. But sustained practice carries substantial opportunity costs, while operational guidance for disengagement and revising high-level goals is incomplete. The score does not assume that effort is harmful; it reflects the conditional value of persistence and the risk of applying it indiscriminately. [book; 24; 26]

Practical Value is 50%: there is real value in selected practices, but generalized grit-building benefits, cross-context transfer, and a sufficiently explicit persist-or-stop decision rule remain incompletely established.

Calculation and scoring anchors

The full audit uses five anchored grades, from 0 to 4. These are the inputs behind the percentages, not an additional reader-facing rating system. At a given criterion, 0 denotes a demonstrated fundamental failure or contradiction; 1 major limitations; 2 partial or mixed support; 3 substantial support with qualifications; and 4 strong support at the defined scope. The series methodology gives category-specific anchors and missing-data rules.

Scientific Accuracy: (2 + 3 + 2) ÷ 12 × 100 = 58.333% ➜ **58%**.

Reference Accuracy: (4 + 3 + 3) ÷ 12 × 100 = 83.333% ➜ **83%**.

Practical Value: (2 + 2 + 2) ÷ 12 × 100 = 50.000% ➜ **50%**.

Overall: (7 + 10 + 6) ÷ 36 × 100 = 63.889% ➜ **64%**. This is exactly the average of the three *unrounded* category scores. Only the final displayed values are rounded.

Uncertainty. A one-grade disagreement on a single criterion changes its category by about 8.3 percentage points and the overall rating by about 2.8 points before rounding. This is a sensitivity calculation, not a confidence interval. Independent duplicate scoring and measured inter-rater reliability are not documented. Small differences between books should not be interpreted as precise scientific rankings.

Missing evidence. No category in this edition is withheld as unrateable, but some individual source passages remain unresolved. Those gaps retain their explicit access labels; they are not zeroes or hidden failed checks. If a whole required criterion could not responsibly be judged, the category and overall score would be withheld rather than silently changing the denominator.

10 Verification record and remaining limits

Research stages and access. The source-access labels describe what was retrieved and inspected at each stated date. The first ledger below records targeted source checks on 8 September 2026. "Primary text" means the stated material was retrieved, not that an entire article, its supplements or its dataset was independently audited. A repeat metadata or abstract check does not upgrade access to full text. The dated checks below identify the research performed and its limits; not every search or retrieval was independently repeated.

Check and source	Finding at the checked scope	Remaining limitation
G01 Duckworth et al. (2007): initial grit studies Primary abstract retrieved	The original studies report predictive relationships and incremental validity, including an average 4% of variance across listed success outcomes. They do not randomize people to a change in grit. Central claims and quantitative audit.	Abstract supports the scoped distinction; all individual model estimates were not reanalysed.
G02 Credé, Tynan & Harms: grit meta-analysis Primary abstract retrieved	The review covers 88 independent samples and 66,807 participants, questions the higher-order construct, and distinguishes perseverance from consistency of interests. It does not show that effort is irrelevant. Central claims and measurement discussion.	Meta-analytic abstract retrieved; no re-estimation of effects or psychometric models.
G03 Complete endnotes and seeded audit supplement Supplied book pages inspected; archived inventory and algorithm executed	The supplement contains 465 note anchors, a 158-ID restricted frame, and the ten stated draws. The draw reproduced exactly. Eight rows were assessed and two remained unscored. The chapter 3 endnotes were inspected; the complete endnote inventory is retained in the private audit record.	The 158-anchor inclusion coding was inherited, not independently recoded. Reproduction is algorithmic, not independent scientific agreement. The two inaccessible source passages remain unresolved.
G04 Duckworth et al. (2019): West Point megaanalysis Primary abstract and article text retrieved	The 11,258-cadet study distinguishes predictors of performance from predictors of persistence. The paper explicitly limits causal inference and cautions about generalization. Quantitative and central-claim discussions.	No participant-level reanalysis; no independent coefficient or confidence-interval audit.
G05 2019 correction: data access, not a retrieved results revision Complete correction text retrieved through the indexed primary PMC record; live page access challenged	The correction supplies contact and authorization procedures for access to the study data and additional variables. It does not provide access to the underlying data. Correction-text access and reference 29.	A separate retrieval returned an unrelated abstract and was discarded. No claim is made to have independently reconstructed every correction or dataset version.
G06 Foliano, Hoskins & Rolfe (2026): perseverance program Primary institutional working-paper abstract retrieved	The 100-school trial's abstract reports changes in beliefs and motivation without a detected effect on the listed achievement tests. This is a working paper and a particular classroom program, not a direct trial of reading Grit. Later-intervention evidence.	Current check is abstract-level; no original-model re-estimation or peer-review certification.

Edition identification and source limits

Edition details are taken from the book's title and copyright pages. They do not authenticate the reviewed copy against a publisher-controlled master or establish correspondence with every commercially distributed edition.

What has not been independently certified

The indexed primary correction text concerns data-access procedures. Access to that text does not constitute access to the dataset, a full version-history audit, or independent validation of coefficients. See check G05 and reference 29.

HTML is the canonical report text; the PDF is generated from that same text without abridgment. The package's validation record documents the executed structural, scoring, citation-target, text-coverage, and rendering checks and their results. A functioning reference link is not proof that the linked study supports an assertion. Nor does a clean PDF layout validate the science.

Remaining source-level boundaries from the complete-endnote review

Item	Remaining boundary
Two sampled source passages	The four-minute treadmill mean in Phillips et al. (1987) and Heath (1945), p. 74, were not verified against the needed original pages. The first paper's broader abstract was read; the historical book's existence was confirmed. Neither item is assigned a score.
Original satisfaction meta-analysis	Morris (2003) is identified, but its full results were not retrieved. Later meta-analytic findings are not substituted for its original estimates.
Google assignment procedure	The final original report specifies volunteer/workshop allocation and condition assignment. It supports a strong but qualified match to the Google account; private correspondence and firm-protected field data remain unavailable. Study 2 data and both studies' code are public but were not executed here.
2019 West Point correction	The indexed primary correction text was retrieved and concerns data-access procedures. This does not constitute access to the dataset, a full version-history audit, or independent validation of coefficients. See check G05 and reference 29.
Animal-study details	The complete official NIH accepted manuscript was read, including methods/results and limits. No independent line-by-line comparison with the final typeset version or rat-level analysis was performed; lifelong human effects remain untested.
Unpublished and private evidence	Original interview transcripts, private correspondence, raw unpublished Grit Grid data, the complete poster materials, and every cited book or historical anecdote are not authenticated here.
Search and reanalysis	A targeted full-document, reference/claim integrity assessment was conducted on 13 September 2026, including exact publication-notice screens. No exhaustive database census, participant-level reanalysis, or re-computation of every source or synthesis was performed.
Independent review	AI-authored, with no independent human peer reviewer or response from the author. The author was not contacted.

Audit method and endnote locator index

Inventory and sampling frame

The 465 entries are distinct bold anchor phrases in the endnotes, not distinct publications. Each receives an ordinal within its chapter. The 158-entry sampling frame was editorially selected for scientific, measurement, and research-informed propositions, including relevant theory, research monographs, and gray literature. It excludes most purely biographical material and private-only communications. This restriction limits generalization; the exact frame, rather than an idealized description of it, is the reproducible definition.

Chapter	Note anchors	In restricted frame
1	19	11
2	33	8
3	46	12
4	34	8
5	28	21
6	46	11
7	58	20
8	46	14
9	43	21
10	37	12
11	26	15
12	37	2
13	12	3
Total	465	158

Reproducing the draw

The frame was fixed before drawing this sample; it was not a blind or preregistered frame. No source was swapped after access failures. Every hash is computed from UTF-8 text, including the separator character. Sort ascending by the hexadecimal SHA-256 digest and select the first ten IDs.

```
seed = "grit-endnote-audit-2026-09-08-v2"
key(note_id) = SHA256(seed + "|" + note_id)
selected = sorted(frame_ids, key=key)[:10]
```

Draw order: N05-13, N04-18, N09-05, N03-38, N07-27, N02-14, N05-23, N04-12, N08-39, N03-40.

The audit supplement includes the 465-entry anchor inventory, the full 158-ID frame, the ten results, and a small verification script. No private interview transcript, unpublished dataset, or copyrighted research PDF is packaged as if it had been independently obtained.

Endnote locator index for this report

The short anchors below allow a reader to locate the exact note without relying on an invented numeric citation in the book. A repeated source is not counted as a new independent study.

N01-06 · Chapter 1, note anchor 6
“West Point admissions”

N01-07 · Chapter 1, note anchor 7
“those with the lowest”

N01-12 · Chapter 1, note anchor 12
“55 percent of the salespeople”

N01-16 · Chapter 1, note anchor 16
“42 percent of the candidates”

N01-17 · Chapter 1, note anchor 17
“Success in the military, business, and education”

N02-14 · Chapter 2, note anchor 14
“more likely to succeed”

N03-15 · Chapter 3, note anchor 15
“talent is how quickly”

N03-38 · Chapter 3, note anchor 38
“for only four minutes”

N03-40 · Chapter 3, note anchor 40
“becomes too severe”

N04-01 · Chapter 4, note anchor 1
“Grit Scale”

N04-02 · Chapter 4, note anchor 2
“how your scores compare”

N04-12 · Chapter 4, note anchor 12
“goals in a hierarchy”

N04-18 · Chapter 4, note anchor 18
“positive fantasizing”

N05-13 · Chapter 5, note anchor 13
“solve a quadratic equation”

N05-16 · Chapter 5, note anchor 16
“researchers in London”

N05-18 · Chapter 5, note anchor 18
“at least 697 different genes”

N05-23 · Chapter 5, note anchor 23
“Grit and age”

N06-06 · Chapter 6, note anchor 6
“fits their personal interests”

N06-07 · Chapter 6, note anchor 7
“happier with their lives”

N06-08 · Chapter 6, note anchor 8
“perform better”

N07-23 · Chapter 7, note anchor 23
“deliberate practice predicted”

N07-24 · Chapter 7, note anchor 24
“benefits to being quizzed”

N07-27 · Chapter 7, note anchor 27
“experienced as supremely effortful”

N07-50 · Chapter 7, note anchor 50
“challenging, effortful practice”

N08-39 · Chapter 8, note anchor 39
“personal loss or adversity”

N08-44 · Chapter 8, note anchor 44
“be a better place?”

N08-45 · Chapter 8, note anchor 45
“calls this idea job crafting”

N09-05 · Chapter 9, note anchor 5
“distinguish optimists from pessimists”

N09-14 · Chapter 9, note anchor 14
“longer-lasting in its effects”

N09-34 · Chapter 9, note anchor 34
“Steve Maier and his students”

N09-40 · Chapter 9, note anchor 40
“resilience training”

N10-20 · Chapter 10, note anchor 20
“authoritative parenting”

N10-21 · Chapter 10, note anchor 21
“a moratorium on further research”

N10-22 · Chapter 10, note anchor 22
“warm, respectful, and demanding parents”

N10-34 · Chapter 10, note anchor 34
“have very high expectations”

N11-07 · Chapter 11, note anchor 7
“beyond standardized tests”

N11-12 · Chapter 11, note anchor 12
“the Grit Grid”

N11-13 · Chapter 11, note anchor 13
“extracurriculars of novice teachers”

N11-25 · Chapter 11, note anchor 25
“Bob Eisenberger”

N12-26 · Chapter 12, note anchor 26
“language is everything”

N13-01 · Chapter 13, note anchor 1
“hand in hand with well-being”

N13-03 · Chapter 13, note anchor 3
“wanted to be grittier”

The seed, frame, draw, and note inventory are included under `publisher-support/data/grit-audit-provenance/` in the series package. The sample’s audit grades and percentage outputs are archived as provenance; they do not determine the headline Reference Accuracy score.

Additional source check · 8 September 2026

Additional source check · Alan, Boneva & Ertaç (2019), Ever failed, try again, succeed better. Author-institution abstract and publication record rechecked on 8 September 2026. The record reports two randomized elementary-school samples and later mathematics benefits from the specific curriculum. This is positive program evidence, not validation of reading Grit or of transfer to every adult achievement domain.

Limit: Abstract-level reconfirmation only; detailed methods and coefficients are not newly re-audited. This check does not constitute a new comprehensive literature search.

Production checks cover score arithmetic, internal links, preservation of the substantive analysis, agreement of HTML and PDF text, and representative browser and PDF rendering. The accompanying production record reports the actual results; these checks do not establish scientific or legal clearance.

Targeted study-integrity assessment · 13 September 2026

The preparation record documents a full read of the reader review and appendix and assessment of the 47 initial source-table rows and 20 registered claims, alongside 49 supplemental study, synthesis, theory, version, or access-lead records. These are not 49 independent replications. It records source-specific searches and shared exact DOI/notice screens. Not every retrieval was independently repeated; absent adverse records do not certify reliability or successful replication.

- The 2018 practice corrigendum was read in full, including its corrected table and formula; no independent meta-analytic reconstruction was performed. [48]
- Credé's complete indexed primary letter and Guo and colleagues' complete institutional manuscript were read. The original authors' reply was accessed in indexed substantive passages. Neither side's raw data or code was executed. [49] [50] [51]
- Kubala's complete official NIH accepted manuscript was read. It was not independently compared line by line with the final typeset version, and no rat-level reproduction was attempted. [41]
- Berg and colleagues' final typeset report was read for substantive methods, results, tables, limitations and disclosures, not every theoretical paragraph or reference. Study 1 data are firm-protected; Study 2 data and both studies' code are public but unexecuted. [54]

Later school, feedback, PRP, reference-bias, retrieval, parenting, and clinical context was added at the access levels specified in the source table. Original methods or exact results for Morris, Heath page 74, the treadmill duration/subset, non-final Grit Grid and well-being sources, and the private survey remain unresolved. Related later publications do not authenticate every inaccessible original estimate.

The relevant practice corrigendum and existing PNAS data-access correction are confirmed. No additional relevant original Grit-source retraction or expression of concern was established in the returned screens and stated searches. This is not an exhaustive notice census, raw-data validation, independent human subject-matter sign-off, or legal clearance. Later evidence is separated from what the author could have known when writing the 2016 book. [29] [48]

11 References and source access

These are the sources cited in the evidence dossiers. Dated access notes identify what was checked and any limits. An abstract or metadata record does not imply full-text verification. Sources 48–63 and the specified access records document the 13 September 2026 assessment; source checks and unsuccessful attempts are distinguished in Section 10. A source's identity or accessible link does not establish every associated claim.

Book citations use chapters and named sections rather than edition-dependent page numbers. Reference numbers are local to this report. Codes such as T/full text, A/abstract, M/metadata, or U/unresolved are defined in the local source ledger; a source can have accessible metadata while a particular passage remains unverified. External links require internet access; the review text and internal navigation work offline.

References and source-access ledger

Access key: T = full-text content was available and relevant sections were inspected; this does not mean every analysis or appendix was audited. A = original-author/publisher or indexed abstract verified; claims are limited accordingly. M = metadata only. T* = the entry's correction-specific note states the actual access limit. A DOI may lead to a paywall even when an accessible manuscript was used. Source-access dates: 8 September and 13 September 2026, as specified. The retrieved PNAS correction concerns data-access procedures, not a results revision; the underlying data were not obtained.

Entries 1–30 and 31–47 document the source and endnote checks; entries 48–63 document the 13 September study-integrity assessment. Dated access notes state what was inspected and what remains unresolved. A publisher's reputation is not proof that an inference is justified.

B / Book under review

Duckworth, Angela. *Grit: The Power of Passion and Perseverance.* Scribner, 2016. Ebook ISBN 9781501111129; complete text and endnotes. Citations marked B identify chapters and named sections. Endnote IDs are assigned by this review and paired with the book's own anchor phrases. The copyrighted book is not redistributed with this report.

Row	Source identity, bibliography, links, and access record
01	01 Red Pen Reviews. (2025). **Review method, version 2.2.** Official methods document. [T] The [T] label records access during the initial source checks. The 13 September repeat fetch of the external PDF failed; the site's own adapted rubric was fully read on that date, without a newly retrieved copy of every external-method provision. https://www.redpenreviews.org/wp-content/uploads/2025/06/Red-Pen-Reviews-method-version-2.2.pdf

Row	Source identity, bibliography, links, and access record
02	02 Duckworth AL, Peterson C, Matthews MD, Kelly DR. (2007). **Grit: Perseverance and passion for long-term goals.** Journal of Personality and Social Psychology, 92(6), 1087–1101. [T] https://doi.org/10.1037/0022-3514.92.6.1087
03	03 Duckworth AL, Quinn PD. (2009). **Development and validation of the Short Grit Scale (Grit–S).** Journal of Personality Assessment, 91(2), 166–174. [T] https://doi.org/10.1080/00223890802634290
04	04 Credé M, Tynan MC, Harms PD. (2017; online 2016). **Much ado about grit: A meta-analytic synthesis of the grit literature.** Journal of Personality and Social Psychology, 113(3), 492–511. [T] https://doi.org/10.1037/pspp0000102
05	05 Rimfeld K, Kovas Y, Dale PS, Plomin R. (2016). **True grit and genetics: Predicting academic achievement from personality.** Journal of Personality and Social Psychology, 111(5), 780–789. [T] https://doi.org/10.1037/pspp0000089
06	06 Duckworth AL, Quirk A, Gallop R, Hoyle RH, Kelly DR, Matthews MD. (2019). **Cognitive and noncognitive predictors of success.** Proceedings of the National Academy of Sciences, 116(47), 23499–23504. Read with the correction in reference 29. [T*] https://doi.org/10.1073/pnas.1910510116
07	07 Lam KKL, Zhou M. (2022; online 2021). **Grit and academic achievement: A comparative cross-cultural meta-analysis.** Journal of Educational Psychology, 114(3), 597–621. [A] https://doi.org/10.1037/edu0000699

Row	Source identity, bibliography, links, and access record
08	08 Duckworth AL, Quinn PD, Tsukayama E. (2021). **Revisiting the factor structure of grit: A commentary on Duckworth and Quinn (2009).** Journal of Personality Assessment, 103(5), 573–575. [A] https://doi.org/10.1080/00223891.2021.1942022
09	09 Jachimowicz JM, Wihler A, Bailey ER, Galinsky AD. (2018). **Why grit requires perseverance and passion to positively predict performance.** Proceedings of the National Academy of Sciences, 115(40), 9980–9985. [T] 13 September 2026: indexed original methods/discussion and the primary 2019 analytical exchange checked at the access levels in references 49–51. The interaction is contested; no independent data/code execution or new participant replication is claimed. [49] [50] [51] https://doi.org/10.1073/pnas.1803561115
10	10 Alan Ş, Boneva T, Ertaç S. (2019). **Ever failed, try again, succeed better: Results from a randomized educational intervention on grit.** Quarterly Journal of Economics, 134(3), 1121–1162. [T] https://doi.org/10.1093/qje/qjz006
11	11 Santos I, Petroska-Beska V, Carneiro P, Eskreis-Winkler L, Muñoz Boudet AM, Berniell I, Krekel C, Arias O, Duckworth AL. (Undated updated manuscript; accessed 8 September 2026). **Can grit be taught? Lessons from a nationwide field experiment with middle-school students.** Author-hosted manuscript, University College London. Publication status and version date not established here; not treated as a verified journal publication. [T] https://www.homepages.ucl.ac.uk/~uctppca/Manuscript.pdf
12	12 Foliano F, Hoskins S, Rolfe H. (2026). **Perseverance in the classroom: Findings from a randomised educational intervention in primary schools in England.** Stockholm School of Economics, Center for Educational Leadership and Excellence, Working Paper 26/5. Working paper, not a verified peer-reviewed journal article. [T] https://swoba.hhs.se/hastel/paper/hastel2026_005.1.pdf

Row	Source identity, bibliography, links, and access record
13	13 Macnamara BN, Hambrick DZ, Oswald FL. (2014). **Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis.** Psychological Science, 25(8), 1608–1618. [T] Earlier access: [T]. 13 September 2026: primary original abstract and complete publisher corrigendum read. Use the corrected 2018 values in the quantitative dossier; source data were not re-executed. [48] https://doi.org/10.1177/0956797614535810
14	14 Macnamara BN, Maitra M. (2019). **The role of deliberate practice in expert performance: Revisiting Ericsson, Krampe & Tesch-Römer (1993).** Royal Society Open Science, 6, 190327. [T] https://doi.org/10.1098/rsos.190327
15	15 Duckworth AL, Kirby TA, Tsukayama E, Berstein H, Ericsson KA. (2011; online 2010). **Deliberate practice spells success.** Social Psychological and Personality Science, 2(2), 174–181. [A] https://doi.org/10.1177/1948550610385872
16	16 Eskreis-Winkler L, Shulman EP, Young V, Tsukayama E, Brunwasser SM, Duckworth AL. (2016). **Using wise interventions to motivate deliberate practice.** Journal of Personality and Social Psychology. [T] https://doi.org/10.1037/pspp0000074
17	17 Yeager DS, Henderson MD, Paunesku D, Walton GM, D'Mello SK, Spitzer BJ, Duckworth AL. (2014). Boring but important: A self-transcendent purpose for learning fosters academic self-regulation. *Journal of Personality and Social Psychology, 107*(4), 559–580. https://doi.org/10.1037/a0037637 [T] Original author-hosted text, including Study 3 methods/results and the relevant rendered page, inspected for this review.
18	18 Yeager DS, Purdie-Vaughns V, Garcia J, Apfel NH, Brzustoski P, Master A, Hessert WT, Williams ME, Cohen GL. (2014; online 2013). **Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide.** Journal of Experimental Psychology: General. [T] https://doi.org/10.1037/a0033906

Row	Source identity, bibliography, links, and access record
19	19 Yeager DS and colleagues. (2019). **A national experiment reveals where a growth mindset improves achievement.** Nature, 573, 364–369. [T] https://doi.org/10.1038/s41586-019-1466-y
20	20 Macnamara BN, Burgoyne AP. (2023; online 2022). **Do growth mindset interventions impact students' academic achievement? A systematic review and meta-analysis with recommendations for best practices.** Psychological Bulletin. [T] https://doi.org/10.1037/bul0000352
21	21 Burnette JL, Billingsley J, Banks GC, Knouse LE, Hoyt CL, Pollack JM, Simon S. (2023; online 2022). **A systematic review and meta-analysis of growth mindset interventions: For whom, how, and why might such interventions work?** Psychological Bulletin, 149(3–4), 174–205. [A] Verified article DOI: 10.1037/bul0000368 https://pubmed.ncbi.nlm.nih.gov/36227318/
22	22 Nye CD, Su R, Rounds J, Drasgow F. (2012). **Vocational interests and performance.** Perspectives on Psychological Science, 7(4), 384–403. [A] https://doi.org/10.1177/1745691612449021
23	23 Hoff KA, Song QC, Wee CJM, Phan WMJ, Rounds J. (2020, online publication). **Interest fit and job satisfaction: A systematic review and meta-analysis.** Journal of Vocational Behavior, article 103503. [A] https://doi.org/10.1016/j.jvb.2020.103503
24	24 Lucas GM, Gratch J, Cheng L, Marsella S. (2015). **When the going gets tough: Grit predicts costly perseverance.** Journal of Research in Personality, 59, 15–22. [T] https://doi.org/10.1016/j.jrp.2015.08.004

Row	Source identity, bibliography, links, and access record
25	25 Hollon SD and colleagues. (2005). **Prevention of relapse following cognitive therapy vs medications in moderate to severe depression.** Archives of General Psychiatry, 62(4), 417–422. [A] https://doi.org/10.1001/archpsyc.62.4.417
26	26 Duckworth AL, Yeager DS. (2015). **Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes.** Educational Researcher. [A] https://doi.org/10.3102/0013189X15584327
27	27 Eskreis-Winkler L, Shulman EP, Beal SA, Duckworth AL. (2014). **The grit effect: Predicting retention in the military, the workplace, school and marriage.** Frontiers in Psychology, 5, 36. [A] https://doi.org/10.3389/fpsyg.2014.00036
28	28 Scribner / Simon & Schuster. (2018). *Grit: The power of passion and perseverance.* Paperback preview: copyright and contents pages. This preview documents the book's publication history; the 2016 source's own notes and copyright page establish the reviewed edition. [T] https://download.boekhuis.nl/9781501111112_fragm-docb.pdf
29	29 Proceedings of the National Academy of Sciences. (2019). **Correction for Duckworth et al., Cognitive and noncognitive predictors of success.** 116(52), 27163. Correction text retrieved during the documented source checks; it concerns data-access procedures. No underlying dataset was obtained. [T*] https://doi.org/10.1073/pnas.1920625117
30	30 Tipton E, Bryan CJ, Murray JS, McDaniel MA, Schneider B, Yeager DS. (2023). **Why meta-analyses of growth mindset and other interventions should follow best practices for examining heterogeneity: Commentary on Macnamara and Burgoyne (2023) and Burnette et al. (2023).** Psychological Bulletin. [T] https://doi.org/10.1037/bul0000384

Row	Source identity, bibliography, links, and access record
31	31 Morris MA. (2003). A meta-analytic investigation of vocational interest-based job fit, and its relationship to job satisfaction, performance, and turnover. Doctoral dissertation, University of Houston. ProQuest dissertation 3089788. [M] Bibliographic identity established from the book and catalog; results not independently retrieved. https://www.proquest.com/psycinfo/docview/305328153/abstract/140273DA7BF460B24
32	32 Oliver B, Harlaar N, Hayiou-Thomas ME, Kovas Y, Walker SO, Petrill SA, Spinath FM, Dale PS, Plomin R. (2004). A twin study of teacher-reported mathematics performance and low performance in 7-year-olds. Journal of Educational Psychology, 96(3), 504–517. [A] Original abstract inspected through the Johns Hopkins institutional publication record. https://doi.org/10.1037/0022-0663.96.3.504
33	33 Oettingen G. (2012). Future thought and behaviour change. European Review of Social Psychology, 23, 1–63. [A] Original abstract; review-wide methods and all underlying experiments not independently audited. https://doi.org/10.1080/10463283.2011.643698
34	34 Peterson C, Semmel A, von Baeyer C, Abramson LY, Metalsky GI, Seligman MEP. (1982). The attributional style questionnaire. Cognitive Therapy and Research, 6, 287–299. [A] Publisher abstract/metadata inspected, supplemented by the creator's University of Pennsylvania instrument description. https://doi.org/10.1007/BF01173577
35	35 Phillips KA, Vaillant GE, Schnurr P. (1987). Some physiologic antecedents of adult mental health. American Journal of Psychiatry, 144(8), 1009–1013. [A] Original abstract inspected through the Weill Cornell publication record. Treadmill mean and subset count not verified. https://doi.org/10.1176/ajp.144.8.1009

Row	Source identity, bibliography, links, and access record
36	36 Ericsson KA, Krampe RT, Tesch-Römer C. (1993). The role of deliberate practice in the acquisition of expert performance. Psychological Review, 100(3), 363–406. [T] Relevant original-text sections on practice and effort constraints inspected; not a fresh audit of every cited experiment. https://doi.org/10.1037/0033-295X.100.3.363
37	37 Tsay C-J, Banaji MR. (2011). Naturals and strivers: Preferences and beliefs about sources of achievement. Journal of Experimental Social Psychology, 47, 460–465. [T] Original article text inspected via an accessible copy; experimental scope retained. https://doi.org/10.1016/j.jesp.2010.12.010
38	38 Duckworth AL, Gross JJ. (2014). Self-control and grit: Related but separable determinants of success. Current Directions in Psychological Science, 23(5), 319–325. [T] Original full text via PubMed Central; a conceptual synthesis, not a randomized test of goal hierarchies. https://doi.org/10.1177/0963721414541462
39	39 Eskreis-Winkler L, Shulman EP, Duckworth AL. (2014). Survivor mission: Do those who survive have a drive to thrive at work? Journal of Positive Psychology, 9(3), 209–218. [T] Original Study 1 methods and results inspected in an accessible article copy; observational mediation, not causal identification. https://doi.org/10.1080/17439760.2014.888579
40	40 Heath CW. (1945). What people are: A study of normal young men. Harvard University Press. Cited page: 74. [M] Catalog confirms the work. The cited page and quotation were not retrieved. https://findit.library.nd.edu/Record/001345026

Row	Source identity, bibliography, links, and access record
41	41 Kubala KH, Christianson JP, Kaufman RD, Watkins LR, Maier SF. (2012). Short- and long-term consequences of stressor controllability in adolescent rats. Behavioural Brain Research, 234, 278–284. [T] 13 September 2026: complete official NIH accepted, unedited manuscript read, including methods, results, discussion and captions. This replaces the earlier metadata-only limit. Not independently line-compared with the final typeset version; no rat-level data/code reproduction. Official NIH accepted manuscript · The complete BioC text used for the source read https://doi.org/10.1016/j.bbr.2012.06.027
42	42 Wood AR and colleagues. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. Nature Genetics, 46, 1173–1186. [A] Publisher abstract verifies 697 variants and 423 loci; no claim of a full GWAS reanalysis. https://doi.org/10.1038/ng.3097
43	43 Wrzesniewski A, LoBuglio N, Dutton JE, Berg JM. (2013). Job crafting and cultivating positive meaning and identity in work. Advances in Positive Organizational Psychology, 1, 281–302. [T] Researcher-authored chapter describes a field quasi-experiment and cites a 2012 working paper, Job crafting in motion. 13 September 2026: the final original report is identified and checked separately in reference 54. Use that source for allocation, outcomes, data access and tool interests; the chapter is earlier context, not an independent replication. https://justinmberg.com/wp-content/uploads/Wrzesniewski-LoBuglio-Dutton-Berg_2013.pdf
44	44 Center for Positive Organizations, University of Michigan. (Undated; current page accessed 13 September 2026). Job Crafting Exercise. [T] Complete current tool-provider page read. It cites the 2013 chapter and final 2023 report but does not itself currently specify the field allocation design. The 8 September page was not archived, so its earlier wording is not inferred. Not independent validation; use reference 54 for study design. https://positiveorgs.bus.umich.edu/cpo-tool/job-crafting-exercise/

Row	Source identity, bibliography, links, and access record
45	45 Stanford Graduate School of Business. (2016, January 22). Should employees design their own jobs? Interview reporting Justin Berg's research. [T] First-person researcher account corroborates the Google setting and reported outcomes; not sufficient to authenticate randomization. 13 September 2026: complete indexed researcher interview read; direct live access was restricted. It corroborates the setting, not the final assignment or long-term outcome details, which are checked in reference 54. https://www.gsb.stanford.edu/insights/should-employees-design-their-own-jobs
46	46 DeRubeis RJ and colleagues. (2005). Cognitive therapy vs medications in the treatment of moderate to severe depression. Archives of General Psychiatry, 62(4), 409–416. [A] Acute-treatment source identified and original abstract checked; distinct from the relapse follow-up in reference 25. https://doi.org/10.1001/archpsyc.62.4.409
47	47 Positive Psychology Center, University of Pennsylvania. (Undated; accessed 8 September 2026). Attributional Style Questionnaire. [T] Creator's description of questionnaire structure and administration; not a new independent validation study. https://ppc.sas.upenn.edu/resources/questionnaires-researchers/attributional-style-questionnaire
48	48 Psychological Science. (2018, 7 May). Corrigendum: Deliberate Practice and Performance in Music, Games, Sports, Education, and Professions: A Meta-Analysis. 29(7), 1202–1204. Publisher correction to Macnamara, Hambrick and Oswald (2014). [T] Accessed 13 September 2026. Complete live publisher notice, corrected table and formula read; no independent meta-analysis reconstruction. https://doi.org/10.1177/0956797618769891 https://journals.sagepub.com/doi/10.1177/0956797618769891

Row	Source identity, bibliography, links, and access record
49	49 Credé M. (2019; online 26 February). Total grit scale score does not represent perseverance. Proceedings of the National Academy of Sciences, 116(10), 3941. [T] Accessed 13 September 2026. Complete indexed primary analytical letter read; live PMC access challenged. Study 2 data were unavailable to the critic; no data/code execution here. https://doi.org/10.1073/pnas.1816934116 https://pmc.ncbi.nlm.nih.gov/articles/PMC6410808/
50	50 Guo J, Tang X, Xu KM. (2019; online 26 February). Capturing the multiplicative effect of perseverance and passion: Measurement issues of combining two grit facets. Proceedings of the National Academy of Sciences, 116(10), 3938–3940. [T] Accessed 13 September 2026. Complete three-page institutional manuscript read; not independently line-compared with typeset text. Their reanalysis used raw Studies 2/3 data; no independent execution here. https://doi.org/10.1073/pnas.1820125116 https://tuhat.helsinki.fi/ws/portalfiles/portal/122514268/Letter_30.01.19.pdf
51	51 Jachimowicz JM, Wihler A, Bailey ER, Galinsky AD. (2019; online 26 February). Reply to Guo et al. and Credé: Grit-S scale measures only perseverance, not passion, and its supposed subfactors are merely artifactors. Proceedings of the National Academy of Sciences, 116(10), 3942–3944. [Earlier indexed-primary access] Initial check, 13 September 2026. Introduction and substantive reply passages read; not a complete live typeset read or independent factor-model execution. https://doi.org/10.1073/pnas.1821668116 https://pmc.ncbi.nlm.nih.gov/articles/PMC6410803/ Further check: complete published PMC HTML and the four-page author-deposited OSF supplement, version 3, were read, including their tables. Archived Mplus output files' model and results sections were inspected but not executed. Participant data were not independently analyzed, and a complete live typeset publisher PDF was not obtained. Author-deposited supplementary information, version 3

Row	Source identity, bibliography, links, and access record
52	52 Balan A, Sjöwall D. (2023; online 22 February 2022). Evaluation of a Deliberate Practice and Growth Mindset Intervention on Mathematics in 7th-grade Students. Scandinavian Journal of Educational Research, 67(4), 549–558. Initial 13 September 2026 access: complete indexed primary abstract and sample/allocation excerpts; direct publisher access restricted. A further check read the publisher-formatted online article through the DiVA repository, including methods, results, Tables 1–3 and limitations. Teachers randomized existing mentor groups to conditions; group composition was unsystematic. The registration, full intervention materials and original data/code were not independently audited or rerun. The attitude narrative and one Table 2 significance marker differ; the summary reports no improvement, not an absence of every significant difference. https://doi.org/10.1080/00313831.2022.2042733 https://www.tandfonline.com/doi/pdf/10.1080/00313831.2022.2042733 Institutional copy of the published online article
53	53 Troy A, Moua H, Van Boekel M. (2024; online 6 February). Wise feedback and trust in higher education: A quantitative and qualitative exploration of undergraduate students' experiences with critical feedback. Psychology in the Schools, 61, 2424–2447. [A] Accessed 13 September 2026. Complete live primary abstract and data-availability statement read. Full article, reasonable-request data and registry not obtained; contextual null at high baseline trust, not an exact school replication. https://doi.org/10.1002/pits.23164 https://onlinelibrary.wiley.com/doi/abs/10.1002/pits.23164
54	54 Berg JM, Wrzesniewski A, Grant AM, Kurkoski J, Welle B. (2023; online 12 May 2022). Getting Unstuck: The Effects of Growth Mindsets About the Self and Job on Happiness at Work. Journal of Applied Psychology, 108(1), 152–166. [T] Accessed 13 September 2026. Final typeset substantive methods, results, tables, limitations and disclosures read. Study 1 data firm-protected/unavailable; Study 2 data and both studies' code public at OSF ka7m3, not executed. Corporation unnamed; qualified match to the Google account. Tool financial/copyright interests disclosed. https://doi.org/10.1037/apl0001021 https://justinmberg.com/wp-content/uploads/Berg-et-al_2023_JAP.pdf https://osf.io/ka7m3/

Row	Source identity, bibliography, links, and access record
55	55 Brunwasser SM, Gillham JE, Kim ES. (2009). A Meta-Analytic Review of the Penn Resiliency Program's Effect on Depressive Symptoms. Journal of Consulting and Clinical Psychology, 77(6), 1042–1054. [T, indexed relevant sections] Accessed 13 September 2026. Primary indexed methods, results and sensitivity discussion read; not independently recoded or re-estimated. Small no-intervention effects and no active-control advantage, already qualified in the book's note. https://doi.org/10.1037/a0017671 https://pmc.ncbi.nlm.nih.gov/articles/PMC4667774/
56	56 Tak YR, Lichtwarck-Aschoff A, Gillham JE, Van Zundert RMP, Engels RCME. (2016; online 25 September 2015). Universal School-Based Depression Prevention "Op Volle Kracht": a Longitudinal Cluster Randomized Controlled Trial. Journal of Abnormal Child Psychology, 44, 949–961. [A plus introductory design context] Accessed 13 September 2026. Complete primary abstract and introductory context read; full sample flow/raw analysis not audited. Nine-school universal adapted program differs from targeted original prevention. https://doi.org/10.1007/s10802-015-0080-1 https://link.springer.com/article/10.1007/s10802-015-0080-1
57	57 Lira B and colleagues. (2022, 10 November). Large studies reveal how reference bias limits policy applications of self-report measures. Scientific Reports, 12, 19189. [T, indexed relevant sections] Accessed 13 September 2026. Primary indexed methods, results and discussion read; no original cohort-model execution. Within-school and between-school comparisons differ; no blanket self-report invalidity is claimed. https://doi.org/10.1038/s41598-022-23373-9 https://pmc.ncbi.nlm.nih.gov/articles/PMC9649615/
58	58 Gillham JE, Reivich KJ, Jaycox LH, Seligman MEP. (1995). Prevention of Depressive Symptoms in Schoolchildren: Two-Year Follow-Up. Psychological Science, 6, 343–351. [A] Accessed 13 September 2026. Complete primary publisher abstract read. At-risk schoolchildren and matched no-treatment comparison; full original methods/raw data not obtained, no individual-randomization claim. https://doi.org/10.1111/j.1467-9280.1995.tb00524.x

Row	Source identity, bibliography, links, and access record
59	59 Seligman MEP, Schulman P, DeRubeis RJ, Hollon SD. (1999). The prevention of depression and anxiety. Prevention & Treatment, 2(1), Article 8. [A] Accessed 13 September 2026. Complete indexed original abstract read. Randomized prevention versus assessment only with outcome-specific benefits; original clinical outcomes/code not executed. https://doi.org/10.1037/1522-3736.2.1.28a https://ppc.sas.upenn.edu/sites/default/files/depressionpreventionseligman1999.pdf
60	60 Voderholzer U and colleagues. (2024, 27 November). Enduring effects of psychotherapy, antidepressants and their combination for depression: a systematic review and meta-analysis. Frontiers in Psychiatry, 15, 1415905. [A plus indexed relevant discussion] Accessed 13 September 2026. Complete primary abstract and comparator/quality passages read; 19 follow-up trials, no independent data/model reconstruction. Treatment termination is not a universal comparison against continuing medication. https://doi.org/10.3389/fpsyt.2024.1415905 https://pmc.ncbi.nlm.nih.gov/articles/PMC11632389/
61	61 Roediger HL III, Karpicke JD. (2006). The Power of Testing Memory: Basic Research and Implications for Educational Practice. Perspectives on Psychological Science, 1(3), 181–210. [A] Accessed 13 September 2026. Complete primary publisher abstract and exact original review identity read. Distinct from the same-year Test-Enhanced Learning experiment; constituent experiments not individually reproduced. https://doi.org/10.1111/j.1745-6916.2006.00012.x
62	62 Agarwal PK, Nunes LD, Blunt JR. (2021). Retrieval Practice Consistently Benefits Student Learning: a Systematic Review of Applied Research in Schools and Classrooms. Educational Psychology Review, 33, 1409–1453. [A] Accessed 13 September 2026. Complete live primary abstract and data-availability page read; full article inaccessible and OSF files not executed. Favorable learning synthesis with limited non-WEIRD coverage, not durable grit validation. https://doi.org/10.1007/s10648-021-09595-9 https://link.springer.com/article/10.1007/s10648-021-09595-9

Row	Source identity, bibliography, links, and access record
63	63 Pinquart M, Kauser R. (2018; online 10 April 2017). Do the associations of parenting styles with behavior problems and academic achievement vary by culture? Results from a meta-analysis. Cultural Diversity and Ethnic Minority Psychology, 24(1), 75–100. [A] Accessed 13 September 2026. Complete live primary PubMed abstract read; 428-study coding and raw models not executed. Generally favorable authoritative-parenting associations with cultural/ethnic variation, not a randomized parenting-for-grit effect. https://doi.org/10.1037/cdp0000149 https://pubmed.ncbi.nlm.nih.gov/28394165/

12 Editorial Review Notice, Disclosures, and Corrections

Purpose and scope.

This report is published for criticism, commentary, education, and discussion of matters of public interest. It evaluates the particular work, edition, claims, and evidence identified in the report. Its conclusions should not be extended to editions, claims, publications, or professional activities that it does not examine. It is not a comprehensive audit of the author's work or a certification of scientific accuracy.

Factual reporting and editorial judgment.

This review contains both factual reporting and editorial analysis. Quotations, descriptions of source material, and reported research findings are presented as factual information. Assessments of evidentiary strength, methodological quality, interpretation, balance, and practical value are the reviewer's evaluative conclusions, based on the sources, reasoning, and criteria identified in the report. Readers are encouraged to examine those materials and assess the conclusions for themselves. Evaluative judgments are not intended to imply possession of undisclosed facts concerning anyone's conduct or motives.

Meaning of critical assessments.

Descriptions such as "unsupported," "overstated," "misleading," or "inconsistent with the evidence" concern the specific claim, presentation, or evidentiary relationship discussed, for the reasons explained in the accompanying analysis. They do not, by themselves, assert that an author or other person knowingly made a false statement, intended to deceive, fabricated evidence, or engaged in professional misconduct. An assessment that the available evidence does not adequately support a claim is distinct from a finding that the claim is necessarily false.

Interpretation of scores.

Scores, grades, rankings, and recommendation categories summarize the reviewer's application of the stated rubric. They are not measurements of the author's honesty, character, intentions, or overall professional competence. They are not estimates of the percentage of a book that is true or false, probabilities that an author is correct, or representations of scientific consensus. The accompanying analysis explains the basis and limitations of each assessment.

Evidence, timing, and verification limits.

Conclusions reflect the materials examined and the state of the evidence assessed through the literature-search cutoff stated in this report. Relevant limitations—including unavailable sources, incomplete access to underlying data, and unresolved verification questions—are identified in the methodology or accompanying analysis. An inability to locate or independently verify a source does not, by itself, establish that the source does not exist or that a claim was fabricated. Evidence published after the reviewed work is distinguished from evidence available when it was written; later developments do not, by themselves, establish what an author knew or should have known at the time. Independent replication of studies, reanalysis of original datasets, and expert peer review are not claimed unless expressly documented.

Preparation and review disclosure.

This paired reader review and technical appendix was prepared with ChatGPT at Jason Hreha's request. AI tools materially assisted research retrieval, comparison, targeted source checking, analysis, drafting, scoring and production. The report presents the evidence, study-integrity findings and source-access limits described in Section 10, with all nine scoring inputs evaluated under the stated method. The evidence cutoff is 13 September 2026; this is not a publication date.

The publisher supplied an additional critique, *The Audit, Audited*, reporting approximately 110 checks of the review series. That document informed preparation; its author, tool identity and independent human-review status were not established in the supplied material. Its reported checks are not relabeled as independent human fact-checking or complete verification. The *Grit* equation locator was checked against the supplied source page. Source and production checks do not constitute a comprehensive literature search or independent replication. The targeted study-integrity screen is described in Section 10.

No independent human subject-matter sign-off or media-law clearance is documented. Jason Hreha is not represented as having personally verified every assertion. No response from a reviewed author or publisher was solicited for this review. Source-target, arithmetic, content-preservation and rendering tests are production checks, not scientific peer review. The extent and limits of source access are stated in Section 10 and the reference records.

Relevant interests and relationships.

Jason Hreha leads The Behavioral Scientist, is the author of the commercially available behavior-change book *Real Change*, and offers behavioral-science consulting and related services. These are relevant commercial and intellectual interests, including potentially

competing books, explanations or products. The publication's book page and services/contact page were checked on 8 September 2026.

The preparation record does not independently establish the presence or absence of additional relationships with the reviewed author or publisher, funding, compensation, review-copy arrangements, sponsorships or affiliate arrangements. No blanket absence-of-interests statement is made. The links included in this report are source and navigation links, not newly created affiliate links.

Affiliation and attribution.

References to individuals, organizations, institutions, publications, and trademarks identify the subjects or sources of discussion and do not, merely by their inclusion, imply sponsorship, approval, or endorsement. This is not an official Red Pen Reviews review and is not affiliated with or endorsed by Red Pen Reviews. Any use or adaptation of its publicly described approach is identified in the methodology; the ratings in this report are not ratings issued by Red Pen Reviews. Copyright in quoted or reproduced third-party material, where applicable, remains with the respective rights holders. Its inclusion in this report does not grant permission for separate reuse.

General information, not individualized advice.

This report does not provide individualized medical, psychological, legal, financial, or other professional advice, and reading it does not establish a professional-client relationship. Discussion of a recommendation's evidentiary support does not establish its suitability for any particular person or circumstance.

Corrections and substantive responses.

The publisher welcomes factual corrections, relevant additional evidence, and substantive responses from authors, publishers, researchers, and readers. Please send submissions through **The Behavioral Scientist contact form**, identifying the passage at issue, the proposed correction or clarification, and supporting sources. The publisher will assess submissions and correct substantiated factual errors. After first public publication, material corrections will be dated and explained in a correction record; changes resulting from new evidence or revised judgments will be identified as updates. A person's failure to respond to a request for comment is not treated as agreement with the review.

Publication details

Reviewer / preparation	ChatGPT-assisted critical analysis, commissioned by Jason Hreha. The extent of AI and human review is disclosed above.
Publisher and responsible editorial contact	The Behavioral Scientist / Haystack Group LLC, as identified in the website's terms. Responsible editorial contact: Jason Hreha.

Book and edition reviewed	Grit — Angela Duckworth. 2016 Scribner ebook; ISBN 9781501111129; complete endnotes.
Literature-search cutoff	13 September 2026. Targeted review and study-integrity assessment; not a guarantee of exhaustive coverage.
Publication date	The first public publication date is recorded on the hosting page when the review is published. The evidence cutoff and source-access dates describe research, not publication.
Evidence date	13 September 2026. Dated source-access records state the actual checks and their limits. This is not a publication date.
Current evidence assessment and contact	The current evidence assessment below summarizes the basis and limits of this review. For the current public version or to submit a correction, use The Behavioral Scientist contact form and identify the book and report ID.

Current evidence assessment

Evidence basis and limits. The targeted evidence cutoff is 13 September 2026. The practice assessment uses the 2018 publisher corrigendum's 24/23/20/5/1% domain figures and 14% overall figure; the notice postdates the 2016 book and does not establish an author omission at publication. The assessment includes the passion critique and reply, scoped mixed/null and favorable later context, the complete accepted rat manuscript, and the final original job-crafting report with separate Study 1/Study 2 data availability. It considers the book's explicit caveats, valid predictive findings and the substantive critique. The nine inputs are [2,3,2]/[4,3,3]/[2,2,2]: this evidence does not establish a universally dominant predictor, a measured effort law, a durable general grit-building package, or unconditional net benefit. Traceability and bounded source fidelity are strengths. Overall 64%; Scientific Accuracy 58%; Reference Accuracy 83%; Practical Value 50%. Current score rationale. No raw-data reproduction, human scientific sign-off or legal clearance is claimed.

Report format. The reader review is accompanied by this technical appendix, with shared policy on the series pages, a claim-navigation register and a source table. The evidence qualifications and numerical scoring inputs accompany the analysis. The displayed equations and their diagram appear in chapter 3; the chapter's discussion of effort's two roles supports the qualitative account. The additional audit supplied by the publisher is acknowledged as a supplied critique, not authenticated human peer review.

Scoring method. The report uses whole-number category percentages and an equally weighted overall rating. Practical Value is assessed from the documented evidence; it is not a relabeled usability score. Scientific Accuracy and Reference Accuracy use the common criteria.

Difficulty is described separately from the scored categories. The substantive claim, chapter and source audits explain the judgments; the numerical scoring inputs and arithmetic are shown in Section 09.

Source and response limits. Source-specific findings and access limits are documented in Sections 08–10. No author or publisher response is recorded in the preparation material.