
Technical appendix — Good Habits, Bad Habits

The evidence, source notes, and scoring rationale behind the Good Habits, Bad Habits review.

Evidence appendix · Commissioned by Jason Hreha for The Behavioral Scientist

Evidence cutoff: 15 September 2026. AI-assisted evidence review; no documented independent human scientific sign-off. Not an official Red Pen review, and not an original-data replication.

84 min read

Contents

The evidence behind the review

Navigable claim register

1. Verdict in context

2. Scope and method

3. Strongest fair reading

4. Three central claims

5. All-chapter assessment

6. Claim and source audit

7. Practical use

8. Evidence over time

9. Scoring rationale

10. Verification and limits

11. References

12. Notices and corrections

Central claims and selection provenance

The evidence behind the review

Companions: reader essay, claim/source CSV, current selection and provenance record, and scoring rationale. The CSV preserves the detailed claim and source records behind this appendix.

Navigable claim register

Exactly three propositions supply Scientific Accuracy. The A-series records are evidence and interpretation dossiers, not extra scientific inputs. The CSV gives smaller claim-block records, source/access mapping and score relationships; it is not an inventory certifying every claim in the book.

Scored proposition	Scope	Grade
C1	Habits as a principal explanation of persistence and much apparent self-control	2
C2	Deliberately building durable, lower-effort recurring behavior	2
C3	Context redesign in addiction recovery, including the Vietnam benchmark	2

Dossier	Question or finding	Navigation verdict	Confidence / limitation
A01	Identity, edition and complete-book access	Supported	high
A02	Strongest fair reading: goals, habits and conscious preparation	Supported	high for attribution; moderate for synthesis
A03	The 43 percent estimate and the diary definition of habit	Overstated	high
A04	Beneficial habits and self-control: association versus causal explanation	Partly supported	moderate
A05	Definitions and experimental identification	Partly supported	high
A06	Popcorn: bounded evidence of cue-sensitive persistence	Supported	moderate
A07	Training-induced habit: failures and a multilaboratory test	Partly supported	high for reported outcome; moderate for theory implication
A08	The author's published response and preserved goal control	Supported	moderate
A09	Habit-formation time: what 66 days means	Supported	high
A10	Routine cues versus clock-time cues	Partly supported	moderate
A11	Friction and practical fit	Practical judgment	high for attribution; moderate for usefulness judgment
A12	Immediate and uncertain rewards: human prescription exceeds the narrow source	Overstated	high for source scope; moderate for human generalization
A13	Monthly incentives: omitted successful arm and unmeasured habit	Incorrect	high
A14	Ten Top Tips: genuine short-term advantage and maintained loss	Supported	high
A15	Smaller weight-loss trial: preserve the comparator over time	Partly supported	moderate
A16	Flexible versus narrowly routinized exercise incentives	Partly supported	moderate
A17	Disruption, exploration and replacement	Partly supported	high for attribution; limited for unexamined natural experiments
A18	Stress: correct replication target and an inadequate critical contrast	Partly supported	high for identities; moderate for inference
A19	Ego depletion is not the same question as stress or habit resilience	Partly supported	moderate
A20	Vietnam: real recovery, non-equivalent comparator and causal overreach	Overstated	high
A21	Rat Park: a pre-book replication failure with important qualifications	Partly supported	moderate
A22	Favorable enrichment evidence that Wood actually cites	Supported	moderate

Dossier	Question or finding	Navigation verdict	Confidence / limitation
A23	Human network-support trials: useful benefits, multiple mechanisms	Supported	moderate
A24	Therapeutic employment: benefits while support is maintained	Supported	moderate
A25	Confirmed unit error in the restaurant portion example	Incorrect	high
A26	Chinese-buffet correction: disclose it, then evaluate the inference	Overstated	high for notice/attribution; moderate for source interpretation
A27	Information is insufficient, not necessarily ineffective	Partly supported	high for attribution; limited for unexamined synthesis
A28	Tian ritual/self-control retraction: precise scope	Supported	high for notice account; exact day unresolved
A29	Brooks ritual/performance retraction: verified notice and bounded reasons	Overstated	high for status and publisher-stated general grounds
A30	A surviving ritual experiment—and Wood's accurate qualification	Partly supported	moderate
A31	Repeated exposure and vegetable acceptance	Supported	moderate
A32	Meaning, fluency, practice and skill are not one mechanism	Partly supported	high for attribution; limited for unexamined individual studies
A33	Policy, defaults and energy feedback: benefit is not mechanism	Partly supported	moderate
A34	2026 deadline replication, retraction and author responses	Partly supported	high for notice/date; moderate for replication scope
A35	Figures, labels and an internal portion-size mismatch	Incorrect	high
A36	Phone appendix: concrete advice with unmeasured program-level effects	Practical judgment	high for advice; moderate for benefit–burden judgment
A37	Whole program, components and durable value	Partly supported	moderate
A38	Later favorable evidence: a hybrid model of social-media posting	Partly supported	moderate
A39	Traceable but unexamined or inaccessible sources	Unresolved	high for access record
A40	Methodology, voice and editorial provenance	Supported	high

The navigation verdict routes the reader to the dossier's qualified conclusion; it is not a label for every assertion in its CSV records or an estimate of the book's error rate. Mixed dossiers preserve the separate findings below. In A28, "Supported" concerns the verified source–status/definitional–citation account, not ritual efficacy. In A29, "Overstated" concerns treating the withdrawn performance source as established current support. Access–limited material remains unresolved rather than false.

Section navigation

1 Verdict · 2 Scope and method · 3 Strongest fair reading · 4 Three claims · 5 All chapters · 6 Source audit · 7 Practical use · 8 Evidence over time · 9 Scoring · 10 Verification and limits · 11 References · 12 Notices

1 Verdict in context

Editorial judgment: the book is substantially useful as a framework for examining the conditions around recurring action. Its scientific explanation is less secure where repeated behavior becomes a proxy for a specific habit mechanism, where correlational mediation becomes a causal account, or where narrow results are translated into broad promises. The most consequential problem is the transfer of the returning-veterans re-addiction figure into a treatment benchmark. The most clear-cut descriptive error is the restaurant example's substitution of kilojoules for calories. These conclusions are supported at different levels and should not be collapsed into a blanket finding that the book is either "proven" or "debunked." A02, A13, A20, A25, A37.

The ratings are **Scientific Accuracy 50%**, **Reference Accuracy 67%**, **Practical Value 67%**, and **overall 61%**, based on exact input totals of 6/12, 8/12, 8/12 and 22/36. They are editorial judgments, not percent-true estimates or probabilities of a reader's success. The complete reasons and neighboring-grade arguments are in Section 9. The relative stability of a numerical grade should not be confused with certainty about every underlying source.

There are substantial favorable findings. Wood makes conscious preparation important, acknowledges mixed control, rejects repetition as a sufficient explanation of excellence, includes favorable environmental treatment evidence, cites the buffet correction, and accurately says the ritual EEG experiment did not improve performance. Related intervention trials support some intended benefits. Later null findings have design limitations and are not treated as universal disproofs. A02, A09, A14, A18, A22, A23, A26, A30.

This is a present-day assessment of a 2019 first edition. Later source withdrawals are identified as later developments, not evidence that Wood knew their eventual status. The verdict applies only within the declared audit and access boundaries; it is not a complete reliability estimate for the bibliography.

2 Scope and method

Book identity and locators

Wendy Wood, *Good Habits, Bad Habits: The Science of Making Positive Changes That Stick*. Farrar, Straus and Giroux, New York. First edition, 2019. Copyright © 2019 Wendy Wood. eISBN 978-1-250-15908-3. Verified from the title image and `OEBPS/xhtml/copyright.xhtml`. See A01. The uploaded file name and OPF source ISBN were not treated as superior to those internal records.

Every book locator in this appendix is an actual EPUB section path, usually with a distinctive anchor or note number. File names below omit the common `OEBPS/xhtml1/` prefix unless necessary. They are not invented print pages. The input hash is
`aa85394ce9062a2ea9f2680e5002041725b62d25381e1d8d589535a1b9d11004.`

Coverage and selection order

All 49 spine sections were inspected: the complete 15 chapters, epilogue, phone-use application, all 16 chapter/application note sections, bibliography, index, acknowledgments, title/copyright material and other front/back matter. All 15 in-chapter illustrations were visually read with their context. The initial extraction duplicated nested note entries; deduplicated notes were used for reading. A check found no nonblank body text omitted by the extraction. Nothing substantive in the EPUB remains inaccessible. Reading the index does not make its entries independently verified sources.

The three central propositions remain unchanged. The preparation record does not verify that this exact selection preceded adverse-evidence searching: an earlier topic list named stress/fatigue/distraction as the third topic, while the final selection uses addiction. This is not proof of preregistration or prior selection order. Research was question-driven and purposive. The all-chapter audit checks the selection for balance; other findings do not become a fourth Scientific Accuracy input.

Operational rubric and search scope

The house methodology, including its Scientific Accuracy and Practical Value anchors, governs the scoring. The supplied methodology text was read during preparation. It controls the nine integer inputs from 0 to 4, three per category, with equal weighting within and across categories. Required missing/unrateable inputs would be NR and withhold affected category and overall scores. Inaccessible papers are not scored as zero.

The research was book-anchored and question-driven, not a preregistered systematic review. It followed cited papers and their primary reports; looked for replications, notices and author accounts; and sought favorable intervention evidence as well as adverse findings. Searches covered habit measurement and induction, repetition/automaticity, incentives, cue planning, habit-based weight interventions, exercise scheduling, stress/depletion, Vietnam recovery, enrichment, network support, rituals, deadlines and selected policy applications. Exact-title and DOI searches were used to reduce study-identity errors. No numerical search-yield or exhaustive-database claim is made because a systematic deduplication/screening corpus was not constructed.

Sources were inspected through journal pages, PubMed/PMC, author/institution repositories and mirrors of primary papers. A search result alone was not silently promoted to a complete article. Notices were checked against authoritative journal records where accessible. The cutoff is **15 September 2026**; “later” means later than the reviewed 2019 edition, not later than the present cutoff. Newly retrieved 2026 material is used only where its date and content were actually visible.

Access labels and inference rules

Full text sections read means the specified sections were read; it does not mean every supplement, every analysis or the original dataset was inspected. The reference register combines initial recorded access and subsequent audit access; these stages are distinguished where material. **Indexed primary sections read** means original text was exposed through a search index despite incomplete direct retrieval. **Full notice text read via index/PubMed** is a notice-specific access statement, not full access to the original study. **Abstract only** supports a bounded summary, not a complete bias assessment. **Metadata only** establishes an identity/status or a book-supplied citation, with its provenance specified. **Unavailable** means the substantive item was not obtained. For several unexamined sources, the bibliographic record in the book is the only access; this is marked explicitly rather than described as an independently retrieved primary record.

A failed retrieval is not a factual error. A retraction is not an automatic finding of misconduct by a named person. A correction is not a retraction. A nonsignificant effect is not proof of equivalence or exactly zero effect. Within-group improvement is not a randomized comparative effect. A follow-up without a control does not retain the original randomized contrast. Model fit and observational mediation do not isolate a causal mechanism by themselves. These rules are applied symmetrically to favorable and adverse evidence.

No original data were reproduced, no author was contacted for this review, and no independent human scientific review is documented. The public voice articles and selected guide material are used for writing/position, not as scientific evidence: P02, P03, P04.

3 Strongest fair reading

Wood's program is not "stop wanting things and let repetition run your life." The charitable, text-supported reading is that conscious goals choose the direction and arrange situations, while learned responses reduce the need to re-decide familiar actions. She acknowledges mixed control and goal-sensitive components, presents enjoyable reward as important, distinguishes excellence from repetition, and discusses both the opportunities and losses created by context change. These are not concessions added by this reviewer: they appear in the book. A02, A11, A17, A32.

The account can therefore be compatible with motivation, selection of a suitable activity and effortful practice. A criticism is warranted when the book's practical or causal conclusion outruns its evidence, not merely when its emphasis differs from Jason's public writing. An automatic start can be valuable without an entire performance being automatic. A useful environment can remain necessary without invalidating an intervention. A financial default can work without being a psychological habit. These distinctions preserve both the strengths and the limits of the strongest version of Wood's argument. A05, A24, A33, A37.

The addiction chapter likewise explicitly treats addiction as multifactorial and recognizes complementary biological and contextual explanations. It acknowledges the uncontrolled Vietnam comparison and the need for continued support in some interventions. The review

does not accuse Wood of prescribing an environment-only cure or of telling readers to discontinue care. It evaluates the actual causal attribution and transferable benchmark that remain after those qualifications. A20, A24.

4 Three central claims

The following wording is retained from the central-claim selection record. There are exactly three scientific inputs. Its selection reasons and passage anchors remain in the current selection and provenance record; its chronology is qualified in Section 2.

C1 — habit as a principal explanation of persistence

For frequently repeated everyday behavior, learned associations between contexts and responses are a principal explanation of persistence beyond conscious intentions. Building and relying on those habits, rather than repeatedly exerting willpower, explains much of the success popularly attributed to self-control and provides a more dependable basis for sustained action.

Grade: 2/4. Confidence: moderate. Habits have meaningful support as one source of persistent action and reduced deliberation. Experiments show learned response tendencies, and observational studies link beneficial habits with self-control and positive outcomes. But these findings do not establish how much of everyday persistence or the self-control advantage is caused by habits rather than goals, preferences, opportunities or other differences between people. The selected claim is partly supported at its stated breadth. Recent favorable preparation/switching evidence is included; it does not provide that population-level causal bridge.

Evidence: A02; A03; A04; A05; A06; A07; A08; A18; A38. Neighboring grades: see SA1.

C2 — deliberate construction of durable, lower-effort behavior

By arranging stable cues, reducing friction, choosing rewarding actions, and repeating them, people can deliberately build goal-consistent habits that make recurring behavior increasingly automatic and less dependent on ongoing effortful decisions. The book presents these principles as usable across eating, exercise, work, household, and financial routines—not merely laboratory button presses.

Grade: 2/4. Confidence: moderate. Longitudinal studies and randomized interventions show that selected repeatable health behaviors can become more automatic or improve under habit-oriented support. This is meaningful evidence for parts of the recommendation. It remains narrower than the cross-domain promise, and attendance or weight change does not itself establish a durable reduction in reliance on effortful decisions. Important measurement, transfer and durability gaps fit the method's partly-supported anchor. A trial of reading the exact book is not required; the existing component evidence is credited.

Evidence: A09; A10; A11; A12; A13; A14; A15; A16; A17; A37. Neighboring grades: see SA2.

C3 — context redesign and addiction recovery

Changing drug-use contexts and supplying rewarding alternatives can interrupt addiction-supporting habits and sustain recovery. Wood uses returning Vietnam veterans and enriched animal environments to argue that environmental redesign deserves a much larger role in treatment, including treating the veterans' roughly 5% first-year re-addiction figure as a benchmark of possibility against the 40–60% relapse figure she assigns to conventional treatment.

Grade: 2/4. Confidence: moderate to high. Partly supported. Controlled animal and human intervention evidence supports useful roles for environment, alternative reinforcement and abstinence-supportive networks. Wood's multifactorial caveats and explicit maintenance limitations deserve credit. The Vietnam cohort does not identify context change as the sole or dominant cause, and its first-year re-addiction outcome is not a transferable benchmark for heterogeneous treatment relapse. Robins herself cautioned against the simplified comparison. That consequential extension prevents a 3 at the selected scope.

Evidence: A20; A21; A22; A23; A24. Neighboring grades: see SA3.

5 All-chapter assessment

Every entry below describes the actual chapter as read, including favorable material. Statements about an unexamined study are book-attributed, not independent verification of its result. Each section's entire text and corresponding note file were read; the linked dossiers identify the deeper checks and limits. Chapter-level evaluations are editorial judgments, not extra scored inputs.

1. Persistence and Change

Book locator: chapter1.xhtml; corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A02; A34.

The chapter establishes the persistence problem and the claim that habit supplies an explanation beyond conscious intention. Its value is to direct attention to recurring conditions rather than moralizing inconsistent action. The review preserves the distinction between choosing a goal and repeatedly executing it. The precommitment citation is now affected by the September 2026 retraction, but that does not retroactively establish negligence in 2019. The chapter's organizing thesis remains subject to the C1 assessment rather than being judged solely by one withdrawn illustration.

2. The Depths Beneath

Book locator: chapter2.xhtml; corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A03; A05.

The diary evidence gives readers a concrete picture of repeated action in familiar places. The leap from that classification to an exact share of behavior performed without conscious thought is not fully supported. The unpublished lifespan manuscript was actually retrieved, so unpublished status is not confused with inaccessibility. Voting and creative-persistence examples are book-reported extensions, not independently reanalyzed causal demonstrations here. In particular, repeating a creative effort does not itself show that the content of the work has become automatic. The chapter is useful on regularity but less secure on the precision and breadth of its explanatory label.

3. Introducing Your Second Self

Book locator: [chapter3.xhtml](#); corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A02; A05; A07; A08.

Wood introduces context-response associations and distinguishes them from other automatic processes. That explicit distinction is favorable material and rules out treating the “second self” as a literal claim about two independent people inside a brain. The Ngram illustration is an historical word-frequency display, not evidence that a mental mechanism exists or has a particular magnitude. The experimental literature supports learned response tendencies but is contested on how reliably additional training produces devaluation-insensitive behavior. Later work on habitual preparation and goal-directed override is relevant to the strongest fair reading rather than a reason to dismiss the chapter wholesale.

4. What About Knowledge?

Book locator: [chapter4.xhtml](#); corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A05; A25; A27.

The useful point is that supplying knowledge need not remove barriers to action. Wood also explicitly recognizes intertwined deliberate and habitual processes. The restaurant passage contains a confirmed kilojoule/calorie error; its relative portion effect survives correction. The brain and Stroop illustrations explain concepts but are not intervention trials. Other eating and nutrition statistics in this chapter are not all independently checked in this review, and the review does not certify them merely because their notes are present. A fair reading says that knowledge alone can be insufficient, not that it cannot contribute to behavior change.

5. What About Self-Control?

Book locator: [chapter5.xhtml](#); corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A04; A11; A19; A39.

The chapter usefully separates apparent self-control from constant in-the-moment resistance. The Galla-Duckworth research provides convergent association evidence, including more than a single cross-sectional survey. It does not, by mediation alone, establish that habit is the causal source of the advantage. The vegetable/joystick source was independently method-audited during revision. Its trained response context increased carrot choice in a short laboratory task, a favorable controlled finding. Small selected samples and combined visual/motor changes limit inference about lasting diet or a unique cue mechanism;

see A04 and U04. The running discussion should retain Wood's acknowledgement of difficult portions of the activity. Reduced deliberation about starting does not imply the physical effort of running disappears.

6. Context

Book locator: [chapter6.xhtml](#); corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A11; A37.

This is among the book's more practically useful chapters. It identifies cues, availability and friction as objects of deliberate design. The review credits that advice as written, not as a newly invented program. Several vivid field and personal examples are illustrations rather than isolated causal estimates; the gym-proximity data were not independently obtained. Context redesign can produce benefit by changing opportunity or attractiveness even when no habit measure is available. Access to a controllable environment varies, which matters for applicability but does not make the advice pointless. No separate numerical grade is assigned to the chapter.

7. Repetition

Book locator: [chapter7.xhtml](#); corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A09; A14; A32.

The Lally discussion is substantially more nuanced than a fixed-day habit slogan and earns credit for emphasizing variability. The review specifies the modeled plateau, follow-up and fit-selection details so that neither the 66-day estimate nor the 39 good-fitting curves is turned into a universal rule or success rate. Wood also explicitly distinguishes repetition from excellence and acknowledges that family-dinner associations are not a simple randomized causal demonstration. That protects the book from the criticism that it erases all deliberate practice. The remaining concern is transferring repeated simple health actions into general promises about complex performance.

8. Reward

Book locator: [chapter8.xhtml](#); corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A06; A12; A13.

The emphasis on genuinely rewarding activities is a substantive strength and keeps motivation inside the account. The popcorn experiments provide bounded support for reduced reward sensitivity in familiar action. The monthly weight-incentive example needs repair because it omits the successful group arm and infers failed habit formation without measuring it. Wood marks part of her behavioral explanation as a guess; the review preserves that. The inspected mouse-striatum experiment does not establish a universal human reward deadline. Animal schedule experiments provide qualified favorable evidence for uncertain reinforcement, and the toy-soap pilot has mixed outcomes rather than demonstrated durable automatic washing. A12, S41, S43. Nor do the Volkswagen demonstrations become controlled sustained habit trials merely because they illustrate enjoyable design.

9. Consistency Is for Closers

Book locator: `chapter9.xhtml`; corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A10; A16; A17; A39.

Stable cues, linking a response to an existing routine, and fitting a substitute into an established context are useful recommendations. The note distinguishing stacking from generic implementation intentions matters and is retained. The flossing study's primary methods were not independently inspected here; its results are not used as a fully verified durable intervention estimate. Later cue-planning and flexible-exercise studies suggest that stability need not mean one rigid clock window. The Magritte illustration is rhetorical, not empirical evidence. The chapter is strongest when consistency describes feasible recurring opportunities rather than inflexible adherence despite changed circumstances.

10. Total Control

Book locator: `chapter10.xhtml`; corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A26; A27; A35; A37; A39.

The chapter makes conscious preparation central and provides practical kitchen and financial examples. Its “passive returns for all time” language exceeds the durable comparative evidence reviewed here. The calorie-label note acknowledges contrary findings, and the buffet note explicitly acknowledges a correction; neither qualification should be erased by an adverse review. The remaining buffet concern is inference from observational differences to causal cue management. The sushi-game manuscript was unavailable for independent checking, which is an access issue rather than a confirmed false report. The chapter's integration is valuable, but that integration has not itself been demonstrated as a single tested program.

11. Jump Through Windows

Book locator: `chapter11.xhtml`; corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A17; A20.

The chapter recognizes that changed circumstances can interrupt old patterns and make alternative actions visible. It also acknowledges that useful habits can be disrupted. That is compatible with exploration and person-context fit rather than blanket rigidity. The strike and other natural-change examples were not all independently method-audited, so the review does not assign exact causal effects to them. Discovering a better option and losing an automatic cue can both lead to change, but evidence for one is not proof of the other. Costs and control over the transition constrain applicability.

12. The Special Resilience of Habit

Book locator: `chapter12.xhtml`; corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A18; A19.

The chapter's resilience argument is practically important but not uniform across experimental paradigms. The later stress replications failed to establish the expected control-group behavior, so they cannot support a clean global null. Their targets differ from the

book's 2011 extinction source, an identity distinction maintained throughout the package. Ego-depletion results are separately labeled and not substituted for direct stress evidence. The phishing illustration and safety examples emphasize that persistent routines can be harmful as well as helpful; this favorable balance survives the critique. Neither the chapter nor this review justifies ignoring changes in circumstance because a response is familiar.

13. Contexts of Addiction

Book locator: `chapter13.xhtml`; corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A20; A21; A22; A23; A24; A39.

This consequential extension receives the third scientific grade. Wood explicitly acknowledges multiple causes, deliberate action and complementary biological/environmental explanations. Favorable animal and human network-support evidence supports a real role for context and alternatives. The Vietnam benchmark remains unjustified because outcome definitions and selection differ, and Robins's own discussion warns against the simplified explanation. Petrie's contrary animal findings and Hofford's favorable, non-identical study are both included. The employment discussion accurately acknowledges loss of advantage after contingencies stop. An unpublished study of successful rehabilitation was not independently obtained and cannot establish a general treatment success rate.

14. Happy with Habit

Book locator: `chapter14.xhtml`; corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A28; A29; A30; A31; A32.

Repeated vegetable exposure gives favorable evidence for acceptance, not automatically for habit. Wood distinguishes the ritual EEG finding from an improvement in performance, which is an important accurate description. Two other cited ritual papers have later retraction notices; those source relationships must now be marked, with Tian's mainly definitional role distinguished from Brooks's substantive singing example and the publisher-indexed Brooks notice now incorporated. The review does not assign a retraction to the Norton-Gino grief paper without a verified notice for that particular work. Mere exposure, fluency, ritual meaning and habitual control remain separate constructs. The broad placebo sentence and other clinical-sounding extensions are not independently validated here and should not be treated as general medical guidance.

15. You Are Not Alone

Book locator: `chapter15.xhtml`; corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A27; A33; A35.

The final chapter usefully expands the focus to social and policy conditions. Defaults, infrastructure, prices and information can operate through different mechanisms; their effects are not all demonstrations of learned automaticity. The electricity trial's abstract supports an information-and-learning contribution, limiting an overbroad anti-knowledge reading. Organ-donation, built-environment and tax examples are source-traceable in the book but not all independently method-audited here; no pooled or current policy effect is

manufactured. Historical maps and device illustrations are not described as current 2026 conditions. The burger text overstates the ratio shown in its own figure, a localized descriptive defect.

E. Epilogue

Book locator: `epilogue.xhtml`; corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A02; A37.

The epilogue restates the advantage of aligning repeated behavior with goals and reduces reliance on moral judgments about persistence. That is a defensible practical emphasis. It is evaluated with the already selected C1/C2 propositions rather than counted as a new empirical claim. No guarantee of automatic permanence or universal success is inferred from its encouraging tone. Its broad synthesis is strongest when read with the book's explicit caveats about preparation, reward and mixed control.

P. How to Stop Looking at Your Phone So Often

Book locator: `howtostop.xhtml`; corresponding chapter/application notes. **Access:** complete chapter text read. **Audit links:** A36.

The entire application appendix and its notes were read. The recommendations make the context–friction–replacement framework concrete, but the complete prescription has not been established as an effective package in the scoped search. The burdens of reduced accessibility differ across readers. One option preserves access for dedicated callers; other options involve turning the phone off or leaving it behind. The review assesses the actual stronger restrictions as well as the simple low-cost ones; it does not quietly soften the instructions and then score the revised advice. The ease of understanding these steps is described separately from evidence of benefit.

Consequential figures and visual material

All paths below are under `OEBPS/images/`, with the host chapter named. This is an image-reading record, not certification of the underlying source datasets. Book source: B; specific numerical issue: A35.

Host	File	Content and treatment
3	<code>illus1.jpg</code>	Ngram trends for habit/goal/evaluation; historical language-frequency illustration, not a measure of habit prevalence.
4	<code>Illus2.jpg</code>	Fruit-and-vegetable campaign logo; campaign illustration, not effectiveness evidence.
4	<code>Illus3.jpg</code>	Brain schematic; explanatory illustration, not a measured treatment result.
4	<code>Illus4.jpg</code>	Animal-name task examples; task explanation read with surrounding text.
4	<code>Illus5.jpg</code>	Interference/task material; not a separate intervention study.

Host	File	Content and treatment
5	illus6.jpg	Vegetable/eating-task illustration; relevant source methods/results read during revision; no long-term dietary result inferred.
5	illus7.jpg	Joystick/task illustration; no general dietary outcome inferred from the image.
9	high_res.jpg	Magritte visual illustration; rhetorical/interpretive role, not empirical evidence.
10	mise_en_place.jpg	Prepared-kitchen example; conveys an arrangement, not a randomized result.
10	illus10.jpg	Sushi task/game illustration; associated unpublished manuscript unavailable.
12	Illus11.jpg	Deceptive-message/phishing example; contextual illustration, not independently validated incidence data.
14	Illus12.jpg	Car-design/typicality illustration; not a direct habit-formation measure.
15	Illus13bw.jpg	Historical alcohol-availability map; not represented as a current 2026 map.
15	Illus14.jpg	Historical portion quantities; burger ratio in prose overstates the displayed change.
15	smart_meter.jpg	Electricity-feedback display; distinguish informational feedback from proven automaticity.

Front-matter title/cover and other decorative images were also inspected for identity/navigation where relevant. No print pagination was inferred from image filenames.

6 Claim and source audit

Each dossier distinguishes the book's position from independently examined evidence and the review's inference. Linked source cards specify exactly what was read. "Book source" means the book itself cites or describes that source; "reader-added" means this review introduces it for comparison, not that Wood relied on it. Source fragments are never silently enlarged into full access.

A01. Identity, edition and complete-book access

Book locator: title.xhtml (title image); copyright.xhtml; complete OPF spine.

Book position: The supplied work identifies Wendy Wood as author, the subtitle as The Science of Making Positive Changes That Stick, Farrar, Straus and Giroux in New York as publisher, and First edition, 2019. Copyright is © 2019 Wendy Wood; e-book ISBN is 978-1-250-15908-3.

Source/access map:

- B — full text sections read; reviewed book

The input identity was checked inside the EPUB rather than inferred from its upload filename. The OPF also contains a distinct source identifier, 9781250159076; it is not substituted for the e-book ISBN printed in the copyright section. The OPF modification date is not treated as the publication date.

The reading record covers all 49 spine sections, including all substantive chapters, notes and consequential illustrations; subsequent chapter auditors independently read their complete assigned chapters and notes. The extraction check found no nonblank body text outside the elements read. Nested note duplication was removed before reading. No missing chapter, unread note section or inaccessible book figure remains. This is a reading-coverage claim, not a claim that every source in the bibliography was independently retrieved. The original EPUB is not redistributed in this package.

Judgment confidence: high. **Scoring relationship:** provenance; no numerical deduction.

A02. Strongest fair reading: goals, habits and conscious preparation

Book locator: chapter1.xhtml, “habit that creates persistence”; chapter3.xhtml, “Although some researchers equate”; chapter4.xhtml, “Very few describable parts”; chapter7.xhtml, “Excellence and repetition”; chapter10.xhtml, “If this sounds like a lot of work”.

Book position: Wood argues that habits are central to persistence while explicitly giving conscious goals, planning and monitoring roles. She distinguishes habit from other automatic processes and rejects repetition as a sufficient explanation of excellence.

Source/access map:

- B — full text sections read; reviewed book

The three selected propositions preserve Wood’s actual explanatory and practical scope. They do not replace it with the truism that repetition sometimes occurs. Conversely, “motivation never matters,” “all skilled performance is unconscious,” and “everyone can be cured by moving” are not acceptable summaries of this book.

The qualifications are substantive. They reduce the force of a categorical criticism of Wood’s two-system language. They do not remove the need to test causal claims about why self-control predicts outcomes, whether an intervention builds automaticity, or how far a recovery statistic transfers. The review therefore accepts her integrated reading while questioning specific overextensions. A familiar start and the demanding activity it initiates are evaluated separately.

Judgment confidence: high for attribution; moderate for synthesis. **Scoring relationship:** C1/C2 fair-reading constraints.

A03. The 43 percent estimate and the diary definition of habit

Book locator: chapter2.xhtml, “For a second study” and “Fully 43 percent”; notes-ch2.xhtml, notes 1–2.

Book position: The book presents roughly 43 percent as an estimate of everyday action performed habitually, extending the diary findings to action without conscious thought.

Source/access map:

- B — full text sections read; reviewed book
- S01 — full text sections read; book source: Chapter 2
- S02 — full text sections read; book source: Chapter 2

Wood, Quinn and Kashy classified diary actions through frequency and contextual stability. Their two study estimates were about 35 and 43 percent; the second study analyzed 209 students. The studies also collected reports about attention and concurrent thoughts, so they offer more than frequency counts alone. Even so, the headline percentages classify actions through repetition and context; they are not a population census or an experimental separation of conscious and habitual contributions to every action. The issue is construct scope, not that the 43 figure was invented.

The available 2005 unpublished manuscript extends the investigation to 100 participants aged 17–79 and reports approximately 47 percent under its daily/same-location classification. It is available despite being unpublished; its age distribution is not representative of all ages. Neither report licenses a precise universal claim that 43 percent of behavior is purely unconscious. The review does not treat the diary method as worthless: it captures a practically important regularity while leaving its underlying control processes incompletely separated.

Judgment confidence: high. **Scoring relationship:** C1 scope qualification; RA3 inference.

Revision crosscheck: The book explains the thought measure and explicitly allows thought during habitual action. Study 2 reported 51% noncorresponding and 44% corresponding thoughts within the frequent/same-location category, with some unclassifiable reports. Those participant-level averages must not be multiplied into a new universal percentage. Related thought does not prove deliberate control; unrelated thought alone does not prove a learned cue–response mechanism. The criticism concerns the broadened headline, not invented data or absent thought measurement. Exact book context: chapter 2 ¶12, ¶18–21 and ¶27.

A04. Beneficial habits and self-control: association versus causal explanation

Book locator: chapter5.xhtml, “The good effects that we popularly ascribe”; notes-ch5.xhtml and notes-ch7.xhtml entries for Galla and Duckworth.

Book position: Wood attributes much of the benefit popularly ascribed to self-control to useful habits, while allowing goals to orient their formation.

Source/access map:

- S37 — revision evidence; exact access and relationship in source record
- S38 — revision evidence; exact access and relationship in source record
- U04 — revision evidence; exact access and relationship in source record
- B — full text sections read; reviewed book
- S04 — relevant primary methods/results/discussion across the six studies read during revision; initial checks used the abstract and overview; book source: Chapter 5

Galla and Duckworth's overview reports six studies, total $N=2,274$, connecting self-control, beneficial habits and outcomes across several domains. It includes prospective and behavioral components; it should not be dismissed as one cross-sectional questionnaire. The source supports the claim that habit measures account statistically for part of the observed relationship.

The causal interpretation remains weaker. A mediation model cannot on its own rule out common causes of habit reports and outcomes, reciprocal influence, or overlap among measures of consistent behavior. The subsequent chapter audit inspected the relevant methods/results/discussion, including the meditation study: 132 enrolled and 109 completed all three assessments. The designs were nonexperimental with respect to forming the beneficial habits, despite prospective and behavioral components. This supports a bounded favorable statement about convergence, with a criticism of treating it as an established account of how much self-control benefit habit causes. Randomizing a habit-building intervention would answer a different question from observing who has beneficial routines.

Judgment confidence: moderate. **Scoring relationship:** C1; RA3—not a claim that the source is fabricated.

Additional original evidence: The chapter's later transition from messy everyday research to controlled laboratory testing introduces de Ridder et al.'s 102-study meta-analysis. Its detailed automatic/controlled comparison used the Self-Control Scale subset: 50 studies, including 34 descriptive, 12 experimental and four mixed designs. Researchers coded behavior categories rather than administering one common automaticity test. For desired behaviors, trait self-control correlations were .36 for automatic and .15 for controlled categories. The authors themselves proposed habits as a possible explanation; Wood accurately retains their word "may." The misleading part is the narrative's increased causal certainty, not a literal statement that all 102 studies were randomized laboratory trials. S37. Book: chapter 5 ¶46–50, notes 16–17.

Initiation and effort: Phillips and Gardner separated deciding to exercise from executing its routine in a prospective study of 123 university students/staff, with 119 meeting diary-completion criteria. Only instigation had a unique association with reported frequency. This does not measure pain, muscular effort or moment-by-moment deliberation; frequency as the outcome may also favor instigation. Wood acknowledges difficult first and last miles in

her six-mile-run example (chapter 5 ¶40). The criticism concerns her diagnosis that continued self-control effort means the runner is doing it wrong, not an invented claim that muscles cease working. S38.

Controlled food-choice finding: Lin, Wood and Monterosso's Study 2 retained 53 of 56 women after exclusions. Two training sessions were 2–4 days apart. In the test, carrot choice was 55% with the trained context versus 37% with a new context. Visual cues and response direction changed together; decisions had a two-second limit. All participants received the same preliminary suppression task, so this is not a randomized depletion contrast. It is favorable evidence for the trained response context under those conditions, not lasting diet improvement or proof of choice without any consideration. U04.

A05. Definitions and experimental identification

Book locator: chapter3.xhtml, “automaticity”; chapter4.xhtml, mixed neural control; chapter8.xhtml, “For scientists, insensitivity to reward”.

Book position: Wood distinguishes habit from other automatic processes and describes reduced sensitivity to reward as a scientific indicator. Elsewhere, familiar repetition and skills support a broader narrative.

Source/access map:

- S33 — revision evidence; exact access and relationship in source record
- S34 — revision evidence; exact access and relationship in source record
- S35 — revision evidence; exact access and relationship in source record
- B — full text sections read; reviewed book
- S06 — full text sections read; reader-added comparison
- S07 — full text sections read; reader-added comparison; later evidence
- S08 — full text sections read; reader-added author position; later evidence

Four constructs must remain separate: repetition frequency; self-reported automaticity; cue-triggered preparation or initiation; and persistence after outcome devaluation. A fifth, computational perseveration, is model-defined rather than a direct behavioral assay. They may converge without becoming interchangeable.

Devaluation tests change an outcome's desirability after learning and assess adjustment under specified test conditions. Continued responding is interpretable only when learning, devaluation and response testing are adequate. A goal-directed override can also prevent a prepared habitual response from appearing in final behavior. Thus neither a questionnaire plateau nor a null devaluation interaction settles every question about habitual control. These distinctions are analytic standards used in the review, not a new demand that useful practical interventions must pass one laboratory test before they can help anyone.

Judgment confidence: high. **Scoring relationship:** C1/C2; avoids both overclaiming and an artificially narrow definition of benefit.

A06. Popcorn: bounded evidence of cue-sensitive persistence

Book locator: chapter8.xhtml, stale-popcorn and usual-hand discussion; notes-ch8.xhtml entry for Neal et al. (2011).

Book position: Wood uses eating stale popcorn in a familiar context, and disruption of the usual response, to illustrate habits persisting despite reduced reward.

Source/access map:

- B — full text sections read; reviewed book
- S05 — full text sections read; book source: Chapter 8

Study 1 involved 98 cinema participants and 60 meeting-room participants; the setting samples were recruited separately, while popcorn freshness was manipulated. Study 2 involved 89 participants and manipulated the eating response as well as freshness. The overall interaction patterns support context/response dependence in people with stronger pre-existing popcorn habits.

Several details limit a categorical account. A nonsignificant fresh–stale contrast is not proof of equal consumption. In Study 2, a relevant conditional contrast was marginal, and adjustment for liking weakened the central interaction. The two experiments are conceptual extensions within one paper, not independent multilaboratory replications. They are compatible with habitual responding despite reduced enjoyment but do not establish a sustained health effect or that habit completely overrides preference. Wood explicitly warns that changing hands is not a dieting method and could increase intake of liked food. That qualification already belongs to her account and must be credited. The remaining limit is how far the local finding supports broader explanations.

Judgment confidence: moderate. **Scoring relationship:** C1 favorable evidence; C2 transfer limit.

A07. Training-induced habit: failures and a multilaboratory test

Book locator: chapter3.xhtml–chapter4.xhtml learning-system argument; chapter7.xhtml repetition; reviewer-added tests not cited by Wood as their own evidence.

Book position: The book presents repetition as a route toward cue-driven habitual control, with many daily-life examples.

Source/access map:

- B — full text sections read; reviewed book
- S06 — full text sections read; reader-added comparison
- S07 — full text sections read; reader-added comparison; later evidence

The 2018 de Wit paper reports five failures across four paradigms to obtain the expected experimental habit-induction effect. This review read the first two experiments in detail and uses the abstract for the five-study summary, not as a substitute for a complete methods audit. Pool and colleagues' later 2022 collaboration tested 327 people and analyzed 306 across four laboratories and five experiments. Its one-versus-three-day training manipulation did not produce the expected overall difference in devaluation sensitivity.

Pool's mixture analyses and stress-related subgroup findings were exploratory and are not promoted into confirmed moderators. These results weaken the assumption that the laboratory transition follows reliably from a simple increase in practice. They neither disprove everyday habitual control nor establish that extra repetition is useless for all outcomes. The useful question is whether the operationalization captures the proposed process, not which side can claim that one null result settled the field.

Judgment confidence: high for reported outcome; moderate for theory implication. **Scoring relationship:** C1/C2 qualification; 2022 marked later.

A08. The author's published response and preserved goal control

Book locator: chapter4.xhtml, "our goal-directed and habit neural systems are interconnected"; later author-position paper.

Book position: Wood already allows mixed control in the book; a later paper develops the distinction between habit preparation and overt responding.

Source/access map:

- B — full text sections read; reviewed book
- S08 — full text sections read; reader-added author position; later evidence

Wood, Mazar and Neal argue that a cue can prepare a familiar response while a goal-based process prevents its execution. Their paper reports a secondary analysis of other researchers' task data supporting competition between these influences. This is a published scientific position, not an author response commissioned for this review.

The argument blocks an overly strong inference from an unexpressed habit to the absence of habit learning. It does not automatically rescue every failed prediction: hypotheses about hidden preparation need independently discriminating measures. The review credits the response as an important qualification, without treating the author's interpretation as an independent replication of the whole book. No underlying task data were downloaded or reanalyzed by this review.

Judgment confidence: moderate. **Scoring relationship:** C1 favorable qualification; later evidence.

A09. Habit-formation time: what 66 days means

Book locator: chapter7.xhtml, “Research provides better insight”; notes-ch7.xhtml, note 3.

Book position: Wood reports substantial variation in the time needed for new behavior to feel automatic and discusses the well-known Lally study rather than endorsing a fixed 21-day rule.

Source/access map:

- B — full text sections read; reviewed book
- S03 — full text sections read; book source: Chapter 7

Ninety-six participants began; 82 supplied sufficient data to attempt modeling, 62 models were fitted, and 39 met the study’s good-fit criterion for an asymptotic curve. The observation period was 84 days. The median 66-day estimate concerned approaching 95 percent of each fitted individual plateau; the 18–254-day range includes extrapolation beyond observed follow-up. Automaticity was self-reported, not measured through devaluation.

The 39/82 figure is not a validated habit-formation success rate. Lack of an acceptable curve fit is not equivalent to failure to develop a useful routine. Similarly, the single-missed-opportunity analysis cannot establish that any length of interruption is harmless. The study is useful longitudinal evidence about change and variability under motivated, self-selected repetition. It cannot prescribe a universal deadline or isolate the causal benefit of the instruction without a comparison group. Wood deserves credit for conveying variation rather than reducing the finding to a guaranteed calendar rule.

Judgment confidence: high. **Scoring relationship:** C2 favorable but bounded; RA2 credit for nuance.

A10. Routine cues versus clock-time cues

Book locator: chapter9.xhtml consistency/stacking; notes-ch9.xhtml, note 18; reader-added trial.

Book position: Wood advocates stable cues and distinguishes stacking onto a preceding action from generic implementation-intention planning.

Source/access map:

- B — full text sections read; reviewed book
- S09 — full text sections read; reader-added intervention; later evidence

Keller and colleagues enrolled 192 adults in a randomized routine-cue versus time-cue trial for a chosen nutrition behavior over 84 days. The longitudinal analysis retained 135 participants (65 routine-cue, 70 time-cue) who met diary-completion requirements. Both

groups received planning support. The reported gains in self-reported automaticity did not show a clear advantage for one cue condition over the other.

The book's published stacking study is a separate comparison that also included inhibition and substitution instructions; see A11 and S42. The Keller comparison supports flexibility in how readers choose a cue but not equivalence under every circumstance. Without a no-planning group, common improvement cannot identify the unique benefit of cue planning, as opposed to monitoring, repeated behavior or time. The trial is a later, component-level test, not validation of the book's package. Its modeled formation-time estimates are not substituted for Lally's results or offered as a second universal deadline.

Judgment confidence: moderate. **Scoring relationship:** C2 favorable component evidence; later 2021.

A11. Friction and practical fit

Book locator: chapter6.xhtml, driving/restraining forces; chapter9.xhtml, product replacement; chapter10.xhtml, kitchen preparation.

Book position: The book recommends reducing obstacles to wanted actions, adding obstacles to unwanted ones, and arranging an environment that makes a preferred response easier.

Source/access map:

- S39 — revision evidence; exact access and relationship in source record
- S42 — revision evidence; exact access and relationship in source record
- B — full text sections read; reviewed book

This is advice the book actually gives, not a revision invented by the reviewer. Its examples connect locations, preparation, availability and response effort to repeated opportunities. They also expose a limit: a clear instruction to change an environment does not establish that a reader can afford or control that change.

The gym-distance example is not treated here as a randomized estimate of the effect of moving closer to a gym. Its underlying data were not independently obtained. The recommendation remains plausible and is supported indirectly by the intervention evidence elsewhere in this review; the anecdote's numerical precision is not needed to award that qualified credit. A recurring action may improve through convenience or greater willingness even before a habit association is established. The score concerns useful results at the advice's real scope, not whether every benefit can be assigned uniquely to habit.

Judgment confidence: high for attribution; moderate for usefulness judgment. **Scoring relationship:** C2/PV1; burden considered separately.

Field example requiring a narrower story: Kirchner et al. followed 475 quitline smokers for a month. GPS exposure meant being within 30 metres of a retailer; craving was summarized across daily reports. Retailer exposure was associated with lapses even on days with zero average reported craving. It did not establish an individual's state immediately before store

entry, observe a pack purchase, or show that cues alone caused sustained relapse. The book's "could" and "Maybe" acknowledge an illustrative reconstruction, but its statement that cues alone led to relapse exceeds those observations. Keep the real zero-craving-day association; do not substitute a fabricated-data accusation. Book: chapter 6 ¶129–33, especially ¶132, notes-ch6_21. S39.

Published replacement experiment: Labrecque et al.'s laundry-product study assigned 70 students to control, standard implementation intentions, or habit-cued planning. The latter also instructed inhibition of the old response and substitution of the new response. Reported mean uses over four weeks were 11.17, 11.05 and 13.28, with partial bottle-weight corroboration. This is favorable intervention evidence, but the bundled design does not isolate automaticity transferred from an old action to a new one. Book: chapter 9, notes 16–19. S42.

A12. Immediate and uncertain rewards: human prescription exceeds the narrow source

Book locator: chapter8.xhtml, ¶15–17, ¶22, ¶35; "All of this means"; notes 7 and 19.

Book position: Wood emphasizes timely enjoyment and uncertainty, acknowledges that payment can motivate work, and makes a stronger claim that delayed, predictable payment prevents dopamine from doing its work in habit learning.

Source/access map:

- B — full chapter and notes read
- U01 — full relevant primary methods/results/discussion read during revision; book source
- S40 — full relevant review sections read; mechanistic context
- S41 — full relevant primary sections read; book's animal schedule evidence

Shindou et al. used adult mouse brain slices, controlled synaptic stimulation and dopamine uncaging. A two-second delay supported potentiation under the relevant protocol. The paper did not compare human pay schedules or everyday habit formation. Its concluding account allows longer-term reward prediction through learning; the short synaptic interval is not the same clock as the elapsed time until a person's final payoff. Schultz's account likewise includes signals at reward-predictive cues and responses involving delayed rewards.

There is real favorable evidence for reinforcement schedules. DeRusso et al. compared fixed and random intervals in mice; after extended training, random-interval responding resisted devaluation and persisted longer under omission. These findings do not establish a universal human requirement for surprising rewards, a salary rule, or a prescription for rewarding a partner.

The warranted criticism is the untested bridge from cellular timing and animal schedules to broad human instructions. Immediate enjoyment can help, and Wood gives it appropriate practical emphasis. The inspected sources neither establish her categorical deadline nor show that every delayed paycheck produces automatic habits. Initial metadata-only access has been superseded for U01, without changing the access history.

Judgment confidence: high for source scope; moderate for human generalization. **Scoring relationship:** C2/RA3; qualified favorable evidence retained.

A13. Monthly incentives: omitted successful arm and unmeasured habit

Book locator: chapter8.xhtml, “Consider a six-month weight-loss program”; “My guess is”; “new habits didn’t form”; notes-ch8.xhtml, note 13.

Book position: Wood describes limited weight loss under monthly individual incentives, speculates about pre-weigh-in dieting, and concludes that new habits did not form. The discussion does not report the study’s successful group-incentive condition.

Source/access map:

- B — full text sections read; reviewed book
- S11 — full text sections read; book source: Chapter 8, note 13

The trial randomized 105 employees to three 35-person groups: control, individual incentives, or group incentives. In the main imputed analysis, 24-week losses were 1.1, 3.7 and 10.7 pounds respectively. Both incentive arms paid monthly; the group arm divided a \$500 pool among successful members of anonymous five-person groups. The individual-control contrast was uncertain ($p=.34$); group incentives outperformed both. At 36 weeks, 12 weeks after incentives ended, the group-control difference remained statistically detectable; group-individual did not pass the multiplicity-adjusted threshold. Observed-completer curves and imputed estimates are different analyses and should not be mixed to manufacture an additional numerical error.

Habit automaticity was not measured. Wood explicitly labels her dieting-pattern explanation a guess; that attribution must survive. The problem is the subsequent explanatory conclusion and the missing arm, which changes the lesson from “monthly incentive timing did not build habits” to “incentive structure mattered in this setting.” The trial does not prove the group arm worked by habit, either. This is a substantive source-selection/inference problem, not evidence that rewards universally work or fail.

Judgment confidence: high. **Scoring relationship:** RA2 completeness; RA3 inference; distinct issues, not duplicate arithmetic penalties.

A14. Ten Top Tips: genuine short-term advantage and maintained loss

Book locator: chapter7.xhtml–chapter10.xhtml practical principles; reader-added program evidence, not a direct citation attributed to Wood.

Book position: The book’s habit-oriented framework can be tested against controlled programs using overlapping principles; this trial is an external comparison, not a test of reading the book.

Source/access map:

- N05 — revision evidence; exact access and relationship in source record
- B — full text sections read; reviewed book
- S12 — full text sections read; reader-added intervention

The English primary-care trial randomized 537 adults with obesity across 14 practices: 267 to Ten Top Tips and 270 to usual care. Three-month follow-up reached 389 participants, with 383 primary weight outcomes. The adjusted weight-loss advantage was 0.87 kg (95% CI 0.27–1.47 kg favoring the intervention). At 24 months, mean losses were 2.15 kg in Ten Top Tips and 2.96 kg in usual care, without a retained intervention advantage. The abstract reports 312 follow-up participants; the indexed flowchart totals 313 and the adjusted Table 5 model uses 290. Those are retained as distinct reported denominators; the one-person abstract/flowchart discrepancy was not resolved and is not treated as a fabricated count.

This is favorable evidence for a brief habit-informed intervention, not evidence of no benefit. Usual care could include referrals to effective programs, so convergence is not equivalent to untreated regression. The intervention combined a consultation, leaflet, logbook and food-label guidance, with further materials at three months. It cannot isolate cue learning from monitoring, information or contact. Maintenance within an arm and superiority over another arm must be reported separately.

Judgment confidence: high. **Scoring relationship:** PV1 positive; PV2 duration/generalization qualification.

Correction: A formal notice published online 7 June 2021 changes Table 1 cholesterol/glucose units from mg/dl to mmol/l. It does not revise the weight-loss results used here. The complete notice was read; this is a unit correction, not a retraction or an outcome reversal. N05. During revision the institutional abstract and indexed original results/flowchart/Table 5 were checked; earlier full-text access remains its own recorded stage.

A15. Smaller weight-loss trial: preserve the comparator over time

Book locator: chapter7.xhtml–chapter10.xhtml; reader-added intervention evidence.

Book position: Other habit-oriented packages provide a test of related practical components, not necessarily Wood's integrated program.

Source/access map:

- B — full text sections read; reviewed book
- S13 — abstract only; reader-added intervention

Cleo and colleagues' abstract reports 75 adults assigned to Ten Top Tips, Do Something Different, or a waiting list. Twelve-week losses were 3.3, 2.9 and 0.4 kg respectively; 66 participants completed that stage. The longer follow-up tracked intervention participants, with 43 of 50 retained at 12 months, rather than retaining an untreated comparison over that period.

The short-term difference is favorable. Long-term improvement among intervention completers cannot be called a randomized long-term advantage over no treatment. Nor can a percentage among follow-up completers be presented as the success probability for everyone starting the program. Access was abstract only: allocation concealment, outcome assessment, missing-data decisions and detailed harms were not independently audited. This study adds qualified support without supporting a high-certainty whole-program claim.

Judgment confidence: moderate. **Scoring relationship:** PV1 corroboration; PV2 limit.

A16. Flexible versus narrowly routinized exercise incentives

Book locator: chapter9.xhtml consistency advice; reader-added later boundary test.

Book position: Wood emphasizes recurring cues and stable circumstances; that should not be equated with a requirement that all exercise occur inside one narrow clock window.

Source/access map:

- B — full text sections read; reviewed book
- S14 — full text sections read; reader-added intervention; later evidence

The workplace study involved 2,508 employees across seven sites, comparing no incentive and two payment levels under flexible versus two-hour routine-window incentives. All chose a two-hour workout window and received reminders; the manipulation changed which visits earned the original incentives. That incentive period lasted four weeks. The main analysis covered the next four weeks, with further analyses for post-intervention weeks 5–10 and 11–40. Flexible conditions retained an advantage in weeks 5–10 (0.11 more visits per week, $p < .01$); differences from routine conditions were no longer statistically significant in weeks 11–40. Daily reminders and common monthly badge-swipe lotteries continued. This was not withdrawal of all external support or all payments.

The result challenges a rigid interpretation of consistency, not the usefulness of contextual cues. Its outcome is attendance, not experimentally established automaticity; the workplace's gym access and follow-up duration constrain generalization. Because the conditions differ in the opportunities they reward, the contrast should not be described as a pure test of whether people possess habits. It supports matching the schedule to feasible opportunities rather than inferring that a stable clock time is always optimal. This is later evidence, not something Wood was obliged to discuss in 2019.

Judgment confidence: moderate. **Scoring relationship:** C2/PV2 boundary condition.

A17. Disruption, exploration and replacement

Book locator: chapter11.xhtml, London Underground strike discussion and context discontinuities; chapter9.xhtml, replacement behavior.

Book position: Wood argues that changes in context can disrupt existing habits and make alternatives easier to discover, while recognizing that useful routines can also be lost.

Source/access map:

- B — full text sections read; reviewed book
- U06 — primary accepted-paper abstract and authors' account read; no full methods/model-estimate audit; book source

The chapter's strongest idea is an opportunity claim: disruption can make an established response less inevitable and prompt reconsideration. Its strike, moving and other natural-change examples do not themselves provide a direct measure of automaticity. The chapter audit read the primary accepted-paper abstract and an authors' account. Full methods and model estimates were not independently audited, and no precise commuted effect size is endorsed here.

There are two different possible benefits: interrupting a cue-response association and discovering a genuinely better option. Either can be valuable. Improvement after a disruption does not distinguish them without further evidence. The review credits Wood's attention to exploration and her recognition that disruption can remove good habits, rather than portraying her as recommending change for its own sake. Feasible replacement opportunities are relevant to applicability, especially when readers cannot readily alter jobs, housing or responsibilities.

Judgment confidence: high for attribution; limited for unexamined natural experiments.
Scoring relationship: C2/PV2; no primary-error finding assigned.

A18. Stress: correct replication target and an inadequate critical contrast

Book locator: chapter12.xhtml; notes-ch12.xhtml, note 4 (Schwabe & Wolf, 2011, extinction) and note 7 (2013 review).

Book position: Wood presents habitual responses as resilient when conscious control is strained, including stress-related evidence.

Source/access map:

- S45 — revision evidence; exact access and relationship in source record
- B — full text sections read; reviewed book
- S22 — full text sections read; reader-added replication comparison; later evidence

Smeets and colleagues' two preregistered replications concern Schwabe and Wolf's 2009 and 2010 outcome-devaluation paradigms. They do not directly replicate the 2011 extinction article specifically cited in Chapter 12 note 4. Their abstract reports successful stress induction and learning/devaluation, but both stressed and non-stressed groups responded without the expected distinction between valued and devalued outcomes.

Because controls did not establish the anticipated goal-directed pattern, the critical stress-induced shift was not cleanly testable. This supports caution about robustness and assay conditions, not a categorical conclusion that stress has no effect. Planned sample sizes in the

methods are not reported as achieved analytic Ns here. Initial checks had limited results access; the chapter audit subsequently inspected relevant primary methods/results/discussion. No original data were reanalyzed. The evidence is explicitly later (2023), and it is not used as a stand-in for direct replication of every stress or depletion experiment in the book.

Judgment confidence: high for identities; moderate for inference. **Scoring relationship:** C1 boundary condition; not a fourth SA claim.

Original experiment in the book: Schwabe and Wolf (2011) randomized brief socially evaluated cold-water stress versus warm-water control in 64 students retained after reward-preference exclusions. Rewarded choices for drinks and extinction blocks alternated. Stressed participants persisted in some extinction blocks; the overall three-way interaction was $p=.09$, with narrower block effects. The authors explicitly cautioned that persistence did not identify habit formation with certainty. The finding is neither “no stress effect” nor proof that every established habit is protected under prolonged crisis. Relevant full primary methods/results/discussion were read during the chapter audit. S45.

A19. Ego depletion is not the same question as stress or habit resilience

Book locator: chapter5.xhtml willpower framing; chapter12.xhtml low-control discussion; reviewer-added later test.

Book position: The book contrasts effortful control with the resilience of established responses; evidence on depletion is relevant background but cannot substitute for habit-specific tests.

Source/access map:

- B — full text sections read; reviewed book
- S23 — indexed primary methods/results/discussion read during chapter audit; initial abstract access upgraded

Vohs and colleagues' preregistered project involved 36 laboratories and 3,531 recruited participants. Preregistered exclusions left 2,463, with 2,461 in the combined confirmatory outcome. Its confirmatory analysis reported a nonsignificant depletion effect, $d=.06$; an exploratory analysis using a broader sample yielded a small $d=.08$. These are findings about performance after prior self-control exertion under the tested procedures.

They do not directly estimate what acute stress does to habits, and they cannot by themselves overturn Wood's habit-resilience experiments. The review uses them only to limit a broad resource-depletion narrative. It does not imply that effort or fatigue is imaginary, that motivation cannot change, or that self-control always has zero carry-over effects. Initial checks used the abstract; the chapter audit additionally inspected indexed primary methods/results/discussion. This remains a bounded test, not a full adjudication of the depletion literature.

Judgment confidence: moderate. **Scoring relationship:** background qualification; no independent score subtraction.

A20. Vietnam: real recovery, non-equivalent comparator and causal overreach

Book locator: chapter13.xhtml, “Robins tracked 470 enlisted men”; “From our habit perspective”; “Why can’t we shoot for 5 percent?”; notes-ch13.xhtml, notes 13 and 15–23.

Book position: Wood reports low re-addiction after return from Vietnam and proposes the roughly 5% figure as a benchmark against the 40–60% relapse figure she gives for conventional treatment. She acknowledges nonrandomization and multifactorial addiction.

Source/access map:

- S32 — revision evidence; exact access and relationship in source record
- B — full text sections read; reviewed book
- S15 — Initially model-reported full access, not independently reproduced; retained historical source identity
- S16 — full text sections read; book source: Chapter 13

Robins’s reports concern a representative returning-enlisted-men sample plus oversampling based on positive urine screening, with weighted population estimates. Among those addicted in Vietnam, approximately 5% were re-addicted in the first year and 12% at some point over three years. These are not rates of all subsequent heroin use. The 2010 article is a reprint of an earlier report from the same project, not independent contemporary replication.

The original cohort report, reprinted in 2010, cautions against interpreting treated-versus-untreated comparisons as evidence that treatment was useless: treatment was brief and more severe cases selected into it. This attribution is based on S16; the initial full-access claim for the 1993 paper was not independently reproduced. The observed changes also do not isolate a conditioned-setting explanation. Many veterans reused heroin without becoming re-addicted; treatment cohorts are selected for persistent or severe problems. Returning home also changes multiple exposures and incentives at once. Wood’s acknowledgement that the contrast is uncontrolled is fair and important, but does not make its numerical benchmark transferable. Wood explicitly acknowledges heroin reuse and the uncontrolled comparison; she does not conceal them. The review rejects the treatment-benchmark transfer, not the observed recovery or the possibility of environmental help. No single uniform follow-up period is asserted for the generic 40–60% figure. The 40–60% figure is attributed to Wood, not endorsed as a current rate for all care.

Judgment confidence: high. **Scoring relationship:** C3 partial support; RA3 inference; PV3 high-stakes boundary.

Separate pre-book source problem: Chapter 13 ¶12, note 9, describes twelve-step programs as no better than standard treatments. Crits-Christoph et al.’s 1999 trial randomized 487 cocaine-dependent patients among four treatment packages. Individual drug counseling

grounded in twelve-step principles plus group counseling produced the strongest drug-use outcomes; psychotherapy retained participants better. All arms included group counseling encouraging twelve-step participation, and the three individual-plus-group arms had the same scheduled session frequency. It does not isolate twelve-step philosophy or community AA, compare medication, or support a general equivalence claim. This is a source-description/inference issue, not later evidence Wood could not have known. S32.

A21. Rat Park: a pre-book replication failure with important qualifications

Book locator: chapter13.xhtml, enriched rat-housing discussion; notes-ch13.xhtml, note 24.

Book position: Wood uses enriched animal environments to support the importance of alternatives and context in drug use.

Source/access map:

- B — full text sections read; reviewed book
- S17 — full text sections read; reader-added replication comparison

Petrie reported two studies with 40 Wistar rats each, comparing animals reared in isolated versus colony housing. The earlier large housing difference in sweetened oral morphine consumption was not reproduced. The isolated animals in these studies consumed relatively little morphine. In the second study, colony consumption was measured in aggregate, preventing an individual-level inferential analysis of that group.

The paper proposes differences associated with a supplier's colony change as a possible explanation; it does not establish a genetic cause. This is pre-book contrary evidence that limits treating Rat Park as a simple, universally reproduced demonstration. It does not establish that enrichment never affects drug taking. Differences in animal population, measurement, drug, dose and administration route matter. The review therefore presents both this failure and the favorable, different enrichment experiment in A22, without turning either into a human recovery rate.

Judgment confidence: moderate. **Scoring relationship:** C3 evidence balance; not a misconduct or retraction allegation.

A22. Favorable enrichment evidence that Wood actually cites

Book locator: chapter13.xhtml and notes-ch13.xhtml, note 24, Hofford et al. (2017).

Book position: The book does not rely only on the original Rat Park account: its note includes later work on remifentanil self-administration.

Source/access map:

- B — full text sections read; reviewed book
- S18 — abstract only; book source: Chapter 13, note 24

Hofford and colleagues' primary abstract describes male rats reared in enriched, standard or isolated conditions and subsequently tested on intravenous remifentanyl self-administration. Enriched animals showed reduced responding under the tested conditions. This is favorable evidence that rearing environment can influence drug-taking behavior, and Wood's citation deserves acknowledgment.

It is not an exact replication of the original sweetened-oral-morphine experiments. The species-level result also cannot quantify recovery in a human treatment population. Because only the primary abstract was read, this review does not supply an unverified sample size or certify every analysis. The properly bounded lesson is that environmental influence survives scrutiny as a plausible and experimentally supported component, while the general clinical promise and the numerical Vietnam comparison need separate evidence.

Judgment confidence: moderate. **Scoring relationship:** C3 favorable component; keeps adverse search balanced.

A23. Human network-support trials: useful benefits, multiple mechanisms

Book locator: chapter13.xhtml, network-support treatment; notes-ch13.xhtml, note 35; reader-added Litt et al. (2016).

Book position: Wood discusses changing social surroundings to support recovery, including network support rather than only removal of drug cues.

Source/access map:

- B — full text sections read; reviewed book
- S19 — full relevant primary methods/results read during chapter audit; book source: Chapter 13, note 35
- S20 — full text sections read; reader-added favorable follow-up

The 2009 trial's abstract reports 210 participants assigned to network support, network support plus contingency management, or case management, with favorable drinking outcomes extending to two years after treatment. The later 2016 trial compared 12 weeks of network support with an active packaged cognitive-behavioral treatment in 193 participants, followed for 27 months. Network support improved proportion of abstinent days and drinking consequences, while other reported drinking endpoints improved similarly in both arms.

AA participation was one means of building abstinence-supportive networks in this treatment; network support should not be portrayed as an entirely separate alternative to established recovery groups. These are meaningful favorable clinical findings, not merely observational correlations between good friendships and recovery. The interventions can affect coping, confidence, opportunities and reinforcement as well as cues. They do not isolate habit as the mechanism or establish uniform superiority over care. Source-specific

access remains visible: the combined audit read the 2009 trial's relevant full sections; the later trial was checked at the sections stated in S20. Neither dataset was independently reanalyzed. The review does not translate an ambiguous "20% more days" phrase into percentage points, an absolute abstinence probability, or a claim of cure.

Judgment confidence: moderate. **Scoring relationship:** C3/PV1 favorable; against dismissing the entire addiction chapter.

A24. Therapeutic employment: benefits while support is maintained

Book locator: chapter13.xhtml, therapeutic workplace and loss of effects after support ends; notes-ch13.xhtml, note 37.

Book position: Wood describes employment-based reinforcement and acknowledges that benefits can diminish when the supporting conditions end.

Source/access map:

- B — full text sections read; reviewed book
- S21 — abstract only; book source: Chapter 13, note 37

The primary abstract reports 51 qualified participants randomized after a training/selection period: 24 to employment alone and 27 to employment with abstinence requirements. During the intervention, cocaine-negative urine-test proportions favored the contingency condition; the subsequent year did not show a retained between-group advantage. These are urine-sample outcomes in a selected group, not population cure rates.

This supports both a potential practical benefit and a maintenance limitation. The book explicitly acknowledges the latter and should not be accused of hiding it. The review does not demand that a useful intervention remain effective after withdrawal to earn any credit. Sustained support can be worthwhile, but its ongoing demands belong in the benefit-burden assessment. Nor can the trial establish that dependence has been permanently replaced by an autonomous habit of abstinence.

Judgment confidence: moderate. **Scoring relationship:** C3/PV2 balanced; no invented requirement for support-free benefit.

A25. Confirmed unit error in the restaurant portion example

Book locator: chapter4.xhtml, paragraph beginning "What's fascinating" and containing "In a cafeteria study"; notes-ch4.xhtml, note 13.

Book position: Wood labels roughly 1,800/1,700 and 2,600/2,400 as calories in a standard-versus-larger pasta-and-cheese comparison.

Source/access map:

- B — full text sections read; reviewed book
- S10 — full text sections read; book source: Chapter 4, note 13

The study involved 180 patrons (89 standard-portion and 91 larger-portion observations) over ten days. Portion size was randomized by study week, not individually by patron; patrons chose whether to order the dish. The standard entrée contained 1,766 kJ, approximately 422 kcal; average consumption was 1,671 kJ, approximately 399 kcal. The larger entrée contained 2,647 kJ, approximately 633 kcal; average consumption was 2,390 kJ, approximately 571 kcal. The book has used rounded kilojoule values as calories.

The relative increase in entrée consumption, about 43%, remains valid. The absolute difference is about 172 kcal for the entrée, not roughly 700 kcal. Whole-meal intake is a separate outcome, with a smaller relative increase, and must not be substituted without saying so. This is a confirmed factual description error with a recoverable source, not a failed retrieval. It affects Reference Accuracy's description input; it does not negate the finding that larger portions increased intake under these conditions.

Judgment confidence: high. **Scoring relationship:** RA2; no separate C1/C2 deduction for the same unit mistake.

A26. Chinese-buffet correction: disclose it, then evaluate the inference

Book locator: chapter10.xhtml, buffet seating/browsing passage; notes-ch10.xhtml, note 13.

Book position: The book's note explicitly says its numbers come from the corrigendum and cites a public checking account. The prose treats thinner diners' choices as examples of managing cues and friction.

Source/access map:

- B — full text sections read; reviewed book
- N04 — full text sections read; book source: Chapter 10 note 13 explicitly cites correction
- U09 — metadata only; book source; correction separately examined as N04

The correction clarifies numbers of observations and missing values in Table 2. It reports varying row denominators rather than a single count for every variable. Its authors state that the adjustments do not alter their conclusions. This review read the correction and table; it did not reproduce the underlying observational analysis or independently read the complete original article.

It would be wrong to say that Wood ignored the correction or to label this particular paper retracted without a verified notice. The surviving concern is causal inference. Observed seating, browsing or napkin differences between body-size groups do not establish that those choices produced lower weight, or that they operated through habit. The passage can

motivate a plausible example without establishing the effect of prescribing that example. The correction's author assurance is reported as their position, not independent scientific sign-off.

Judgment confidence: high for notice/attribution; moderate for source interpretation.

Scoring relationship: RA2 favorable disclosure; RA3 inference qualification.

A27. Information is insufficient, not necessarily ineffective

Book locator: chapter4.xhtml, knowledge discussion; chapter10.xhtml, calorie labeling; notes-ch10.xhtml, note 2; chapter15.xhtml, electricity feedback.

Book position: Wood contrasts information-centered approaches with changing conditions, but her notes acknowledge contrary calorie-label evidence and her final chapter gives information a practical role.

Source/access map:

- B — full text sections read; reviewed book
- S26 — abstract only; book source: Chapter 15
- U07 — metadata only; book source, not primary-verified here

Chapter 10 note 2 includes an explicit “although see” reference to Zlatevska, Neumann and Dubelaar’s 2018 analysis. This review did not independently read that entire review and does not adopt its numerical results. The presence of the qualifying citation nevertheless matters when assessing Wood’s fairness: it is not accurate to say that she gives readers no indication of competing evidence.

Chapter 15’s electricity example also prevents a coherent reading in which information can never change behavior. The primary abstract reports a randomized information/price intervention and learning-related interpretation. The appropriate distinction is between providing information and solving all the obstacles to acting on it. Failure of information alone in one setting does not show that information has no useful role elsewhere, and a useful information intervention does not disprove every claim about habit.

Judgment confidence: high for attribution; limited for unexamined synthesis. **Scoring relationship:** RA3 interpretation; no new calorie-label score component.

A28. Tian ritual/self-control retraction: precise scope

Book locator: chapter14.xhtml, ritual/self-control discussion; notes-ch14.xhtml, note 17, Tian et al. (2018), DOI 10.1037/pspa0000113.

Book position: Chapter 14 ¶126, note 17, cites Tian et al. for a definition of ritual as repeated, fixed actions. Wood does not recount the calorie-intake experiment affected by the notice. Its role is more peripheral than the Brooks singing experiment.

Source/access map:

- B — full text sections read; reviewed book
- N02 — complete notice text read through PubMed/XML; later evidence
- U10 — abstract only; book source; retraction N02

The journal notice is recorded in the February 2024 issue; its detailed account is available in the PubMed abstract. Following the authors' 2023 review, 15 participants' condition codes differed between the analyzed file and participants' reported actions in Qualtrics. Reanalysis no longer supported Study 1's calorie-intake benefit (reported $p=.576$). The authors agreed to retraction. The notice reports no identified discrepancies in the datafiles for Studies 2–16; that is not independent certification of those studies.

The exact first-online day was not verified and is not invented. The original paper's full dataset was not analyzed here. This changes the cited source's current status but does not establish that its ritual definition is false, or that Wood misdescribed the calorie experiment. It is not proof of her knowledge in 2019 or a finding by this review about who caused the discrepancies. Nor does a nonsignificant reanalysis prove a precisely zero effect or an effect in the opposite direction.

Judgment confidence: high for notice account; exact day unresolved. **Scoring relationship:** current evidential support/RA discussion; not another SA proposition.

A29. Brooks ritual/performance retraction: verified notice and bounded reasons

Book locator: chapter14.xhtml, ¶150 singing example; notes 22 and 24, DOI 10.1016/j.obhdp.2016.07.004.

Book position: Wood describes reduced anxiety and improved singing performance following a draw/salt/tear/count ritual.

Source/access map:

- B — full book context and notes read
- N03 — full publisher-indexed notice text read during crosscheck; later evidence
- U11 — relevant original experiments read in author PDF during chapter audit

The original paper reported the broad findings Wood describes. The entire article was later retracted. The publisher notice, in the November 2024 issue, reports that the authors initiated an audit and no longer trusted the data's integrity. The notice identifies unavailable original data for several studies and concerns particularly involving Study 1b. The combined audit recovered the full publisher-indexed notice text after v1's unsuccessful direct retrieval; it did not obtain an independently downloaded notice PDF.

The affected article should no longer serve as established support for the singing/performance claim. This is not a finding that Wood misquoted the original reported result, knew of later concerns, or caused any data problem. The exact first-online day

remains unresolved. No inference about individual responsibility, unrelated papers, or all ritual research follows from the notice.

Judgment confidence: high for status and publisher-stated general grounds. **Scoring relationship:** current evidential support; no speculative or retroactive misconduct penalty.

A30. A surviving ritual experiment —and Wood’s accurate qualification

Book locator: chapter14.xhtml, EEG/error-monitoring discussion, “nor did it improve”; notes-ch14.xhtml, note 19.

Book position: Wood reports that a ritual changed an error-related brain response without improving task performance.

Source/access map:

- B — full text sections read; reviewed book
- S24 — full text sections read; book source: Chapter 14, note 19

Hobson, Bonk and Inzlicht used a preregistered pre/post design with an arbitrary ritual practiced over a week and a laboratory performance task during EEG recording. The reported result concerned attenuation of the error-related negativity; the authors did not demonstrate a behavioral performance improvement. Wood explicitly preserves that distinction. The self-report measures also did not show improved affect/motivation, and their changes were not correlated with the neural response. Her subsequent use of selected participant comments to suggest calmed fears therefore exceeds the measured emotional result.

This is favorable evidence that the ritual discussion is not wholly dependent on the retracted papers and that Wood can report a limitation accurately. A neural response associated with errors is not identical to distress, and neither is a direct measure of better real-world performance. The failure to detect improved self-reported emotion is not proof of exactly zero emotional effect. The study cannot replace the withdrawn evidence for every ritual benefit. It also does not prove a general clinical anxiety treatment. The review therefore retains a bounded favorable finding rather than either discarding the chapter or using neuroscience language to upgrade the outcome.

Judgment confidence: moderate. **Scoring relationship:** RA2 credit; limits global negative verdict about rituals.

A31. Repeated exposure and vegetable acceptance

Book locator: chapter14.xhtml, opening vegetable study; notes-ch14.xhtml, note 1.

Book position: Wood uses repeated exposure to unfamiliar vegetables as evidence that repetition can increase acceptance and reward.

Source/access map:

- B — full text sections read; reviewed book
- S25 — full text sections read; book source: Chapter 14, note 1

Caton and colleagues studied 72 children aged 9–38 months assigned to repeated exposure, flavor–flavor learning or flavor–nutrient learning conditions. The primary report describes increased intake of a novel vegetable across exposure conditions, with follow-up. The relevant result supports repeated exposure as a route to acceptance without requiring an extra flavor or energy pairing.

The experiment measures consumption and acceptance, not persistence after outcome devaluation. It therefore supports a component of Wood’s practical story—an action or food may become more acceptable with experience—without establishing an automatic habit. The chapter audit reconciled ten training exposures plus pre/post and three follow-up assessments, explaining approximately fifteen total exposures. Follow-up included 45 children over several weeks. There is no established ten-versus-fifteen error to add. This is a favorable chapter finding with a clear construct boundary.

Judgment confidence: moderate. **Scoring relationship:** C2 reward component; all-chapter favorable coverage.

A32. Meaning, fluency, practice and skill are not one mechanism

Book locator: chapter 7 closing excellence/practice discussion, note 16; chapter 14 routines/meaning discussion, note 27.

Book position: Wood links familiar activity with comfort, meaning and flow while distinguishing repetition from excellence. Family-meal associations may reflect stable homes rather than cause their benefits; she acknowledges that alternative.

Source/access map:

- B — full relevant chapters and notes read
- S36 — indexed primary methods/results read; routines and meaning
- S44 — primary abstract; deliberate–practice meta-analysis
- N06 — complete formal 2018 correction read

Heintzelman and King related routine to meaning in 317 participants and, in a second study, 85 participants supplying 2,590 episodes. Within-person associations survived affect adjustment; one between-person association weakened. These observations do not establish that imposing a routine causes a meaningful life. Norton–Gino’s grief paper is a separate book reference (U12), not the source for the routine-and-meaning findings.

Macnamara et al.’s correction repaired a dependent-sample adjustment and changed the overall variance estimate from 12% to 14%, with domain estimates also altered. The authors reported no substantive conclusion change. This preserves the distinction between repetition and excellence. Unexplained variance cannot simply be attributed causally to native talent.

Liking, fluency, meaning, skill and learned automatic response are related but different outcomes. They should not be counted as interchangeable proof of a single habit mechanism. Equally, none is invalidated merely because it lacks a devaluation assay. Wood's express qualification about excellence is important favorable evidence.

Judgment confidence: high for book attribution; limited by source-specific access for empirical details. **Scoring relationship:** C1/C2 fair reading and construct boundaries.

A33. Policy, defaults and energy feedback: benefit is not mechanism

Book locator: chapter15.xhtml, organ donation, saving and electricity; notes-ch15.xhtml, notes 2–3 and 24.

Book position: Wood broadens the framework to social settings, defaults, infrastructure and feedback rather than assigning all responsibility to individuals.

Source/access map:

- B — full text sections read; reviewed book
- S26 — abstract only; book source: Chapter 15

The Jessoe–Rapson primary abstract reports a randomized trial of high-frequency electricity-use information and temporary price changes. Informed households responded more strongly to prices, with conservation also observed beyond the pricing events. The authors discuss learning and possible habit formation; the abstract does not provide a direct automaticity assay. This review did not independently audit all measures and follow-up analyses.

A default that enrolls someone or executes a payment is a system arrangement, not necessarily a newly learned psychological habit. Similarly, a policy can change prices, opportunities or incentives while behavior remains deliberate. These distinctions do not subtract value from a successful policy. They prevent Chapter 15's diverse outcomes from being counted as repeated independent validation of one mental mechanism. Organ-donation and built-environment source relationships are recorded, but their numerical causal effects are not independently certified in this review.

Judgment confidence: moderate. **Scoring relationship:** C2 scope; PV1 favorable attention to external conditions.

A34. 2026 deadline replication, retraction and author responses

Book locator: chapter1.xhtml, precommitment discussion; notes-ch1.xhtml, note 1, Ariely & Wertenbroch (2002).

Book position: The book cites the 2002 precommitment paper. Its later evidential status must be disclosed without implying foreknowledge by the book's author.

Source/access map:

- B — full text sections read; reviewed book
- S28 — metadata only; book source: Chapter 1, note 1
- S29 — full text sections read; reader-added replication; later evidence
- N01 — full text sections read; notice affecting book source S28; later evidence
- R01 — full text sections read; published author response; later evidence
- R02 — full text sections read; published author response; later evidence

The official notice, first published 2 September 2026, says the editor can no longer attest to the original findings' reliability after questions concerning underlying datasets and an investigation. Both authors agreed to retraction. This is the journal's reliability judgment, not this review's independent attribution of misconduct.

Hyndman and Bisin's replication, published 15 July 2026, reports no reproduction of the original deadline-related performance advantages while retaining evidence of demand for costly self-imposed deadlines. Its main study had 124 participants; procedures and materials differed in identified ways. The full supplement was not read. Ariely's 7 August statement acknowledges anomalies and an insufficient surviving record to resolve questions. Wertenbroch's published response says he lacked the published studies' data and requested retraction, while separating demand for commitment from its efficacy. These are attributed author accounts, not findings independently investigated here.

The later developments weaken the affected illustration; they do not refute all commitment devices or habits. The review does not turn a demand result into proof of improved performance. Nor does it infer that Wood could have known the 2026 outcome in 2019.

Judgment confidence: high for notice/date; moderate for replication scope. **Scoring relationship:** later evidence; no retroactive claim of negligence or fourth SA grade.

A35. Figures, labels and an internal portion-size mismatch

Book locator: All in-chapter images; chapter15.xhtml, paragraph preceding Illus14.jpg.

Book position: The EPUB includes 15 in-chapter illustrations. Chapter 15 describes burgers as having quadrupled while the associated illustration labels 3.9 oz and 12 oz.

Source/access map:

- B — full text sections read; reviewed book

The 15 illustrations were inspected visually, including the brain/task diagrams, Stroop material, contextual illustrations, phishing example, car-design display, historical alcohol-availability map, portion graphic and electricity display. Their content was read alongside the

surrounding prose and credits; this is not independent validation of all original figure datasets.

In the portion graphic, 12 divided by 3.9 is approximately 3.08, not four. The accompanying burger sentence therefore overstates the ratio shown by its own illustration. This is a small internal descriptive error, separate from the much larger kilojoule/calorie error in A25. The map and historical device examples are treated as historical book material, not current 2026 descriptions. Decorative or illustrative content is not counted as a controlled scientific experiment.

Judgment confidence: high. **Scoring relationship:** minor RA2 finding; no freestanding additional score subtraction.

A36. Phone appendix: concrete advice with unmeasured program-level effects

Book locator: howtostop.xhtml, complete; notes-bm2.xhtml; copyright.xhtml health disclaimer.

Book position: The phone appendix recommends manipulating access, notifications, friction, replacements and rewards, including stronger restrictions for persistent checking.

Source/access map:

- B — full text sections read; reviewed book
- U08 — primary abstract read; no full methods or complete phone-program audit; book source

The advice is concrete enough to test in daily use. But no controlled evaluation of the complete appendix prescription was identified in the scoped search. A laboratory notification study in the notes cannot by itself establish that the full combination improves well-being, productivity or long-term phone use. The primary abstract was read during the chapter audit; its full methods were not independently examined.

Burden varies by application. Reducing a notification is different from making oneself unavailable when work, caregiving or safety depends on contact. The book allows dedicated callers through its Do Not Disturb option, but also suggests turning the phone off or leaving it behind; the review therefore does not rate all its suggestions as burden-free. This is an editorial appraisal of plausible trade-offs, not evidence of a measured adverse-event rate. The review does not silently replace Wood's instructions with a safer custom protocol and then award the book credit for that improved version.

Judgment confidence: high for advice; moderate for benefit-burden judgment. **Scoring relationship:** PV1/PV2/PV3; ease cannot raise score.

A37. Whole program, components and durable value

Book locator: chapter6.xhtml–chapter10.xhtml; chapter10.xhtml, “passive returns for all time”; howtostop.xhtml.

Book position: Wood offers an integrated practical framework rather than a standardized, dose-specified clinical treatment manual. The prose sometimes implies that initial investment can remove continuing attention indefinitely.

Source/access map:

- S30 — revision evidence; exact access and relationship in source record
- S31 — revision evidence; exact access and relationship in source record
- B — full text sections read; reviewed book
- S03 — full text sections read; book source: Chapter 7
- S09 — full text sections read; reader-added intervention; later evidence
- S12 — full text sections read; reader-added intervention
- S13 — abstract only; reader-added intervention
- S14 — full text sections read; reader-added intervention; later evidence
- S21 — abstract only; book source: Chapter 13, note 37

No controlled trial of reading this exact book or implementing its entire cross-domain program was identified in the scoped search. That is a search result with limits, not proof that none exists. The evidence includes longitudinal automaticity reports, interventions using overlapping habit principles, controlled outcomes for particular behaviors, and examples of benefits that require continued support.

The absence of a book trial is not an automatic zero: useful advice can have convergent component support. Conversely, a collection of positive components does not establish the added value, mechanism or durability of the integrated package. This distinction supports a qualified positive benefit grade and a lower applicability/durability grade. A maintained environment is allowed to count as success; the review does not require cues, support or monitoring to vanish. It asks whether the actual package is feasible and beneficial, and whether stronger promises of unattended permanence have adequate evidence.

Judgment confidence: moderate. **Scoring relationship:** C2/PV1/PV2; prevents rewarding reviewer-improved advice.

Revision evidence: Controlled physical-activity trials improve reported habit strength on average, and this favorable evidence is included in the later-audit section. It does not equate reported automaticity with effort-free whole performances. The book also says context changes can disrupt good habits: its permanence language should be evaluated with that condition, not as a guarantee that survives every life change. A trial of reading the exact book is not required to credit useful, faithfully described components.

A38. Later favorable evidence: a hybrid model of social-media posting

Book locator: howtostop.xhtml and chapter8.xhtml; reviewer-added 2026 comparison.

Book position: Wood's framework predicts roles for repetition, context and reward; later work can support mixed mechanisms without proving her practical program.

Source/access map:

- B — full text sections read; reviewed book
- S27 — full text sections read; reader-added favorable mixed-process evidence; later evidence

Turner and colleagues' paper was published 4 June 2026. It modeled posting from 2,696 Twitter users, divided into 1,138 discovery and 1,558 confirmatory users. Hybrid reward-learning/perseveration models outperformed the alternatives examined. The confirmation was preregistered, but it used another portion of the same underlying dataset; the model-falsification procedure was described as exploratory.

The result is favorable to considering both past behavior and reward sensitivity. Its "habit" component operationalizes perseveration/autocorrelation rather than a devaluation experiment. Model comparison is relative to the candidates considered and does not alone identify a unique causal mental process. This later finding prevents the evidence-over-time section from becoming only a list of failures, while keeping its methodological limits explicit. This review did not reproduce the model fits, inspect every supplement, or infer that the paper validates Wood's phone-use instructions.

Judgment confidence: moderate. **Scoring relationship:** C1 favorable mixed-process evidence; clearly later.

A39. Traceable but unexamined or inaccessible sources

Book locator: notes-ch5.xhtml; notes-ch8.xhtml; notes-ch9.xhtml; notes-ch10.xhtml, note 14; notes-ch13.xhtml, note 40; notes-bm2.xhtml.

Book position: The book uses published papers, syntheses, interviews, media accounts and unpublished manuscripts/data. These have different evidential roles and different access status.

Source/access map:

- B — full text sections read; reviewed book
- U01 — full relevant primary sections read during revision; initial metadata access superseded
- U02 — unavailable; book source, not primary-verified here
- U03 — unavailable; book source, not primary-verified here

- U04 — full relevant primary sections read in chapter audit; initial metadata access superseded
- U05 — metadata only; book source, not primary-verified here
- U06 — primary accepted-paper abstract and authors' account read; no full methods/model-estimate audit; book source
- U07 — metadata only; book source, not primary-verified here
- U08 — primary abstract read; no full methods or complete phone-program audit; book source

Items still not independently method-audited include Judah's flossing work, the London disruption study, several policy sources and the notification study. The London study's primary abstract and authors' account, and the notification study's primary abstract, were read during the chapter audit; those sources are not metadata-only. Revision audits obtained relevant primary sections for Shindou's cellular timing study and Lin, Wood and Monterosso's eating-task study; their earlier access gaps are not current gaps. The book's "Overthinking Habit" manuscript under revision (Labrecque, Lee and Wood, 2018) and "Habits of Successful Rehabilitation" unpublished data (Monterosso and Wood, 2017) were not obtained. They are not counted as verified primary experiments here.

A traceable citation is not the same as full access. Equally, an unavailable manuscript is not a fabricated study. The retrieved Quinn-Wood unpublished manuscript is treated differently because relevant text was actually available. Interviews and private datasets are used only as book-reported illustrations unless independently corroborated. The selected-source audit therefore cannot supply an error percentage for the full bibliography. These limits are carried into the confidence of the editorial ratings rather than coded as zero scientific evidence or automatically as reference errors.

Judgment confidence: high for access record. **Scoring relationship:** RA1 scope limit; no automatic negative grade for missing access.

A40. Methodology, voice and editorial provenance

Book locator: User commission; requested methodology URL; supplied EPUB; public voice articles.

Book position: This is an independent commissioned adaptation of Red Pen Reviews using the user-specified nine-input rubric, not an official Red Pen Reviews review.

Source/access map:

- B — full text sections read; reviewed book
- P01 — supplied house methodology text read; editorial method, not independent scientific evidence
- P02 — full text sections read; voice/position calibration only
- P03 — full text sections read; voice/position calibration only

- P04 — full text sections read; voice/position calibration only

The supplied house methodology was read and its operational anchors applied. Source-access upgrades and corrections are recorded; repeated AI analysis is not described as independent human review. The central propositions are retained, but their selection before adverse searching was not verified.

The public articles and relevant guide sections informed prose and the person-behavior-context, motivation and practice distinctions. They are not independent evidence against Wood. No Jason anecdote, personal scientific endorsement, author correspondence, undisclosed relationship, public posting date or publisher approval is invented. The output is AI-assisted, with no documented independent human scientific sign-off and no original-data replication. The completed checks concern document structure, source access and arithmetic; they do not certify every scientific assertion.

Judgment confidence: high. **Scoring relationship:** provenance and verification boundary.

Evidence added or strengthened during revision

This supplement preserves the forty original dossiers and supplies the later/current evidence needed to assess them fairly. It is not a fourth scientific claim. Dates refer to source publication; retrieval occurred by 15 September 2026.

Controlled automaticity evidence. Ma et al. (2023) synthesized ten randomized physical-activity trials with 2,349 participants. Reported habit strength improved relative to controls, SMD 0.31, 95% CI 0.14–0.48, $I^2=64\%$. The outcome was SRHI/SRBAI self-reported habit strength/automaticity, not a 31% activity increase, clinical improvement or an effortless workout. Longer-follow-up studies had smaller pooled effects, but between-study comparisons do not establish each person's rate of decline. Component associations were exploratory. This is meaningful favorable evidence for a narrower outcome. S30.

Formation times remain variable. Singh et al. (2024) included twenty studies and 2,601 participants, but only four reported formation times and eleven were rated high risk of bias under its PEDro rule. Its pooled SMD 0.69 was pre-to-post change, not a controlled treatment effect. Different thresholds and modeled estimates do not yield a universal deadline. This supports Wood's own rejection of a single correct number. S31.

Prepared habits and deliberate control. Hardwick et al. (2019) reported old practiced responses under short preparation times after remapping, with current goals able to guide later responses. The inspected abstract supports distinguishing preparation from execution; the final paper appeared close to book production, so no omission-negligence claim follows. S35. Wood and colleagues' later theory also allows interacting goals and habits; it is an author position and secondary analysis, not a response solicited by this review. A08.

Favorable 2026 mechanistic evidence. Michiels et al. used three-day keypress/reward training and a 600-ms response deadline. More practiced responses showed greater switching cost relative to training, with convergent sensorimotor/control-network findings. Reported final samples were 28 for fMRI and 25 for active/sham TMS. Regional analysis supported posterior-putamen activity, but neural-behavior correlations did not survive

multiplicity correction; model parameters were not uniquely recoverable. Stimulation improved accuracy generally, without a selective habit interaction or significant reaction-time effect. This supports learned responses competing with goals, not a selective causal localization or broad everyday-effect estimate. The recruitment/exclusion arithmetic did not reconcile the reported fMRI N. Extensive indexed original sections, not a complete downloaded PDF/supplement audit, underlie this account. S33.

A 2026 measurement challenge. Vázquez-Millán et al. randomized one versus three days of monetary keypress training. Of 123 completers, 107 met preregistered inclusion criteria. Both groups adjusted after instructed devaluation; the training-by-devaluation interaction was nonsignificant ($p=.166$). Exploratory analyses related residual responding to less effective devaluation and reanalyzed older data; clustering did not match the preregistered single-cluster prediction. This conceptual replication permits deliberate control to suppress a learned tendency; it cannot establish whether habits formed or whether every persistent response is goal-directed. S34.

These 2026 papers share a coauthor and are not wholly independent research teams. Their results can coexist: changing a practiced response can carry a cost while persistence alone is an ambiguous habit measure. Neither determines the fraction of daily life governed by habit, lasting health benefits, or treatment success.

Pre-book treatment correction. The 1999 cocaine-treatment trial is a newly audited source, not later evidence: all groups encouraged twelve-step participation, and individual drug counseling plus group counseling produced the strongest drug-use outcomes. It does not support Wood's broad equivalence comparison. The three individual-plus-group conditions had the same scheduled frequency, so their contrast is not explained by calling one simply the "most intensive." Details and fair limits are in A20, S32.

Other bounded chapter additions. The toy-soap pilot randomized 203 households. Favorable parent reports of soap use coexisted with nonsignificant observed unprompted-washing comparisons; 42% reportedly obtained toys by other means such as cutting/dissolving soap. It is promising mixed pilot evidence rather than durable habit demonstration. S43. The published stacking trial and retailer-exposure evidence are separately explained in A11. The unadjudicated 2025 contingency-degradation paper is not used as a settled additional finding.

7 Practical use

What is actually being rated

The rated advice is Wood's: choose goal-consistent recurring behavior, arrange cues and opportunities, change friction, make repetition rewarding, preserve useful consistency, and use disruptions or replacements where appropriate. The scoring does not incorporate a new diagnostic protocol invented by the reviewer. The reader essay's hypothetical alternatives illustrate a boundary of the advice; they do not earn the book additional evidence credit. Book locators: `chapter6.xhtml–chapter10.xhtml`, `chapter11.xhtml`, `howtostop.xhtml`; A11, A17, A36.

For simple repeatable actions, the framework has a defensible practical foundation. Related controlled interventions produce useful outcomes, although some combine habit advice with monitoring, information, rewards or professional support. The intervention evidence is strongest for the concrete populations, settings and outcomes tested, not every domain named in the book. A14, A15, A16, A23.

Durability and continuing support

A maintained behavior can be worthwhile even if stable cues, ongoing support or reinforcement remain necessary. This review does not require cue removal, withdrawal of effective support, or support-free permanence as a universal criterion for benefit. It does require transparency about whether an effect was observed while an intervention continued, after it ended, or only among follow-up completers. Comparative advantage can disappear while both groups improve. A14, A15, A24, A37.

The same reasoning applies to automaticity. A repeated start can support a demanding activity; the whole performance need not become effortless. Evidence that a person attends a gym more often should not be translated into evidence that exercise no longer requires effort. Conversely, lack of a devaluation assay does not invalidate a real attendance benefit. A05, A16, A32.

Applicability, burdens and risks

Some actual suggestions are inexpensive and reversible. Others require more control over location, schedules, family cooperation, social relationships or communication access. The practical-value assessment includes those demands rather than rating only the easiest examples. These are reasoned implementation judgments, not measured cost-effectiveness or adverse-event estimates. A11, A17, A36.

The addiction material is a high-stakes boundary. The book's copyright matter also advises physician consultation before certain diet/exercise changes, especially with existing health problems. Its multifactorial and complementary-care qualifications are preserved, but the numerical benchmark should not be used to judge an individual's recovery or the effectiveness of a different treatment population. The review makes no individualized medical recommendation and does not treat this self-help book as a clinical protocol. A20–A24.

Ease of application — described, not scored

The instructions are generally understandable and often concrete; a reader can identify a cue or obstacle without specialized equipment. That is editorial assessment of the text. Actual implementation ranges from a small rearrangement to substantial lifestyle or social change. Ease does not raise Practical Value, compensate for weak benefit evidence, or turn an untested program into an effective one. The three practical grades remain benefit, applicability/durability, and benefit relative to burden/risk.

8 Evidence over time

The timeline is selective and claim-linked, not a history of all habit research. “Later” means later than the 2019 book. Issue dates and exact online dates are not silently interchanged.

Period	Evidence	Interpretation in this review
1977 report / 2010 reprint	Original Vietnam cohort report	Real low renewed addiction; treatment-selection cautions verified in the reprint. The 1993 paper’s v1 full-access claim was not independently reproduced. A20, S16.
1996	Petrie’s enrichment comparison	Pre-book contrary animal result with population/measurement limits. A21.
2002 / 2005	Diary studies, including unpublished lifespan manuscript	Repeated-context behavior; not a universal unconscious-action fraction. A03.
2009–2011	Formation-time, popcorn and therapeutic-workplace studies	Positive but different constructs and outcomes; support-maintenance limitation preserved. A06, A09, A24.
2013–2017	Incentives, network support, brief weight intervention, enrichment, neural ritual study	Favorable evidence with comparator/mechanism limits; not all positive evidence disappears. A13, A14, A22, A23, A30.
June/July 2017	Buffet corrigendum	Predates book and explicitly acknowledged in its notes. Exact online day provenance described in A26/N04.
2018	de Wit habit-induction failures	Pre-book limits on specified laboratory procedures; not universal disproof. A07.
2021–2022 — later	Cue planning, flexible exercise, author response and multilaboratory induction work	Both favorable and constraining evidence; distinguish initiation, attendance and devaluation. A07, A08, A10, A16.
2021 — later	Preregistered depletion project	Small/nonsignificant confirmatory result; not a direct stress–habit test. A19.
20 February 2023 — later	Preregistered stress replications	Critical control behavior failed; no clean global stress-null conclusion. A18.
February 2024 issue — later	Tian retraction	Detailed notice account read through PubMed; exact online day unresolved. A28.
November 2024 issue — later	Brooks retraction	Full publisher-indexed notice read during revision; exact online day unresolved. A29.
4 June 2026 — later	Social-media computational model	Favorable mixed-process account, with same-dataset confirmation and model limits. A38.
15 July 2026 — later	Hyndman–Bisin deadline replication	Demand and performance results separated. A34.
7 August 2026 — later	Ariely public statement	Author account included without independent provenance adjudication. R01.
2 September 2026 — later	Official Ariely–Wertenbroch retraction	Complete journal notice and exact date read; both authors agreed. A34.
1999	Twelve-step counseling trial	Pre-book source newly checked; all groups encouraged participation. A20, S32.
7 June 2021 — later	Beeken unit correction	Baseline lipid/glucose units corrected; weight results unchanged. N05.

Period	Evidence	Interpretation in this review
2023–2024 — later	Controlled automaticity and habit-time syntheses	Positive reported automaticity; timing variation and design limits. Later audit.
January/July 2026 — later	Neural switching and devaluation measurement studies	Qualified favorable evidence and measurement cautions coexist. Later audit.
15 September 2026	Review cutoff	No claim about findings published or notices issued after this date.

Later evidence changes current support without rewriting what could have been known in 2019. The review does not assign a separate numerical “2019-only” score because that would require its own consistently time-restricted evidence judgment.

9 Scoring rationale

Scale and calculation

The house method evaluates scientific propositions at their stated breadth, reference relationships at the examined scope, and practical benefit separately from ease of use. Anchor 2 means meaningful narrower/component support with important gaps; anchor 3 requires substantial support at the proposition's scope. Access failures are not traceability failures. Practical benefits can be demonstrated without proving a unique mechanism or testing the act of reading the book.

Each category is $100 \times (\text{sum of its three inputs}) / 12$. Overall is $100 \times (\text{sum of all nine inputs}) / 36$. No intermediate grade or total is rounded; displayed percentages use whole numbers. An NR input would withhold its category and the overall rather than contribute a zero.

Input	Grade
C1 persistence mechanism	2
C2 deliberate habit construction	2
C3 context and addiction recovery	2
Reference traceability	4
Reference description/completeness	2
Reference inference	2
Intended practical benefit	3
Applicability/durability	2
Benefit versus burden/risk	3

Exact ratios: Scientific 6/12; Reference 8/12; Practical 8/12; Overall 22/36. Display: **50 / 67 / 67; overall 61**. These are editorial ratings, not percentages of true statements or probabilities of success.

SA1. C1: habit as a principal explanation of persistence — 2/4

Confidence: moderate.

Habits have meaningful support as one source of persistent action and reduced deliberation. Experiments show learned response tendencies, and observational studies link beneficial habits with self-control and positive outcomes. But these findings do not establish how much of everyday persistence or the self-control advantage is caused by habits rather than goals, preferences, opportunities or other differences between people. The selected claim is partly supported at its stated breadth. Recent favorable preparation/switching evidence is included; it does not provide that population-level causal bridge.

Adjacent lower grade: 1 would undercredit the cue/response experiments, converging observations and meaningful later mechanistic evidence.

Adjacent higher grade: 3 would require appropriate evidence for much of the claimed relative explanatory importance, not simply further demonstrations that habitual tendencies exist.

Evidence: A03, A04, A05, A06, A07, A08, A38, later evidence.

SA2. C2: deliberate construction of durable, lower-effort behavior — 2/4

Confidence: moderate.

Longitudinal studies and randomized interventions show that selected repeatable health behaviors can become more automatic or improve under habit-oriented support. This is meaningful evidence for parts of the recommendation. It remains narrower than the cross-domain promise, and attendance or weight change does not itself establish a durable reduction in reliance on effortful decisions. Important measurement, transfer and durability gaps fit the method's partly-supported anchor. A trial of reading the exact book is not required; the existing component evidence is credited.

Adjacent lower grade: 1 would undercredit the controlled automaticity evidence and actual useful intervention outcomes.

Adjacent higher grade: 3 would require support for much of the stated cross-domain, lower-effort proposition. Grading a narrower useful framework while treating its stated generality as optional would change the claim being scored.

Evidence: A09, A10, A11, A12, A13, A14, A15, A16, A37, later evidence.

SA3. C3: context redesign and addiction recovery — 2/4

Confidence: moderate to high.

Controlled animal and human intervention evidence supports useful roles for environment, alternative reinforcement and abstinence-supportive networks. Wood's multifactorial caveats and maintenance qualifications are substantial. The returning-veterans cohort does not identify cue disruption as the dominant cause, and its first-year renewed-addiction outcome is not a transferable benchmark for heterogeneous treatment relapse. That consequential extension remains partly supported at the selected scope.

Adjacent lower grade: 1 would fit a narrow unsupported numerical comparison better than the whole selected proposition; it would undercredit favorable treatment evidence.

Adjacent higher grade: 3 would require a defensible bridge for the benchmark and causal extension, or a transparently different narrower claim. The delivered proposition is retained.

Evidence: A20, A21, A22, A23, A24.

RA1. Traceability — 4/4

Confidence: moderate within the examined scope.

Important examined claims lead to identifiable sources through the book's notes, including papers whose limitations or later corrections weaken their use. No consequential source-identity failure was established after the chapter audits. Named private data, interviews and unfinished manuscripts can be identifiable even when their evidence is unavailable. Access is therefore not used as a substitute traceability penalty. This is the highest anchor at the examined scope, not certification of every bibliography item or its data.

Adjacent lower grade: 3 would require a concrete limitation in tracing an important examined claim to its cited work. The prior rationale identified unavailable material rather than such a defect.

Adjacent higher grade: 4 is the scale ceiling. A broader bibliography audit could change confidence or reveal new defects; lack of that audit is not itself an observed traceability error.

Evidence: A01, A26, A28, A29, A34, A39.

RA2. Accurate description and material completeness — 2/4

Confidence: moderate to high.

The restaurant example confuses kilojoules with calories; the incentive account omits the successful group condition needed to assess its lesson; and the burger ratio overstates its own figure. The twelve-step comparison also needs the actual treatment design. Against these problems, Wood preserves formation-time variability, several experimental qualifications and the corrected buffet numbers. Later withdrawals update source status without demonstrating a 2019 misquotation or negligence.

Adjacent lower grade: 1 would overstate the breadth of confirmed misdescription; several findings and qualifications are faithfully reported.

Adjacent higher grade: 3 is plausible if the confirmed consequential errors are weighted as localized. The current judgment gives their scale and omitted comparisons substantial weight.

Evidence: A09, A13, A20, A25, A26, A30, A35.

RA3. Appropriate inference — 2/4

Confidence: moderate to high.

Meaningful narrower findings are repeatedly extended into stronger explanations: correlational mediation becomes a broad self-control mechanism, retailer exposure becomes a causal purchase/relapse sequence, cellular timing becomes a human reward rule, and unlike recovery endpoints become a treatment benchmark. The ritual EEG result is also extended beyond its unchanged subjective-affect measure. Mixed-control qualifications and useful component evidence prevent a uniformly negative assessment.

Adjacent lower grade: 1 would require predominantly unsupported reasoning across the examined relationships; the favorable evidence and explicit caveats do not support that judgment.

Adjacent higher grade: 3 would require more restrained conclusions at the major translation points, especially self-control, reward timing, addiction comparisons and subjective ritual benefit.

Evidence: A03, A04, A11, A12, A13, A20, A26, A30, A37.

PV1. Evidence of intended benefit — 3/4

Confidence: moderate.

Several consequential recommendations have appropriate evidence of useful benefits, including related habit-oriented weight interventions and social/reinforcement support. The method credits faithfully described useful components without requiring a trial of reading the book. A measured practical benefit need not depend on identifying habit as its unique mechanism. The grade does not validate an invented matching protocol, the whole phone appendix, or equal effectiveness across all named domains.

Adjacent lower grade: 2 would be more appropriate if the demonstrated benefits transferred less well to important everyday applications than this assessment allows. The question is relevance and breadth, not mandatory testing of the exact whole book.

Adjacent higher grade: 4 would require stronger, convergent and directly relevant evidence for the important benefits at the stated scope, with clearer comparative outcomes and follow-up. Proof of a unique mental mechanism is not an additional practical-benefit requirement.

Evidence: A11, A14, A15, A16, A23, A24, A31, A36, A37.

PV2. Applicability and durability — 2/4

Confidence: moderate.

Useful applications exist, but transfer, maintenance and resource assumptions remain incomplete. Stable opportunities and control over one's setting cannot be assumed. Automatic initiation differs from skilled performance, and maintained within-group improvement differs from a sustained comparative advantage. The exercise scheduling advantage attenuated later in follow-up. The book acknowledges disrupted good habits and some continuing-support needs; effective support is not disqualified because it remains necessary.

Adjacent lower grade: 1 would undercredit maintained outcomes, usable components and the book's attention to adaptation and support.

Adjacent higher grade: 3 would need stronger evidence that the advice is matched to its intended settings and horizons, with clearer limits across the claimed populations and domains.

Evidence: A09, A14, A15, A16, A17, A24, A36, A37.

PV3. Benefits relative to burdens and risks — 3/4

Confidence: moderate; no measured adverse-event synthesis.

Expected benefits are reasonably proportionate to many everyday recommendations, and the book supplies meaningful safeguards: physician-consultation advice for some diet/exercise changes, complementary rather than exclusive addiction explanations, recognition of continuing support and some phone-access exceptions. Other recommendations require substantial control over settings or relationships. The recovery benchmark can encourage unrealistic judgments. This is an editorial assessment of actual advice, not a measured harm rate or proof of zero risk.

Adjacent lower grade: 2 is plausible if demanding context changes and the high-stakes recovery benchmark receive greater weight than the common lower-burden applications and explicit safeguards.

Adjacent higher grade: 4 would require stronger evidence of net value and better integrated boundaries for heterogeneous readers and major foreseeable trade-offs.

Evidence: A02, A11, A20, A24, A36, A37.

Avoiding unexplained double deductions

These nine inputs are holistic judgments, not a ledger subtracting a point per error. The pasta units and burger ratio chiefly inform RA2. The incentive example contains a source-completeness problem and a distinct unmeasured-mechanism inference; neither shows that

incentives never work. The Vietnam benchmark can inform scientific support, inferential use and high-stakes practical burden because those are different questions. No statistical independence of the ratings is claimed.

Retractions update current support without an automatic per-paper penalty or a retroactive misconduct inference. Traceability is graded separately from ability to inspect a named source. The review's own mapping errors were repaired rather than charged to Wood. Continued support is not a defect by itself; its expected benefit and demands matter.

Revision and sensitivity

Scientific Accuracy is 2/2/2 because the stated propositions receive narrower or component support with important scope gaps. Reference Accuracy is 4/2/2: no consequential identity-traceability defect was established, while use and interpretation retain the limitations described above. Practical Value is 3/2/3. These criterion-based inputs total 22/36, displayed as 61%; the grades do not target an overall score.

One input point changes its category by 100/12 percentage points and the overall by 100/36. Raising only SA2 from 2 to 3 would produce Scientific 7/12 (58%) and overall 23/36 (64%). Lowering only RA1 from 4 to 3 would produce Reference 7/12 (58%) and overall 21/36 (58%). These are adjacent editorial judgments, not confidence intervals.

10 Verification and limits

Completed checks

The reading record covers book identity, whole-spine coverage, note reading and 15 in-chapter image inspection. Chapter auditors subsequently read their complete assigned text and notes. The delivered selection file remains unchanged, but its claimed pre-adverse-search chronology is not verified. The review distinguishes actual book citations from reader-added sources, original cohorts from later reprints, absolute quantities from relative effects, and active versus untreated comparators. Primary text was checked for the main descriptive error, the missing incentive arm, the Vietnam inference, and the accessible notice records. The appendix contains forty dossiers, supporting evidence and source records, with checks for duplicate anchors and unresolved internal links. This structural check does not certify every scientific assertion or the rendering of separate artifacts.

Limits that remain material

This is not a systematic review with an exhaustive search corpus, not independent verification of every bibliography entry, and not a full risk-of-bias assessment of every included trial. Some clinical/intervention summaries are abstract-only. The full publisher-indexed Brooks notice was obtained during revision; its exact online date and Tian's exact online day remain unresolved. The buffet correction's retrieved exact-day metadata is not confused with a separately confirmed publisher timestamp. The supplied house methodology was read during preparation. Initial and revised access are distinguished rather than silently conflated.

The selected full-text label always names the sections actually read. Full-text availability somewhere online is not equivalent to this review having inspected it. No access is claimed to unpublished source data that were not retrieved, and no source is declared fabricated because retrieval failed. Author statements are attributed statements, not independent findings about responsibility.

No one has been contacted for an author response to this review. Published responses within the scientific disputes are separately described. There is no documented independent human scientific sign-off, no publisher approval and no original-data replication. AI-assisted source and editorial checks are documented with their remaining limits in this section.

11 References

This is the **review's examined or explicitly access-limited source register**, not a reproduction or certification of Wood's entire bibliography. Near-claim links in Section 6 identify which records support each dossier. Entries labeled book-metadata-only are not passed off as independently read primary studies. Publication chronology and source relationship are given separately from access.

B. Wood, Wendy. Good Habits, Bad Habits: The Science of Making Positive Changes That Stick. First edition. Farrar, Straus and Giroux, 2019. eISBN 978-1-250-15908-3.

Source: Supplied EPUB; exact paths and anchors accompany claims.

Access: full text sections read. All 49 spine sections, all 15 chapters, epilogue, phone appendix, notes, bibliography, and 15 in-chapter illustrations. **Relationship:** reviewed book. **Date:** 2019.

Limit: Whole-book reading is not whole-bibliography verification. No print-page locators used.

S01. Wood, Quinn & Kashy (2002). Habits in everyday life: Thought, emotion, and action.

Source: <https://doi.org/10.1037/0022-3514.83.6.1281>.

Access: full text sections read. Introduction, study methods, relevant results and discussion; study table visually inspected in author-hosted PDF. **Relationship:** book source: Chapter 2. **Date:** 2002.

Limit: Diary classification based on repetition and context; not experimental devaluation or a representative estimate of all actions.

S02. Quinn & Wood (2005). Habits across the lifespan. Unpublished manuscript.

Source: <https://gwern.net/doc/psychology/2005-quinn.pdf>.

Access: full text sections read. Methods and relevant results, including age/sample description. Relationship: book source: Chapter 2. Date: 2005; unpublished.

Limit: Available manuscript is not peer-reviewed publication; convenience sample and repeated-context classification.

S03. Lally, van Jaarsveld, Potts & Wardle (2010). How are habits formed: Modelling habit formation in the real world.

Source: <https://doi.org/10.1002/ejsp.674>.

Access: full text sections read. Methods, results, discussion and Table 2; repository PDF read and relevant table inspected. Relationship: book source: Chapter 7. Date: online 2009; issue 2010.

Limit: Self-reported automaticity, selected actions, no untreated control; modeled plateau times are not universal thresholds.

S04. Galla & Duckworth (2015). More than resisting temptation: Beneficial habits mediate the relationship between self-control and positive life outcomes.

Source: <https://doi.org/10.1037/pspp0000026>.

Access: relevant primary methods/results/discussion across the six studies read during revision. The chapter audit inspected prospective and behavioral outcomes and the meditation follow-up; initial checks used the abstract and manuscript introduction/overview. No raw-data reanalysis was performed. Relationship: book source: Chapter 5. Date: 2015.

Limit: Detailed study-level causal claims are restricted; mediation does not by itself establish a causal habit mechanism.

S05. Neal, Wood, Wu & Kurlander (2011). The pull of the past: When do habits persist despite conflict with motives?

Source: <https://doi.org/10.1177/0146167211419863>.

Access: full text sections read. Both experiments' methods/results, relevant discussion and figure in author-hosted PDF. Relationship: book source: Chapter 8. Date: 2011.

Limit: Pre-existing habits measured rather than randomly assigned; several conditional contrasts imprecise; not a weight-loss trial.

S06. de Wit et al. (2018). Shifting the balance between goals and habits: Five failures in experimental habit induction.

Source: <https://pmc.ncbi.nlm.nih.gov/articles/PMC6033090/>.

Access: full text sections read. Abstract, introduction, Experiments 1A/1B methods and results; partial Experiment 2, not complete audit of every experiment/supplement.

Relationship: reader-added comparison. Date: 2018.

Limit: Five-experiment overview comes from abstract; failure of particular induction procedures is not proof that humans lack habits.

S07. Pool et al. (2022). Determining the effects of training duration on the behavioral expression of habitual control in humans: a multilaboratory investigation.

Source: <https://doi.org/10.1101/lm.053413.121>.

Access: full text sections read. Methods, results and discussion in PMC8686594.

Relationship: reader-added comparison; later evidence. Date: 2022.

Limit: No overall training-duration effect; subgroup/model explorations not promoted to confirmatory findings.

S08. Wood, Mazar & Neal (2022). Habits and goals in human behavior: Separate but interacting systems.

Source: <https://doi.org/10.1177/1745691621994226>.

Access: full text sections read. Relevant argument, response-preparation discussion, secondary-analysis account and qualifications; author-hosted manuscript. Relationship: reader-added author position; later evidence. Date: online 20 July 2021; issue March 2022.

Limit: Published response within a theoretical debate, not a response to this commissioned review; no reanalysis performed here.

S09. Keller et al. (2021). Habit formation following routine-based versus time-based cue planning: A randomized controlled trial.

Source: <https://doi.org/10.1111/bjhp.12504>.

Access: full relevant primary methods, participant flow, intervention and results sections read during revision; initial checks used selected repository-manuscript sections. Analysis retained 135 of 192 enrolled participants. Relationship: reader-added intervention; later evidence. Date: 2021.

Limit: Both arms receive planning; no no-planning control. Self-report automaticity is not devaluation.

S10. Diliberti et al. (2004). Increased portion size leads to increased energy intake in a restaurant meal.

Source: <https://doi.org/10.1038/oby.2004.64>.

Access: full text sections read. Methods, results and discussion including portion and consumption quantities. **Relationship:** book source: Chapter 4, note 13. **Date:** 2004.

Limit: Serving condition varied by day; patrons were not individually randomized. Unit correction does not negate relative intake finding.

S11. Kullgren et al. (2013). Individual-versus group-based financial incentives for weight loss: A randomized, controlled trial.

Source: <https://doi.org/10.7326/0003-4819-158-7-201304020-00002>.

Access: full text sections read. Methods, main results, Table 2, discussion and relevant imputation/analysis details in PMC/author PDF. **Relationship:** book source: Chapter 8, note 13. **Date:** 2013.

Limit: Use main imputed estimates separately from observed-weight curves. No automaticity measure; group incentive is a distinct active arm.

S12. Beeken et al. (2017). A brief intervention for weight control based on habit-formation theory delivered through primary care: results from a randomised controlled trial.

Source: <https://www.nature.com/articles/ijo2016206>.

Access: Initial checks reported full methods/results/discussion; revision checked the primary abstract and indexed original results, flowchart and Table 5. The 24-month abstract/flowchart count differs by one and remains unresolved; separate adjusted-model denominators are retained. Complete correction N05 read. **Relationship:** reader-added intervention. **Date:** online 21 November 2016; issue 2017.

Limit: Multicomponent program; long-term advantage over active usual care not established; substantial missing follow-up. The 2021 lipid/glucose unit correction does not change the weight-loss results.

S13. Cleo et al. (2019). Habit-based interventions for weight loss maintenance in adults with overweight and obesity: a randomized controlled trial.

Source: <https://doi.org/10.1038/s41366-018-0067-4>.

Access: abstract only. Publisher abstract. Relationship: reader-added intervention. Date: online 23 April 2018; issue 2019.

Limit: No full methodological audit. Twelve-month intervention follow-up lacks continuing untreated comparator.

S14. Beshears et al. (2021). Creating exercise habits using incentives: The trade-off between flexibility and routinization.

Source: <https://doi.org/10.1287/mnsc.2020.3706>.

Access: full text sections read. Selected introduction, methods and results in PMC8734590.

Relationship: reader-added intervention; later evidence. Date: 2021.

Limit: One workplace environment with free gyms; attendance, not a direct automaticity assay. Flexible/routine advantage through ten post-intervention weeks, no statistically clear difference during weeks 11–40; reminders and common lotteries continued.

S15. Robins (1993). Vietnam veterans' rapid recovery from heroin addiction: A fluke or normal expectation?

Source: <https://doi.org/10.1111/j.1360-0443.1993.tb02123.x>.

Access: Initially model-reported full access; not independently reproduced. The initial access register stated that the complete article was read; the independent chapter crosscheck did not retrieve it in full. Current treatment-selection cautions are grounded in S16, the inspected original report reprinted in 2010, rather than claimed as independently verified 1993 wording. Relationship: book source: Chapter13 notes 13/16/18 and others. Date: 1993.

Limit: Historical observational cohort; no randomized context intervention. Re-addiction, reuse and treatment relapse cannot be interchanged.

S16. Robins et al. (2010). Vietnam veterans three years after Vietnam: How our study changed our view of heroin. Reprint of earlier report.

Source: DOI record; inspected primary reprint.

Access: full relevant primary sections read and independently crosschecked. Sampling, addiction/remission, treatment and reuse discussion on journal pages 203, 206–207; Figure 4 inspected. A separate AI source check also rechecked the reprint's relevant sections/page images. **Relationship:** book source: Chapter 13. **Date:** 2010 reprint of 1977 report.

Limit: Same historical cohort, not a second independently collected replication. Weighted estimates and subgroup denominators differ.

S17. Petrie (1996). Environment is not the most important variable in determining oral morphine consumption in Wistar rats.

Source: <https://doi.org/10.2466/pro.1996.78.2.391>.

Access: full text sections read. Complete paper; methods/results/discussion and relevant table visually inspected in mirror PDF. **Relationship:** reader-added replication comparison. **Date:** April 1996.

Limit: Two studies with changed animal population/measurement; second colony outcome is aggregate. Proposed genetic explanation not proved.

S18. Hofford et al. (2017). Effects of environmental enrichment on self-administration of the short-acting opioid remifentanyl in male rats.

Source: <https://doi.org/10.1007/s00213-017-4734-2>.

Access: abstract only. Author-institution repository abstract, https://uknowledge.uky.edu/psychology_facpub/239/. **Relationship:** book source: Chapter 13, note 24. **Date:** 15 September 2017.

Limit: Different opioid, administration route and design from Rat Park; no human clinical effect size inferred.

S19. Litt et al. (2009). Changing network support for drinking: Network Support Project 2-year follow-up.

Source: <https://doi.org/10.1037/a0015252>.

Access: full relevant primary methods/results read during the chapter audit. Initial checks used the abstract; combined revision access includes the original trial sections in PMC2661035. **Relationship:** book source: Chapter 13, note 35. **Date:** 2009.

Limit: No independent data reanalysis; endpoint windows and percentage of abstinent days are not converted into complete-abstinence probability.

S20. Litt et al. (2016). Network Support II: Randomized controlled trial of Network Support treatment and cognitive behavioral therapy for alcohol use disorder.

Source: <https://doi.org/10.1016/j.drugalcdep.2016.06.010>.

Access: full text sections read. Abstract, introduction and selected participant/discussion excerpts in publisher preview; not full methods/results. Relationship: reader-added favorable follow-up. Date: 1 August 2016 issue.

Limit: Endpoint-specific benefits, not superiority on every drinking measure or proof of habit mediation.

S21. DeFulio & Silverman (2011). Employment-based abstinence reinforcement as a maintenance intervention for the treatment of cocaine dependence: post-intervention outcomes.

Source: <https://doi.org/10.1111/j.1360-0443.2011.03364.x>.

Access: abstract only. Primary abstract, PubMed 21226886. Relationship: book source: Chapter 13, note 37. Date: 2011.

Limit: Selected participants who qualified for employment randomization; urine-test proportions are not participant cure rates.

S22. Smeets, Ashton, Roelands & Quaedflieg (2023). Does stress consistently favor habits over goal-directed behaviors? Data from two preregistered exact replication studies.

Source: <https://doi.org/10.1016/j.ynstr.2023.100528>.

Access: full relevant primary methods/results/discussion read in the combined chapter audit. Initial access stopped earlier in the methods; the revision retains the actual outcome and control-response limits. Relationship: reader-added replication comparison; later evidence. Date: 20 February 2023 online; March issue.

Limit: Abstract reports failure of expected goal-directed behavior in controls, compromising the critical test. Replicates 2009/2010, not book's 2011 extinction study.

S23. Vohs et al. (2021). A multisite preregistered paradigmatic test of the ego-depletion effect.

Source: <https://doi.org/10.1177/0956797621989733>.

Access: indexed primary methods/results/discussion read during chapter audit. Initial checks used the official/author-hosted abstract; the fuller combined access distinguishes recruited and analytic samples. **Relationship:** reader-added comparison; later evidence. **Date:** 14 September 2021 online; October issue.

Limit: Sequential self-control exertion is not interchangeable with acute stress or a direct habit experiment.

S24. Hobson, Bonk & Inzlicht (2017). Rituals decrease the neural response to performance failure.

Source: <https://doi.org/10.7717/peerj.3363>.

Access: full relevant primary methods/results/discussion read in the chapter audit. The revision rechecked self-report results and Table 3 through the indexed original PDF; no dataset reanalysis. **Relationship:** book source: Chapter 14, note 19. **Date:** 30 May 2017.

Limit: Neural error-related response changed; neither behavioral performance nor self-reported affect/motivation improved. This does not establish a clinical anxiety benefit.

S25. Caton et al. (2013). Repetition counts: Repeated exposure increases intake of a novel vegetable in UK pre-school children compared to flavour-flavour and flavour-nutrient learning.

Source: <https://doi.org/10.1017/S0007114512004126>.

Access: full relevant primary exposure methods/results read during the chapter audit. The procedure and training/assessment exposure counts were reconciled in the Cambridge original. **Relationship:** book source: Chapter 14, note 1. **Date:** online 30 October 2012; issue 2013.

Limit: Repeated exposure/acceptance is not an outcome-devaluation demonstration of habitual action.

S26. Jessoe & Rapson (2014). Knowledge is (less) power: Experimental evidence from residential energy use.

Source: <https://doi.org/10.1257/aer.104.4.1417>.

Access: abstract only. American Economic Association abstract and article metadata.

Relationship: book source: Chapter 15. **Date:** April 2014.

Limit: Information/price intervention and subsequent use patterns; no direct habit assay independently verified from full text.

S27. Turner et al. (2026). A computational model of reward learning and habits on social media.

Source: <https://doi.org/10.1038/s41467-026-73547-6>.

Access: full text sections read. Abstract, introduction, model definitions, comparison and relevant confirmatory/exploratory descriptions; not complete methods/supplements.

Relationship: reader-added favorable mixed-process evidence; later evidence. **Date:** 4 June 2026.

Limit: Observational model comparison; confirmatory users from same dataset; model-based perseveration not identical to experimental devaluation.

S28. Ariely & Wertenbroch (2002). Procrastination, deadlines, and performance: Self-control by precommitment.

Source: <https://doi.org/10.1111/1467-9280.00441>.

Access: metadata only. Book source relationship and original identity confirmed through journal notice; original experiments not independently audited here. **Relationship:** book source: Chapter 1, note 1. **Date:** 2002; retracted 2 September 2026.

Limit: Do not retain as ordinary affirmative evidence; no original-data replication performed in this review.

S29. Hyndman & Bisin (2026). Replication of “Procrastination, Deadlines, and Performance: Self-Control by Precommitment”.

Source: <https://doi.org/10.1177/09567976261460772>.

Access: full relevant publisher methods/results/discussion read and independently checked during revision. The main study and changed procedures were checked; the supplement remains unread. **Relationship:** reader-added replication; later evidence. **Date:** 15 July 2026 online; August issue.

Limit: New data and several procedural differences. Demand for commitment and performance efficacy have different results.

N01. Psychological Science (2026). Retraction: Procrastination, deadlines, and performance: Self-control by precommitment.

Source: <https://journals.sagepub.com/doi/full/10.1177/09567976261488042>.

Access: full text sections read. Complete official retraction notice and first-online date. **Relationship:** notice affecting book source S28; later evidence. **Date:** 2 September 2026.

Limit: Editor's reliability judgment; both authors agree. Not an independent finding here about misconduct or Wood's knowledge.

N02. Journal of Personality and Social Psychology (2024). Retraction of Tian et al. (2018).

Source: <https://pubmed.ncbi.nlm.nih.gov/38466335/>.

Access: full notice text read through PubMed/XML. The detailed journal notice is reproduced there; DOI 10.1037/pspa0000374. **Relationship:** notice affecting Chapter 14 note 17; later evidence. **Date:** February 2024 issue; exact online day unresolved.

Limit: Condition-code discrepancies in Study 1; authors' reanalysis invalidated its prior calorie-benefit conclusion. Do not extend anomaly to all studies.

N03. Organizational Behavior and Human Decision Processes (2024). Retraction of "Don't stop believing: Rituals improve performance by decreasing anxiety".

Source: <https://doi.org/10.1016/j.obhdp.2024.104377>.

Access: full publisher-indexed notice text read during revision. Direct publisher retrieval remained blocked; no separately downloaded notice PDF is claimed. Initial access was metadata/excerpts only. **Relationship:** notice affecting Chapter 14 notes 22/24; later evidence. **Date:** November 2024 issue; exact first-online day unresolved.

Limit: Author-requested withdrawal after an audit undermined confidence in data integrity; exact online day unresolved. No independent attribution of responsibility or generalization to other papers.

N04. Wansink & Payne (2017). Corrigendum: Eating behavior and obesity at Chinese buffets.

Source: <https://doi.org/10.1002/oby.21906>.

Access: full text sections read. Entire correction text and revised Table 2. **Relationship:** book source: Chapter 10 note 13 explicitly cites correction. **Date:** retrieved record: 26 June 2017; July 2017 issue.

Limit: Clarifies observation counts/missingness. Authors say conclusions unchanged; this is not independent reanalysis or a verified retraction.

R01. Dan Ariely (2026). Public statement concerning the 2002 procrastination study.

Source: <https://danariely.com/dan-ariely-statement-on-2002-procrastination-study/>.

Access: full text sections read. Complete short public statement. **Relationship:** published author response; later evidence. **Date:** 7 August 2026.

Limit: Author's own account of anomalies, incomplete surviving record and cooperation; not independent resolution of provenance.

R02. Klaus Wertenbroch (2026).

Response to Data Colada posts 138–139.

Source: <https://datacolada.org/wp-content/uploads/Wertenbroch-Response-to-Data-Colada-Posts-138-139.pdf>.

Access: full text sections read. Complete response text as retrieved; bibliography not independently followed. **Relationship:** published author response; later evidence. **Date:** undated response; linked from August/September 2026 posts.

Limit: Statements about his role/data access attributed to him; no independent investigation here.

P01. The Behavioral Scientist. Book-review methodology (requested URL).

Source: <https://www.thebehavioralscientist.com/book-reviews/methodology>.

Access: supplied house methodology text read. the house book-review methodology, sections on claim scope, evidence, all nine anchors and arithmetic. **Relationship:** governing editorial method, not scientific evidence. **Date:** 15 September 2026.

Limit: Local access is not a verified successful live fetch; scores are editorial judgments under this stated rubric.

P02. Jason Hreha. Habits vs. practices.

Source: <https://psychology.substack.com/p/habits-vs-practices>.

Access: full text sections read. Public article. **Relationship:** voice/position calibration only. **Date:** accessed during this review.

Limit: Not independent proof of any empirical criticism of Wood; no public posting date asserted.

P03. Jason Hreha. Behavior change is a matching problem.

Source: <https://psychology.substack.com/p/behavior-change-is-a-matching-problem>.

Access: full text sections read. Public article. **Relationship:** voice/position calibration only. **Date:** accessed during this review.

Limit: Not independent scientific evidence; no opinion or approval attributed to Jason by commissioning alone.

**P04. The Behavioral Scientist.
Behavior-change guide.**

Source: <https://www.thebehavioralscientist.com/guides/behavior-change>.

Access: full text sections read. Relevant introduction and fit-focused sections, not every linked source. **Relationship:** voice/position calibration only. **Date:** accessed during this review.

Limit: Website format/methodology not inferred beyond material actually accessed.

**P05. The Behavioral Scientist. Current local
book-review methodology, Edition 3.0.**

Source: the house book-review methodology.

Access: local source read. Scientific Accuracy, Reference Accuracy and Practical Value anchors, source-evidence rules and score arithmetic inspected during revision. **Relationship:** operational rubric check, distinct from P01's initial failed live-URL request. **Date:** September 2026.

Limit: Reading the local source does not verify its live publication or establish a successful fetch of the public methodology URL.

**U01. Shindou et al. (2018). A silent eligibility
trace enables dopamine-dependent
synaptic plasticity for reinforcement
learning in the mouse striatum.**

Source: <https://doi.org/10.1111/ejn.13921>.

Access: full text sections read during revision. Mouse preparation, timing/pharmacology methods, relevant results and final discussion in complete saved primary PDF. Initial access was book metadata only. **Relationship:** book source. **Date:** online 14 April 2018; March 2019 issue.

Limit: Cellular protocol and theoretical temporal-difference account do not directly test human reward schedules or a universal habit deadline.

**U02. Labrecque, Lee & Wood (2018).
Overthinking Habit. Manuscript under revision.**

Source: Book: notes-ch10.xhtml, note 14.

Access: unavailable. Book citation and narrative only; manuscript not obtained. **Relationship:** book source, not primary-verified here. **Date:** 2018 as cited.

Limit: No primary methodological finding or falsehood judgment assigned.

U03. Monterosso & Wood (2017). Habits of Successful Rehabilitation. Unpublished data.

Source: Book: notes-ch13.xhtml, note 40.

Access: unavailable. Book citation and narrative only; data/report not obtained. **Relationship:** book source, not primary-verified here. **Date:** 2017 as cited.

Limit: Selected successful-rehabilitation description cannot establish a population recovery probability.

U04. Lin, Wood & Monterosso (2016). Healthy eating habits protect against temptations.

Source: <https://doi.org/10.1016/j.appet.2015.11.011>.

Access: full relevant primary sections read in the chapter audit. Complete original obtained; saved source identifier: Lin, Wood and Monterosso (2016); methods/results inspected. Initial access was book metadata only. **Relationship:** book source. **Date:** 2016.

Limit: Controlled short laboratory cue/response-context result verified; no long-term dietary outcome, isolated visual-cue mechanism or independent depletion contrast established.

U05. Judah, Gardner & Aunger (2013). Forming a flossing habit: An exploratory study of the psychological determinants of habit formation.

Source: <https://doi.org/10.1111/j.2044-8287.2012.02086.x>.

Access: metadata only. Book notes-ch9.xhtml, note 15 and corresponding prose. **Relationship:** book source, not primary-verified here. **Date:** 2013.

Limit: Study methods/follow-up not independently audited.

U06. Larcom, Rauch & Willems (2017). The benefits of forced experimentation: Striking evidence from the London Underground network.

Source: <https://doi.org/10.1093/qje/qjx020>. Accepted-paper abstract; authors' account.

Access: primary accepted-paper abstract and authors' account read during the chapter audit. Sources were the Cambridge accepted-paper abstract and the authors' CEPR account; full methods and model estimates were not independently audited. Initial checks recorded book metadata only. **Relationship:** book source: chapter 11, note 1. **Date:** 2017.

Limit: No exact causal commuter estimate independently endorsed.

**U07. Zlatevska, Neumann & Dubelaar (2018).
Mandatory calorie disclosure: A comprehensive
analysis of its effect on consumers and retailers.**

Source: <https://doi.org/10.1016/j.jretai.2017.09.007>.

Access: **metadata only.** Book notes-ch10.xhtml, note 2, including its explicit contrary-evidence qualifier. **Relationship:** book source, not primary-verified here. **Date:** 2018.

Limit: Evidence of Wood's disclosure, not independent verification of a pooled effect.

**U08. Stothart, Mitchum & Yehnert
(2015). The attentional cost of
receiving a cell phone notification.**

Source: <https://doi.org/10.1037/xhp0000100>. Inspected primary abstract.

Access: **primary abstract read during the chapter audit.** The PubMed abstract supports the bounded attention-task summary; full methods and the complete phone-use prescription were not audited. Initial checks recorded book metadata only. **Relationship:** book source: phone appendix, note 4. **Date:** 2015.

Limit: No detailed task result or entire phone-program outcome certified.

**U09. Wansink & Payne (2008). Eating
behavior and obesity at Chinese buffets.**

Source: <https://doi.org/10.1038/oby.2008.286>.

Access: **metadata only.** Original identity from book/correction; original complete text not independently read. **Relationship:** book source; correction separately examined as N04. **Date:** 2008.

Limit: Correction access does not confer full original-paper access.

**U10. Tian et al. (2018). Enacting
rituals to improve self-control.**

Source: <https://doi.org/10.1037/pspa0000113>.

Access: **abstract only.** Original abstract reproduced in retraction record; no complete original-data or all-study audit. **Relationship:** book source; retraction N02. **Date:** 2018; retracted 2024.

Limit: Do not treat the retracted source as current ordinary affirmative evidence.

U11. Brooks et al. (2016). Don't stop believing: Rituals improve performance by decreasing anxiety.

Source: <https://doi.org/10.1016/j.obhdp.2016.07.004>.

Access: relevant original experiments read during revision in the HBS author PDF. Initial access was abstract/introduction and selected publisher-preview sections; the chapter audit added singing-study methods/results. **Relationship:** book source; retraction N03. **Date:** 2016; retracted 2024.

Limit: Original reported results are distinguished from the later withdrawal; no original-data rerun. Notice access is separately recorded under N03.

U12. Norton & Gino (2014). Rituals alleviate grieving for loved ones, lovers, and lotteries.

Source: <https://doi.org/10.1037/a0031772>.

Access: metadata only. Book notes-ch14.xhtml, notes 21 and 23; not a complete primary-source audit. **Relationship:** book source, not primary-verified here. **Date:** 2014.

Limit: No retraction assigned without a verified notice for this specific article; no claim of a complete clean-status audit.

S30. Ma et al. (2023). Effects of habit formation interventions on physical activity habit strength: meta-analysis and meta-regression.

Source: Primary source.

Access: Full relevant methods, effect-size calculation, study characteristics, results and limitations read by the root audit. **Relationship:** reader-added favorable synthesis; later evidence. **Date:** 2023.

Limit: Controlled reported automaticity, not clinical benefit; no pooled-analysis rerun.

S31. Singh et al. (2024). Time to Form a Habit: A Systematic Review and Meta-Analysis of Health Behaviour Habit Formation and Its Determinants.

Source: Primary source.

Access: Full relevant eligibility, synthesis, timing and discussion sections read by root audit. **Relationship:** reader-added timing synthesis; later evidence. **Date:** 2024.

Limit: Four timing studies; heterogeneous definitions; pre-post change is not a controlled effect.

**S32. Crits-Christoph et al. (1999).
Psychosocial treatments for cocaine
dependence: National Institute on Drug Abuse
Collaborative Cocaine Treatment Study.**

Source: Primary source.

Access: Full relevant Treatments, results and Comment sections read and independently crosschecked. Relationship: book source: chapter 13 note 9. Date: 1999.

Limit: Treatment packages; every arm encouraged twelve-step participation; no medication contrast.

**S33. Michiels, Man, Luque & Obeso (2026).
The neural basis of habit formation measured
in goal-directed response switching.**

Source: Primary source.

Access: Extensive indexed primary methods/results/discussion plus PubMed abstract; direct full-page retrieval challenged. Relationship: reader-added favorable mechanistic evidence; later evidence. Date: 2 January 2026; February issue.

Limit: No full downloaded-PDF, supplement or raw-data audit; stimulation result not selective.

**S34. Vázquez-Millán et al. (2026). The
evaluation of devaluation: Deficient
outcome devaluation leads to wrongly
considering goal-directed actions as habits.**

Source: Primary source.

Access: Accessible full primary participants, procedure, preregistered/exploratory analyses, results and discussion read. Relationship: reader-added measurement evidence; later evidence. Date: 7 July 2026.

Limit: Conceptual replication; exploratory older-data reanalyses are not independent new experiments.

**S35. Hardwick et al. (2019). Time-
dependent competition between goal-
directed and habitual response preparation.**

Source: Primary source.

Access: University primary abstract and publication record only; publisher full page not obtained. Relationship: reader-added mechanistic comparison. Date: 30 September 2019 online.

Limit: No full methods/data audit; close to book production rather than established preproduction omission.

S36. Heintzelman & King (2019; online 2018). Routines and Meaning in Life.

Source: Primary source.

Access: Publisher-indexed original methods/results read in chapter audit; title/issue independently verified. **Relationship:** book source: chapter 14 note 27. **Date:** online 2018; 2019 issue.

Limit: Observational routine/meaning association; no imposed-routine causal estimate.

S37. de Ridder et al. (2012). Taking stock of self-control: A meta-analysis of how trait self-control relates to a wide range of behaviors.

Source: Primary source.

Access: Full relevant primary methods, coding, tables and discussion read and independently crosschecked. **Relationship:** book source: chapter 5 notes 16–17. **Date:** 2012.

Limit: Mixed designs; researcher-coded automaticity categories, not experimentally assigned trait self-control.

S38. Phillips & Gardner (2016; online 2015). Habitual exercise instigation (vs. execution) predicts healthy adults' exercise frequency.

Source: Primary source.

Access: Relevant author-uploaded original methods, measures, results and limitations read during crosscheck; PubMed also checked. **Relationship:** reader-added instigation/execution comparison. **Date:** online 2015; 2016 issue.

Limit: Observational reported frequency; no direct measure of mile-by-mile effort or pain.

S39. Kirchner et al. (2013). Geospatial exposure to point-of-sale tobacco: Real-time craving and smoking-cessation outcomes.

Source: Primary source.

Access: Full relevant primary methods, results and limitations read and independently crosschecked. **Relationship:** book source: chapter 6 note 21. **Date:** 2013.

Limit: Daily averages and retailer proximity; no observed store-entry/purchase sequence or causal relapse estimate.

S40. Schultz (2016). Dopamine reward prediction-error signalling: A two-component response.

Source: Primary source.

Access: Full relevant review sections on timing, predictive cues, values and delayed rewards read. Relationship: mechanistic context for reward discussion. Date: 2016.

Limit: Review of neural evidence, not direct human delayed-pay habit experiment.

S41. DeRusso et al. (2010). Instrumental Uncertainty as a Determinant of Behavior Under Interval Schedules of Reinforcement.

Source: Primary source.

Access: Full relevant primary methods/results/discussion read. Relationship: book source: chapter 8 note 19. Date: 28 May 2010.

Limit: Mouse fixed/random-interval schedules; does not establish universal human prescription.

S42. Labrecque, Wood, Neal & Harrington (2017; online 2016). Habit slips: when consumers unintentionally resist new products.

Source: Primary source.

Access: Full relevant primary text, Study 2 instructions/results and limitations read. Relationship: book source: chapter 9 stacking discussion. Date: 2017 issue; 2016 online.

Limit: Planning bundled cue, inhibition and substitution; bottle-weight corroboration limited to subset; P&G funding disclosed.

S43. Burns et al. (2018). Washing with hope: evidence of improved handwashing among children in South Africa from a pilot study of a novel soap technology.

Source: Primary source.

Access: Full relevant primary methods/results/limitations read. Relationship: book source: chapter 8 toy-soap example. Date: 7 June 2018.

Limit: Mixed observed/reported outcomes; no established long-term automatic habit.

S44. Macnamara, Hambrick & Oswald (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis.

Source: Primary source.

Access: Primary abstract plus complete formal correction N06; no independent meta-analysis rerun. Relationship: book source: chapter 7 note 16. Date: 2014; corrected 2018.

Limit: Use corrected estimates; unexplained variance is not a causal estimate of talent.

S45. Schwabe & Wolf (2011). Stress increases behavioral resistance to extinction.

Source: Primary source.

Access: Full relevant primary methods/results/discussion read in chapter audit. Relationship: book source: chapter 12 note 4. Date: 3 March 2011 online; October issue.

Limit: Persistence during extinction does not uniquely identify habit formation; later 2009/2010 replications use different targets.

N05. Beeken et al. (2021). Correction to the brief habit-based weight-control trial.

Source: Primary source.

Access: Complete two-page publisher correction read. DOI 10.1038/s41366-021-00862-x. Relationship: notice affecting reader-added source S12. Date: 7 June 2021 online.

Limit: Table 1 cholesterol/glucose units mg/dl→mmol/l; weight-loss results unchanged.

N06. Macnamara et al. (2018). Corrigendum to the deliberate-practice meta-analysis.

Source: Primary source.

Access: Complete formal notice read. Relationship: notice affecting book source S44. Date: 7 May 2018.

Limit: Dependent-sample adjustment corrected; authors report no substantive conclusion change.

12 Notices and corrections

Notice ledger

Book-linked source	Verified status and chronology	Reason/access boundary	Author position and fair inference
Ariely & Wertenbroch (2002), Chapter 1 note 1	Retracted 2 September 2026 , exact journal first-online date	Complete notice read: editor could no longer attest to reliability given questions about underlying datasets and the reported investigation. N01.	Both authors agreed. Published statements are included in R01/R02, attributed to their authors. No independent misconduct or Wood-knowledge finding.
Tian et al. (2018), Chapter 14 note 17	Retracted; February 2024 issue verified; exact online day unresolved	Journal account read in PubMed abstract: Study 1 condition-code discrepancies and non-supportive reanalysis. N02.	Authors agreed; notice says no discrepancies identified in other datafiles. Do not turn that into certification or generalize Study 1's problem to all studies.
Brooks et al. (2016), Chapter 14 notes 22/24	Retracted; November 2024 issue record; exact online day unresolved	Full publisher-indexed notice text read during revision; author audit undermined confidence in data integrity. N03.	Authors requested withdrawal; no independent person-level responsibility attribution. Exact first-online day remains unresolved.
Wansink & Payne (2008), Chapter 10 note 13	Correction , retrieved record dated 26 June 2017 , July issue; exact-day provenance limited to retrieved record	Full correction/table read; clarifies observation counts and missingness. N04.	Authors say conclusions unchanged. Wood explicitly cites corrected data. No verified retraction assigned to this paper.

Additional correction records

Source	Status and date	Consequence
Beeken weight-control trial	Correction, 7 June 2021 online; full notice read. N05.	Cholesterol/glucose units changed from mg/dl to mmol/l; weight-loss conclusions used here unchanged.
Macnamara deliberate-practice meta-analysis	Corrigendum, 7 May 2018; full notice read. N06.	Dependent-sample adjustment and estimates corrected; authors report no substantive conclusion change.

Evidence used in the assessment

The assessment includes the later exercise follow-up, Keller/Lally denominators, source relationships, primary reward evidence, the distinct ritual source roles and subjective-affect result, the pre-book twelve-step trial, and favorable current automaticity/mechanistic evidence. Scores reflect the stated claim scope, and selection before adverse-evidence searching is not certified. The Beeken 2021 and Macnamara 2018 corrections do not overturn the outcome conclusions used here.

Corrections required in a revised account of the book

The restaurant example should identify its quantities as kilojoules or convert them to kilocalories while retaining the supported relative portion effect. The incentive example should include the group arm and distinguish speculation about dieting patterns from

measured habit formation. The Vietnam discussion should preserve the genuine low re-addiction finding while dropping the inference that it supplies a comparable treatment benchmark. The twelve-step comparison should reflect the treatment packages actually compared, and the ritual discussion should distinguish neural response from measured emotional benefit. The burger-size sentence should match the roughly threefold ratio in its figure. Current presentations of the affected ritual and deadline sources should disclose their later withdrawals. These are specific review findings, not changes already made to Wood's book. A13, A20, A25, A28, A29, A34, A35.

Corrections to avoid introducing

Do not say Wood ignored the buffet correction; that assertion is contradicted by her note. Do not call the popcorn studies Wansink studies. Do not convert the Lally fit-selection fraction into a habit success rate. Do not identify the 2023 stress replications as direct tests of the book's 2011 extinction study. Do not treat the 2010 Robins reprint as an independent contemporary cohort. Do not label a failed retrieval a false book claim, or require every useful support to be removed before counting its benefit. Do not assign the retraction of one ritual paper to every paper with a shared author.

Preparation and author responses

This appendix was prepared with AI-assisted research, drafting, separate chapter audits, challenges of consequential findings, and source/access checks. No author response to this review has been solicited or received. The published author responses described above concern their own research disputes. No independent human scientific sign-off or book-publisher approval is claimed. Any future correction should identify its exact claim IDs, old and new wording, source basis and effect on the nine grades rather than silently changing the review's selected propositions.

Central claims and selection provenance

The three propositions below are retained from the central-claim selection record. Their wording has not been narrowed to improve the grades.

Selection-order limitation: an early topic list named stress, fatigue and distraction as its third claim, while the final selection uses addiction recovery. The preparation record therefore cannot certify one unchanged selection made before adverse-evidence searching. Additional source checks ran during the initial preparation. This is a question-driven editorial review, not a preregistered systematic review.

The adopted scientific grades are 2,2,2 at the retained scope. Supporting evidence, uncertainty and all nine score explanations appear in the technical appendix.

C1 — Habits as the principal engine of everyday persistence

Proposition to be assessed: For frequently repeated everyday behavior, learned associations between contexts and responses are a principal explanation of persistence beyond conscious intentions. Building and relying on those habits, rather than repeatedly exerting willpower, explains much of the success popularly attributed to self-control and provides a more dependable basis for sustained action.

Why this proposition: This is the organizing thesis of Part I and the epilogue. It preserves the book's claim about an explanatory mechanism and its relative practical importance; it is not reduced to the uncontroversial observation that people sometimes repeat actions. Nor is it inflated into the claim that motivation never matters or that every action is a habit.

Core locators:

- `chapter1.xhtml`, paragraph containing “We now know that it’s habit that creates persistence”; preceding discussion of the Ouellette–Wood synthesis and critics.
- `chapter3.xhtml`, paragraph beginning “Our research thereby linked memory, action, context, and persistence,” including the context–response definition and shorthand “automaticity in lieu of conscious motivation.”
- `chapter5.xhtml`, paragraphs beginning “If you talk to someone who runs six miles regularly” and “The good effects that we popularly ascribe to ‘self-control’”; read the latter through its explicit allowance for goals to orient formation.
- `chapter12.xhtml`, closing paragraphs beginning “What we can learn from the demonstrations of resilience” and “Habits are not malleable or creative.” Stress resilience is a supporting extension of C1, not a fourth scored claim.
- `epilogue.xhtml`, opening paragraph and paragraph beginning “More than this, understanding habits normalizes the trials of changing behavior.”

Fair-reading constraints: Wood gives executive control important work, explicitly distinguishes habit from other automatic processes (`chapter3.xhtml`, “Although some researchers equate habits with automaticity”), and says few behaviors are purely habitual or purposeful (`chapter4.xhtml`, discussion following the memory–system evidence). Her position is not a literal two-person brain theory. The assessment must distinguish past-behavior prediction, contextual stability, self-reported automaticity, and experimental evidence of diminished sensitivity to outcomes. Evidence that self-control and habits correlate does not by itself establish that habits causally explain the association.

Evidence that would support a high grade: Convergent, appropriately measured evidence that learned cue–response associations account for substantial persistence and reduced deliberation, including evidence robust to measurement overlap and alternative explanations.

Evidence that would constrain the grade: Treating repeated action or expertise as equivalent to habit; correlational mediation presented as a mechanism established; generalizing task-specific results to the predominance of habit across complex life outcomes.

C2 — Deliberately constructing durable, lower-effort behavior

Proposition to be assessed: By arranging stable cues, reducing friction, choosing rewarding actions, and repeating them, people can deliberately build goal-consistent habits that make recurring behavior increasingly automatic and less dependent on ongoing effortful decisions. The book presents these principles as usable across eating, exercise, work, household, and financial routines—not merely laboratory button presses.

Why this proposition: It is the distinctive practical promise of Part II, named in its context—repetition—reward structure and integrated explicitly in Chapter 10. It tests the link from mechanism to useful results rather than the truism that easier actions are easier. It does not require an implausible guarantee that every reader succeeds or that expert performance becomes entirely unconscious.

Core locators:

- `chapter6.xhtml`, closing application of driving/restraining forces to food and daily contexts.
- `chapter7.xhtml`, Lally habit-formation discussion beginning “Research provides better insight,” and the later distinction “Excellence and repetition are not the same.”
- `chapter8.xhtml`, paragraphs beginning “All of this means one big thing for our purposes” (reward timing), “The workplace wisdom builds trust” (uncertain rewards), and “For scientists, insensitivity to reward” (identification criterion).
- `chapter9.xhtml`, sections on stable cues, stacking after brushing, and swapping products; `notes-ch9.xhtml`, note 18 distinguishes stacking from generic implementation intentions.
- `chapter10.xhtml`, paragraph beginning “If this sounds like a lot of work for your executive, conscious mind,” and the two concluding integrations beginning “Remember that goal of getting your whole family eating together” and “The same principles apply to making the goal of taming your finances.”
- `howtostop.xhtml`, the concrete phone-use application, especially cue removal, added friction, replacement action, and reward.

Fair-reading constraints: Wood acknowledges the effort of getting started, attrition and missed repetitions in the habit-formation study, the importance of genuine enjoyment, the need for conscious preparation, and the difference between repetition and excellence. She recognizes the advantage of fitting a substitute to an existing context rather than merely piling on another task. The review must distinguish initiating a workout from performing it, reducing deliberation from eliminating exertion, observed persistence from automaticity, and component studies from a trial of the book's whole program. An intervention need not make its cues disappear to be useful.

Evidence that would support a high grade: Controlled, sustained improvements in the intended behaviors/outcomes under feasible versions of the book's principles, accompanied where claimed by suitable habit measures. **Evidence that would constrain the grade:** Strong

scope claims resting on short self-report studies, narrow tasks, or physiological findings that do not validate the stated human prescription; insufficient evidence of durability, generalizability, or added benefit of the whole package.

C3 — Environmental change as a major route to addiction recovery

Proposition to be assessed: Changing drug-use contexts and supplying rewarding alternatives can interrupt addiction-supporting habits and sustain recovery. Wood uses returning Vietnam veterans and enriched animal environments to argue that environmental redesign deserves a much larger role in treatment, including treating the veterans' roughly 5% first-year re-addiction figure as a benchmark of possibility against the 40–60% relapse figure she assigns to conventional treatment.

Why this proposition: Chapter 13 is a consequential extension of the book's context-and-habit framework to a high-stakes domain, with explicit treatment implications and a numerical ambition. It joins Chapter 11's discontinuity argument to Part III's broader insistence that outcomes are not solely personal failures. Addiction is selected over a peripheral numerical anecdote because the inference could change how readers evaluate care. Chapters on stress, well-being, and social policy will still receive substantive assessment.

Core locators:

- `chapter13.xhtml`, paragraphs beginning “However, one significant way that addictions and habits differ” and “The habit-like aspects of addiction may lead us into insights.”
- Same file, the Robins discussion beginning “Robins tracked 470 enlisted men,” through “From our habit perspective, the soldiers’ return home.” Supporting book notes: `notes-ch13.xhtml`, notes 13, 15–18.
- Same file, Rat Park and subsequent enrichment discussion; notes 24–26.
- Same file, paragraphs beginning “In this model of drug use disorder, friction accounts for so much,” “The more conventional model for drug use,” and “Why can’t we shoot for 5 percent?”
- Same file, community reinforcement, network support, and therapeutic-workplace discussion; notes 33–38, including the explicit post-treatment limits.

Fair-reading constraints: Wood explicitly says addiction differs from ordinary habit, is multifactorial, involves purposeful behavior as well as automaticity, and is partly hereditary. She says relying on external forces alone is simplistic, describes brain and context accounts as complementary, recognizes the Vietnam comparison was not controlled, and acknowledges loss of benefit when some supports cease. She is not claiming that moving house universally cures addiction or telling readers to stop medication. Those caveats do not, by themselves, validate a comparison of differently defined outcomes, populations, or treatment histories.

Evidence that would support a high grade: Well-characterized original cohort findings together with controlled evidence that feasible context/reinforcement interventions improve recovery, and a defensible basis for the numerical comparison. **Evidence that would constrain the grade:** Confusing drug exposure/use, physiological dependence, and re-addiction;

treating a urine screen as a complete addiction diagnosis; attributing a multi-change cohort outcome uniquely to habit discontinuity; comparing unlike outcomes or portraying established care as a single low-performing alternative.