Athens Journal of Philology 2026, volume (13): 1-21
https://doi.org/10.30958/ajp.X-Y-Z

Do English Proverbs Stand Scrutiny? Twenty Familiar Sayings and the Loss of Conditions

By Aneesah Nishaat[*] & Wasim Parvez Syed[±]

Proverbs, maxims and familiar quotations carry cultural knowledge with great authority, yet they are rarely asked to show their evidence. This paper examines twenty well-known English sayings, grouped into eighteen cases, against findings from psychology, medicine, epidemiology, communication research and meteorology. Each saying is traced to its origin, read both literally and in its conventional sense, set against the best available evidence, and given a graded verdict with a rating of evidence strength. The results are mixed. "Birds of a feather flock together" is well supported; "Let off steam" is not supported and "Opposites attract" is contradicted. The most frequent verdict, under both readings, is that a saying holds only under conditions it does not state. We argue that incompleteness, not falsity, is the characteristic failure of proverbial wisdom, and we identify four routes by which conditions are lost: truncation, universalisation, recontextualisation and compression. The findings also give empirical content to the distinction between a proverb's literal reading and its standard proverbial interpretation. Research on processing fluency helps explain why sayings keep their authority once their conditions are gone.

Keywords: *proverbs, paremiology, folk wisdom, loss of conditions, evidence-based evaluation*

Introduction

Every language keeps a store of short, memorable sentences that its speakers reach for when they want to advise, warn, console or settle an argument. In English these range from anonymous folk proverbs ("An apple a day keeps the doctor away") through the maxims of named writers ("Early to bed and early to rise…") to lines of poetry that have escaped their poems ("Where ignorance is bliss, 'tis folly to be wise"). Such sayings share a peculiar kind of authority. Their credibility rests on their age, on how often we have heard them, on the fact that no individual seems to own them, and on how pleasing they sound, rather than on evidence.

Over the past century the sciences have gathered a large body of evidence on precisely the questions proverbs claim to answer: how to stay healthy, how to raise children, what makes people happy, who falls in love with whom, and when groups outperform individuals. This paper places twenty familiar sayings beside that evidence. It is not an exercise in debunking. Proverbs often work figuratively, and a saying should not be faulted for failing a literal test it never set for itself. For that reason we assess each saying in its conventional sense as well as its literal one. But

[*]Lecturer, Higashi Nippon International University, Japan.
[±]Independent Researcher, India.

proverbs are also offered as practical advice, and in that use it is fair to ask whether the advice holds.

Our central argument is that the typical fault of a familiar saying is not falsity but incompleteness, and that the missing conditions are lost in identifiable ways.

The study is intended as a contribution to paremiology and to the linguistic study of fixed expressions. It asks whether the distinction between a proverb's literal reading and its standard proverbial interpretation (Norrick 1985) matters for truth as well as for meaning, and which of the conditions that proverbs, as generic sentences, leave unstated (Krifka et al. 1995) the evidence can supply.

The paper is organised as follows. The next section reviews research on how proverbs mean, earlier tests of proverbial truth and why sayings feel true. The Method section explains how the sayings were selected and assessed. The Findings present the eighteen cases, and the Discussion examines the pattern of verdicts, the four routes by which conditions are lost, and what the findings mean for proverb studies and linguistics. The paper ends with its limitations and conclusions.

Literature Review

Proverbs and How They Mean

We use *saying* as an umbrella term for proverbs, maxims, aphorisms and familiar quotations. The boundaries are porous: lines by Gray, Nietzsche and Keats now circulate with the anonymous confidence of folk proverbs, and we retain these literary items deliberately, because a sentence that loses its author tends also to lose its qualifications.

A central insight of paremiology is that proverbs mean in context. Norrick (1985), by contrast, deliberately set context aside and treated the proverb as a discrete text. He distinguished a proverb's literal reading from its standard proverbial interpretation, the customary meaning speakers attach to it, and classified the figures that relate the two: synecdoche, metaphor, metonymy, hyperbole and paradox. Yankah (1984) applied the contextual view to contradictory proverbs. He distinguished the proverb as an isolated "concept", a cultural truism open to contradiction, from the proverb in context, where it serves to support an opinion the speaker has already formed and its truth becomes irrelevant. He further argued that many pairs labelled contradictory are not logically opposed even as concepts, and that proverb contradiction is largely a scholarly construct. We accept this account of conversational use. Our question concerns the other use: when a saying is offered as general advice, does the evidence support it, and under what conditions?

Two further strands of linguistic work bear on this question. Studies of proverbiality have identified the cues that make a sentence sound like a proverb, including metaphor, rhyme, alliteration and parallelism (Arora 1984). These cues favour short, balanced sentences, and so leave little room for qualification. Work on the semantics of generic sentences, such as "Birds fly", has shown that such sentences express regularities that tolerate exceptions, and that the conditions restricting them

are usually left implicit (Krifka et al. 1995). Most proverbs are generic sentences of this kind, which is why the question of their unstated conditions arises at all.

Previous Tests of Proverbial Truth

The idea of testing proverbs is not new. Furnham (1987) found that people often endorsed both members of an antonymous pair, such as "Absence makes the heart grow fonder" and "Out of sight, out of mind". When the proverbs were given explicit conditions, however, ratings diverged sharply: absence was judged to strengthen close friendships and to weaken casual ones. Gibbs and Beitel (1995) offered a cognitive account of proverb comprehension, arguing that understanding a proverb draws on conceptual metaphors pervasive in everyday thought, rather than on a special capacity for abstraction. They also reviewed psychological evidence on proverbs' truth. They judged some proverbs inaccurate ("Spare the rod and spoil the child", "Familiarity breeds contempt"), others accurate ("Birds of a feather flock together"), and many true only in some contexts. Their conclusion was that proverbs are best regarded as neither true nor false but as applicable or inapplicable to a situation. This work shows that belief in a proverb depends less on its wording than on the circumstances people imagine for it, and that evidence-based scrutiny of proverbs is feasible. What it has not explained, to our knowledge, is why the circumstances drop out, and why proverbs that fail scrutiny tend to fail in the same way. That is the contribution of this paper.

Why Sayings Feel True

People tend to judge a statement as truer simply because they have heard it before, a phenomenon known as the truth effect. Hasher et al. (1977) first demonstrated it. Students rated plausible but little-known statements, some true and some false, in three sessions held two weeks apart. Some of these statements were repeated across sessions and others were new. Ratings of the repeated statements rose from "uncertain" toward "possibly true", regardless of whether the statements were actually true or false, while ratings of new statements did not change. A later meta-analysis of 51 studies by Dechêne et al. (2010) found the effect to be of medium size and robust across delays, presentation formats and age groups. The effect depends on the statements being ambiguous, meaning people cannot tell from their own knowledge whether they are true, and it was stronger when statements recurred word for word than when only their gist was repeated.

Although these studies used unfamiliar trivia rather than lifelong sayings, proverbs fit both findings: their claims are rarely checkable, and their wording is fixed. They are therefore natural beneficiaries. McGlone and Tofighbakhsh (2000) found that unfamiliar aphorisms in their original rhyming form were rated more accurate than the same aphorisms with the rhyme broken, an advantage that disappeared when readers were warned to separate form from content. More generally, statements that are easier to process are judged truer (Dechêne et al. 2010), and the same fluency makes things seem more attractive (Reber et al. 2004). Fluency explains why a

saying is believed, however, not whether it is false, a distinction we return to in the Discussion.

Method

This study examines twenty English sayings, assessing each in four steps: tracing its origin, stating its literal and conventional claims, reviewing the research evidence, and assigning a verdict. The sayings were chosen according to three criteria. Each had to be in wide use in contemporary English and to make a claim that research could, in principle, test; together, the sayings had to cover a range of domains. One saying was included deliberately as an exception: Keats's "Beauty is truth, truth beauty", which makes no empirical claim, serves as a boundary case to show where the method stops.

Selection of Sayings

We included several contradictory pairs and three literary items. The selection was purposive and biased in a predictable direction: sayings on which research exists, and which researchers suspect to be false, are over-represented. The verdicts below therefore do not estimate the accuracy of English proverbs in general. Candidate sayings were drawn from widely known English sayings and retained only if they met all three criteria. Each retained saying was then checked against *The Oxford Dictionary of Proverbs* (Speake 2015) to confirm that it is an established saying and to fix its standard wording. The twenty sayings form eighteen cases because two contradictory pairs, "Absence makes the heart grow fonder" / "Out of sight, out of mind" and "Too many cooks spoil the broth" / "Many hands make light work", are tested against the same studies. The cases are grouped by domain, from health and child-rearing to relationships, groups and perception.

Literal and Conventional Readings

Each saying was restated as a literal claim and, where the two differed, as a conventional claim: the sense in which the saying is commonly used as advice, which Norrick (1985) calls the standard proverbial interpretation. Because proverbs are generic sentences, which tolerate exceptions and usually leave their restricting conditions unstated (Krifka et al. 1995), the two readings are used to identify which of those conditions the evidence supplies. Where only one claim is given in the Findings, the literal and conventional readings coincide. Conventional readings were checked against *The Oxford Dictionary of Proverbs* (Speake 2015), and, following the philological practice of tracing a saying through its dated attestations (e.g., Payne 2017), origins are taken from the dated attestations it records.

Evidence and Its Grading

For each reading we sought the strongest available evidence, giving priority to meta-analyses and systematic reviews, then to large cohort or nationally representative studies, and then to experiments and smaller studies. Where a finding had been replicated, corrected or jointly reanalysed, the later version was used and the revision noted. Where studies disagreed, both positions are reported, and the disagreement is reflected in the verdict rather than settled by choosing one side. For the three literary items, the origin is the original text rather than a dictionary attestation. The evidence was rated **A** (meta-analyses or several large, consistent studies), **B** (one large study or several smaller ones) or **C** (a single study or indirect evidence). Where the literal and conventional readings rest on different evidence, both ratings are given, literal first (e.g., B/A in Table 1).

Verdicts and Coding

Each reading received one of seven verdicts: *supported*; *partially supported* (right in direction, overstated in size or scope); *conditionally supported* (true under some identifiable conditions and false or reversed under others); *not supported* (adequately tested, no reliable support); *contradicted* (evidence points the other way); *insufficient direct evidence*; or *not empirically testable*. Not supported and contradicted are distinguished by direction: the first means that adequate tests found no reliable effect, the second that they found an effect in the opposite direction. Partially and conditionally supported are distinguished by kind: the first applies when a claim is right in direction but overstated, the second when the evidence identifies conditions under which the claim holds and others under which it fails. Both authors assigned verdicts independently and resolved disagreements by discussion. The verdicts remain interpretive and provisional. Each saying was also coded for four sound patterns, namely rhyme, alliteration, repetition and chiasmus; nine of the twenty carry at least one (see the note to Table 1). The four mechanisms by which conditions are lost (Table 2) were not fixed in advance. They were identified after the verdicts had been assigned, by comparing each saying with its earliest attested or original form and with the conditions that the evidence supplied.

Findings: Eighteen Cases

Health and Daily Habits

1. "An Apple a Day Keeps the Doctor Away"

Origin. The earliest version is a Pembrokeshire rhyme printed in 1866: "Eat an apple on going to bed, and you'll keep the doctor from earning his bread" (Phillips 1866).

Claim. Literally, a daily apple reduces the need for medical care; conventionally, eating fruit and vegetables protects health.

Evidence. In a nationally representative sample of 8,399 US adults, the 9% who ate at least one small apple a day were more likely than others to "keep the doctor away" (no more than one visit a year) in crude comparisons. The difference disappeared once education, smoking and other characteristics were controlled, and apple eaters were only marginally more likely to have taken no prescription medication in the past month (Davis et al. 2015). This cross-sectional result is null rather than negative: it fails to support the claim without showing it to be false. In a meta-analysis of 95 prospective cohort studies, people eating the most apples and pears had lower risks of coronary heart disease, stroke and death than those eating the least, but no clear dose-response was found, and the authors caution that such subtype results rest on few studies (Aune et al. 2017). The conventional claim is better supported. Each additional 200 g of fruit and vegetables a day was associated with an 8–16% lower risk of coronary heart disease, stroke and cardiovascular disease and a 10% lower risk of death from any cause, though with only a 3% lower risk of cancer (Aune et al. 2017). These associations are observational, and residual confounding cannot be ruled out.

Verdict. Literal: **not supported** (B). Conventional: **supported** as an association (A).

2. "Early to Bed and Early to Rise Makes a Man Healthy, Wealthy and Wise"

Origin. Popularised by Benjamin Franklin in *Poor Richard's Almanack*, but recorded a century earlier (Speake 2015).

Claim. Early risers will be healthier, richer and wiser.

Evidence. Among 1,229 Britons aged 65 and over, morning types ('larks') were no healthier, no richer and no better on a short cognitive test than others at baseline; if anything, evening types ('owls') had slightly higher incomes (Gale & Martyn 1998). Over 23 years of follow-up, larks and owls both had slightly lower mortality than the rest of the sample, but this was explained by their shorter time in bed. The authors concluded that when people sleep matters little. Their sample was elderly, however, and "wisdom" was measured only by a brief cognitive screen. In the far larger UK Biobank cohort of 433,268 adults aged 38–73, followed for about six and a half years, definite evening types reported more illness of every kind than definite morning types and had a 10% higher risk of death. The difference held after adjustment for sleep duration, socioeconomic status and other factors (Knutson & von Schantz 2018). The authors suggest it may partly reflect a mismatch between evening types' body clocks and society's schedules, but note that the illness comparisons cannot show causal direction.

Verdict. **Partially supported** (B): a modest health advantage; no wealth; little wisdom.

Child-Rearing

3. "Spare the Rod and Spoil the Child"

Origin. The idea goes back to the biblical Book of Proverbs (13:24), and English forms appear from the fourteenth century (Speake 2015).

Claim. Literally, children who are not physically punished grow up badly behaved; conventionally, undisciplined children are spoiled.

Evidence. An early meta-analysis of 88 studies found parental corporal punishment associated with ten undesirable outcomes, including more aggression and weaker moral internalisation. The one desirable association was with immediate compliance, and that rested on only five studies (Gershoff 2002). A later meta-analysis confined to ordinary spanking, covering 160,927 children, found 13 of 17 outcomes significantly associated with spanking, all in the harmful direction. On reanalysis, spanking no longer predicted children's immediate compliance either way (Gershoff & Grogan-Kaylor 2016). Larzelere and Kuhn (2005) had argued that this literature, including Gershoff's (2002) meta-analysis, is predominantly correlational. Their meta-analysis of 26 studies compared physical punishment directly with other disciplinary tactics. Harms that appeared in correlational designs shrank almost to zero when statistical controls were applied. "Conditional spanking" (two swats, used sparingly with 2- to 6-year-olds to back up milder tactics such as time-out) reduced noncompliance and antisocial behaviour more than most alternatives, although this finding rested largely on a few small clinical trials. Ordinary spanking was about as effective as the alternatives, and only severe or predominant use compared unfavourably. Gershoff and Grogan-Kaylor (2016), however, found effects in longitudinal and experimental studies similar to those in cross-sectional ones. The disagreement, then, is not over whether harsh punishment harms children but over whether mild, occasional spanking does.

Verdict. Literal: **contradicted** (A). Conventional: **partially supported** (A). Frequent or severe physical punishment is linked to harm, while the value of mild, occasional use is disputed.

Mind and Emotion

4. "Money Can't Buy Happiness"

Origin. No single source; the idea recurs across moral traditions.

Claim. Literally, income does not raise happiness; conventionally, money is not sufficient for happiness.

Evidence. Using some 450,000 responses to a US Gallup survey, Kahneman and Deaton (2010) distinguished two kinds of well-being. People's evaluations of their lives rose steadily with the logarithm of income, with no plateau. Their day-to-day emotional well-being, however, stopped improving at around $75,000 a year. Low income also made misfortunes such as illness and divorce more painful. Killingsworth (2021) suggested that this plateau reflected a ceiling in Kahneman

and Deaton's yes/no measure of emotion. Using 1.7 million real-time reports on a 10-point scale from over 33,000 employed US adults, collected through a smartphone app from a sample that was not nationally representative, he found that both feelings and life evaluation rose with log income without a plateau. In his data, people who equated money with success were less happy on average. In a joint reanalysis of Killingsworth's data, the two teams found that the plateau held only for the least happy 15–20% of people, whose happiness stopped rising at around $100,000. For the rest, happiness rose with log income, and for the happiest 30% it rose faster still (Killingsworth et al. 2023). They also stressed that the link is weak. The correlation between happiness and log income was 0.09, and median happiness at a household income of $250,000 was only about five points higher, on a 100-point scale, than at $15,000.

Verdict. Literal: **contradicted** as an association (A), although the association is weak; the verdict holds for both life evaluation and, in most people, day-to-day emotion. Conventional: **partially supported** (A): money helps, with sharply diminishing returns, but is not sufficient.

5. "Ignorance Is Bliss"

Origin. Thomas Gray's "Ode on a Distant Prospect of Eton College" (Gray 1747) ends: "where ignorance is bliss, / 'Tis folly to be wise." The proverb dropped the conditional "where".

Claim. Not knowing makes people happier.

Evidence. Choosing not to know serves several functions, among them regulating emotion and preserving suspense (Hertwig & Engel 2016). In nationally representative surveys in Germany and Spain, 85–90% of adults said they would not want to know when or how they or their partner would die, or whether their marriage would end in divorce. Many also preferred not to learn good news in advance, so as to keep the surprise, and only 1% consistently wanted to know. Those who preferred ignorance were more risk-averse, consistent with a wish to avoid anticipated regret (Gigerenzer & Garcia-Retamero 2017). These findings concern what people say they would choose, however, not whether ignorance leaves them happier. Ignorance also has costs. People who avoid a diagnosis of a treatable disease may face serious health consequences, and such avoidance is likely to be most harmful when something can be done and when it persists (Sweeny et al. 2010). Weighing such cases, Hertwig and Engel (2016) conclude that whether not knowing helps or harms must be judged case by case, depending partly on whether anything can be done with the information.

Verdict. **Conditionally supported** (B), exactly as Gray's conditional wording implied.

6. "Let Off Steam"

Origin. The image of a boiler that must release pressure carries an implicit theory of catharsis, the "hydraulic model" of anger that Bushman (2002) traces to Breuer and Freud.

Claim. Expressing anger reduces it.

Evidence. Bushman (2002) compared two groups of angered participants. Those who hit a punching bag while thinking of their provoker reported more anger, and were more aggressive toward that person, than those who sat quietly for two minutes. The effects were small to moderate. A meta-analysis of 154 studies found that calming activities such as deep breathing, relaxation and mindfulness reliably reduced anger and aggression, whereas arousing activities were ineffective overall (Kjærvik & Bushman 2024). They found that punching or kicking an object had no significant effect, jogging tended to increase anger, and only playful activities such as ball sports reduced it. Putting experiences into words is a different practice from venting. In a meta-analysis of 146 randomised studies, most of them private writing over several short sessions, disclosing thoughts and feelings about personal experiences produced a small but reliable benefit: roughly 54% of participants improved, against 46% of controls (Frattaroli 2006). Benefits were larger when people wrote privately, at home, and without anyone collecting what they wrote. This evidence concerns private expression rather than confiding in another person, which is what English calls "getting it off your chest". The saying, however, describes releasing anger, not reflecting on it.

Verdict. **Not supported** (A): venting through arousing activity does not reduce anger, and in some studies it increases it.

7. "What Doesn't Kill You Makes You Stronger"

Origin. Nietzsche (1889/1990, "Maxims and Arrows", §8): "Out of life's school of war: What does not kill me makes me stronger." The proverb universalised a first-person boast.

Claim. Surviving adversity makes a person stronger.

Evidence. In a national US sample of 2,398 adults followed over several years, people who had experienced some adverse life events reported less distress and higher life satisfaction than those who had experienced none or a great many, and were the least affected by new adversity (Seery et al. 2010). The pattern was U-shaped, and the data, being correlational, cannot show that adversity caused the difference. The usual evidence for growth, people's own reports, is weaker than it looks. Among 122 students who experienced a traumatic event during a two-month study, self-reported growth was largely unrelated to measured change in the same domains, and it accompanied rising rather than falling distress (Frazier et al. 2009). The authors do not conclude that growth never happens, only that asking people whether they have grown does not measure it.

Verdict. **Conditionally supported** (C): some adversity is associated with resilience; a great deal is not.

Learning and Skill

8. "Practice Makes Perfect"

Origin. Earlier English forms such as "use makes mastery" date from the late Middle Ages (Speake 2015).

Claim. Enough practice produces mastery.

Evidence. Ericsson et al. (1993) argued that deliberate practice, rather than innate talent, explains most individual differences in expert performance. However, they defined deliberate practice narrowly as effortful activity designed to improve performance, with immediate feedback and usually a teacher's guidance. They stressed that mere repetition does not automatically bring improvement. A meta-analysis by Macnamara et al. (2014; figures as corrected in Macnamara et al. 2018) later found that deliberate practice explains about 14% of performance variance overall. That breaks down to roughly a quarter in games and music, a fifth in sports, and very little in education (5%) and the professions (1%, not significant). The corrigendum fixed a misapplied adjustment for dependent effect sizes; the figures shifted by a few percentage points without changing the authors' conclusions. A preregistered replication of the original violin study found that solitary practice explained 26% of the variance in performance, against 48% in the original (Macnamara & Maitra 2019). It also found that the best violinists had not practised significantly more than the good ones.

Verdict. **Partially supported** (A): deliberate practice makes people better, not perfect; mere repetition does not reliably improve performance.

9. "You Can't Teach an Old Dog New Tricks"

Origin. Sixteenth-century advice about training dogs, later applied to people (Speake 2015).

Claim. Literally, older people cannot learn new skills; conventionally, learning becomes harder with age.

Evidence. Park et al. (2014) studied older adults who took up demanding new activities. Adults aged 60 to 90 spent about 16 hours a week for three months learning quilting or digital photography, and they mastered the skills. Their episodic memory also improved more than that of comparison groups who socialised or did undemanding tasks, although most other abilities did not. Research on cognitive ageing, however, shows that some abilities decline with age. Examining when this decline begins, Salthouse (2009) found that reasoning, speed, spatial ability and memory decline gradually from the twenties or thirties, and faster after 60, while vocabulary and general knowledge keep growing until at least 60. Older adults, then, can learn new skills, but the abilities on which learning draws decline with age, and Park et al.'s participants spent a substantial amount of time learning theirs.

Verdict. Literal: **contradicted** (B). Conventional: **partially supported** (B).

Love and Relationships

10. "Opposites Attract"

Origin. A metaphor from magnetism.

Claim. People are drawn to partners unlike themselves.

Evidence. A meta-analysis of 313 studies found that both actual (r = .47) and perceived (r = .39) similarity predict attraction (Montoya et al. 2008). The effect of actual similarity was strongest in laboratory studies where people never met, much weaker after a short interaction, and undetectable in existing relationships. Perceived similarity, however, predicted attraction in all three settings. No setting showed that dissimilarity attracts. In about 79,000 UK Biobank couples, partners resembled each other significantly on 115 of 133 traits, although the correlations were mostly modest (mean r = .19). Only three traits showed a significant tendency for partners to differ, the largest being chronotype: morning types were somewhat more often partnered with evening types (Horwitz et al. 2023). Such correlations show that partners resemble each other, but not why. People may choose similar partners, meet through shared circumstances, or grow alike over time.

Verdict. **Contradicted** (A).

11. "Birds of a Feather Flock Together"

Origin. Recorded from the sixteenth century (Speake 2015).

Claim. People associate with others who resemble them.

Evidence. The evidence against "Opposites attract" supports this saying. In their meta-analysis of partner similarity, Horwitz et al. (2023) found that resemblance is strongest for political and religious attitudes, education and some kinds of substance use, and weakest for personality. Correlations between partners ranged from r = .08 for extraversion to r = .58 for political values.

Verdict. **Supported** (A): scrutiny does not always end in rejection.

12. "Absence Makes the Heart Grow Fonder" / "Out of Sight, Out of Mind"

Origin. Both are long established and often cited as a contradictory pair (Yankah 1984).

Claim. Separation strengthens affection; separation weakens it.

Evidence. In a one-week diary study of 63 dating couples, mostly students, long-distance partners reported feeling more intimate after their interactions than geographically close partners did. The difference was linked to their disclosing more about themselves, and to their perceiving their partners as disclosing more than the partners themselves reported, especially in text-based media (Jiang & Hancock 2013). In a separate study, Stafford and Merolla (2007) found that long-distance dating relationships were more stable than geographically close ones while the partners remained apart, but were likely to end once the partners lived near each other. Extreme idealisation during separation predicted this instability. The evidence therefore supports the first saying, with conditions. It offers no direct test of the

second: affection in these committed couples faded when partners came back into sight, not while they were apart. Casual ties, for which Furnham's (1987) participants judged absence harmful, were not studied.

Verdict. "Absence…": **conditionally supported** (C). "Out of sight…": **insufficient direct evidence**.

13. "Familiarity Breeds Contempt"

Origin. Found in Aesop and in Chaucer (Speake 2015).

Claim. The better we know people, the less we like them.

Evidence. Norton et al. (2007) found that people who read more randomly selected traits about an unknown person liked that person less, and that online daters liked their dates less after meeting them than before. They attributed both results to the discovery of dissimilarities. Reis et al. (2011) found that same-sex strangers who interacted more, whether face to face or in online chats, liked each other more. After publishing replies to each other's work, the Norton and Reis teams wrote a joint review (Finkel et al. 2015). In it, they proposed that familiarity's effect depends on the stage of a relationship, from mere awareness of someone, through surface contact, to mutual interdependence. They also reinterpreted the trait-list result as a problem of coherence rather than dissimilarity. A long list of random traits is harder to form into a clear impression. The dispute was thus settled not by choosing a winner but by specifying conditions, although the proposed model has yet to be tested.

Verdict. **Conditionally supported** (B).

Groups and Cooperation

14. "Two Heads Are Better Than One"

Origin. Recorded in Heywood's proverb collection (Speake 2015).

Claim. Two people deciding together outperform one.

Evidence. In a brief visual discrimination task, pairs of similar sensitivity who could discuss their disagreements outperformed their better member, even without feedback on accuracy (Bahrami et al. 2010). Pairs allowed to share only their choices gained nothing. When one member's sensitivity was reduced to less than about 40% of the other's, most pairs did worse than the better member alone.

Verdict. **Conditionally supported** (C): two heads are better when they are similarly competent and can communicate their confidence.

15. "Too Many Cooks Spoil the Broth" / "Many Hands Make Light Work"

Origin. Both recorded in English from at least the sixteenth century (Speake 2015).

Claim. Adding people spoils the result; adding people eases the work.

Evidence. A meta-analysis of 78 studies found that people exert less effort when their work is pooled with others' than when they work alone, a moderate and

robust effect (Karau & Williams 1993). Loafing was reduced when individual contributions could be evaluated and when the task was meaningful, and it disappeared in highly valued groups. The finding cuts both ways. Each member works less hard, so the work is lighter for each hand, but the group falls short of what its members could achieve. Steiner (1972, as cited in Karau & Williams 1993) traced such group shortfalls to reduced motivation or to losses of coordination, the more obvious way in which "too many cooks" might spoil the broth. Neither saying, however, has been tested directly as stated.

Verdict. "Too many cooks…": **insufficient direct evidence**, although social loafing is one mechanism by which groups underperform. "Many hands…": **conditionally supported** (B): the work is lighter for each person, but less gets done than the group could achieve unless contributions are visible.

The Physical World and Perception

16. "Lightning Never Strikes the Same Place Twice"

Origin. An American saying, first recorded in P. H. Myers's *The Prisoner of the Border* (Myers 1857, cited in Speake 2015).

Claim. Literally, lightning never strikes the same place twice; conventionally, where never is understood as "seldom" (Norrick 1985), misfortunes rarely repeat.

Evidence. The US National Weather Service lists the saying as a myth: lightning often strikes the same place repeatedly, especially tall, pointed and isolated objects, and the Empire State Building is struck an average of 23 times a year (National Weather Service n.d.). The conventional claim is a form of the gambler's fallacy (Tversky & Kahneman 1971). Past misfortune does not lower future risk, and may signal a persistent one. Victims of burglary and other crimes, for example, face a raised risk of being victimised again (Farrell & Pease 1993).

Verdict. Literal: **contradicted** (B). Conventional: **not supported** (C).

17. "Seeing Is Believing"

Origin. First recorded in 1609, in a manuscript by S. Harward (Speake 2015).

Claim. What we see can be trusted.

Evidence. Research on attention shows that perception is selective. In a study of inattentional blindness, Simons and Chabris (1999) asked observers to count basketball passes in a video. More than half of them failed to notice a person in a gorilla suit walking through the scene, and the rate ranged from 17% to 92% depending on how hard the counting task was and which team they were watching. Research on memory shows that recollections of what we have seen are also reconstructed. In a classic experiment, Loftus and Palmer (1974) showed students a film of a car accident and asked how fast the cars were going when they "smashed" into, or "hit", each other. A week later, students who had been asked the "smashed" question were more than twice as likely to report broken glass that was not in the film (32% vs. 14%). Yet in both studies many observers were accurate. A substantial share noticed the gorilla, and most students, even in the "smashed" condition, reported

no broken glass. Seeing fails under particular conditions, when attention is engaged elsewhere and when memory is shaped by later suggestion, rather than in general.

Verdict. **Conditionally supported** (A).

18. "Beauty Is Truth, Truth Beauty"

Origin. The closing line of John Keats's "Ode on a Grecian Urn" (Keats 1820).

Claim. Beauty and truth are one.

Evidence. The line is a philosophical proposition, not an empirical one. Yet it describes human judgment well. The fluency that makes a stimulus feel beautiful also makes it feel true (Reber et al. 2004, Dechêne et al. 2010).

Verdict. **Not empirically testable**, but apt as a description of how the mind confuses beauty with truth.

Table 1. *Summary: Evidence Strength and Verdicts for the Twenty Sayings*

No.	Saying	Strength	Literal verdict	Conventional verdict
1	An apple a day…	B / A	Not supported	Supported
2	Early to bed…	B	Partially supported	Partially supported
3	Spare the rod…	A	Contradicted	Partially supported
4	Money can't buy happiness	A	Contradicted	Partially supported
5	Ignorance is bliss	B	Conditionally supported	Conditionally supported
6	Let off steam	A	Not supported	Not supported
7	What doesn't kill you…	C	Conditionally supported	Conditionally supported
8	Practice makes perfect	A	Partially supported	Partially supported
9	You can't teach an old dog…	B	Contradicted	Partially supported
10	Opposites attract	A	Contradicted	Contradicted
11	Birds of a feather…	A	Supported	Supported
12a	Absence makes the heart…	C	Conditionally supported	Conditionally supported
12b	Out of sight, out of mind	—	Insufficient evidence	Insufficient evidence
13	Familiarity breeds contempt	B	Conditionally supported	Conditionally supported
14	Two heads…	C	Conditionally supported	Conditionally supported
15a	Too many cooks…	—	Insufficient evidence	Insufficient evidence
15b	Many hands…	B	Conditionally supported	Conditionally supported
16	Lightning never strikes the same place twice	B / C	Contradicted	Not supported
17	Seeing is believing	A	Conditionally supported	Conditionally supported
18	Beauty is truth…	—	Not testable	Not testable

Note: A = meta-analyses or several large, consistent studies; B = one large or several smaller studies; C = a single study or indirect evidence. Where two ratings are given, the first refers to the literal and the second to the conventional reading. Sayings marked by rhyme, alliteration, repetition or chiasmus: nos. 1, 2, 3, 8, 11, 12b, 15b, 17 and 18.
Source: Authors' compilation.

Discussion

The Pattern of Verdicts

Under the conventional reading (Table 1), two sayings were supported, five partially supported, seven conditionally supported, two not supported, one contradicted, two lacked direct evidence, and one was not testable. The literal reading was harsher, with five contradicted. Two points stand out. First, under both readings the largest group consists of sayings true only under conditions they do not state: proverbs fail mainly by being incomplete. Second, four sayings that fail literally survive conventionally, so testing only literal claims would exaggerate the failure of proverbial wisdom. The two clearest failures, "Let off steam" and "Opposites attract", share a feature. Each rests on a metaphor, steam pressure or magnetism, that imports a false theory.

Contradictory pairs, which Yankah (1984) argued are unproblematic in conversation, look different when treated as advice. The evidence sometimes takes a side (similarity over difference), sometimes assigns each saying its conditions, and sometimes shows that one member of a pair has simply never been tested ("Out of sight", "Too many cooks").

How Conditions Are Lost

The cases suggest four routes by which qualifications disappear (Table 2).

Table 2. *Mechanisms by Which Sayings Lose Their Conditions*

Mechanism	What happens	Examples
Truncation	The qualifying part of a sentence is dropped	"Ignorance is bliss" (Gray's "where" lost)
Universalisation	A statement about oneself becomes a statement about everyone	"What doesn't kill you…" (Nietzsche's "me")
Recontextualisation	A sentence moves into a new domain	"Old dog" (from dogs to people); "Let off steam", "Opposites attract" (from machines and magnets to emotion)
Compression	A conditional regularity is squeezed into a short, balanced sentence	"Practice makes perfect"; "Two heads…"; "Familiarity breeds contempt"; "Absence makes the heart…"

Source: Authors' compilation.

The first three mechanisms can be observed directly because the source survives, which is why the literary items illustrate them so clearly. Compression is inferred, but its effect is visible in the evidence. The conditions research uncovers ("when partners are similarly competent and can talk", "if the practice is deliberate") are exactly what a short, balanced sentence has no room for. When conditions are specified, listeners' judgements change. Furnham's (1987) participants judged "Look before you leap" true of marriage and "He who hesitates is lost" false of it. The proverb itself carries no such qualifier, and out of context each is offered as a general truth. Yankah (1984) made a similar observation from a folkloristic standpoint. Rhyme, balance and brevity lend proverbs the character of "absolute,

unqualified rules encapsulated in abbreviated language" (p. 10). Apparently contradictory pairs carry an unspoken qualification that each applies only in some contexts. A qualified form such as "Absence sometimes makes the heart grow fonder" would hardly pass as a proverb in English. The qualifier that evidence restores is exactly what proverbial form excludes.

The Role of Form

McGlone and Tofighbakhsh (2000) found that rhyme makes sayings seem more accurate. If rhyme lends this kind of unearned authority, a false saying with rhyme might survive more easily than a plain one, and we might therefore expect sound-patterned sayings to fare worse against the evidence. In our sample they did not. Under the conventional reading, of the nine sayings marked by rhyme, alliteration, repetition or chiasmus, two were supported and three partially supported, and none was contradicted. In the case of the eleven plain sayings, none was supported and one was contradicted. The sample is small, but the comparison clarifies what fluency can and cannot explain. Fluency explains why a saying is believed (Dechêne et al. 2010), and why it may still be believed after a correction has been forgotten (Skurnik et al. 2005). It does not explain why a saying is wrong.

Transmission and Revision

Proverbs transmit beliefs without inviting questions. Adults who report having been spanked as children are more likely to endorse spanking (Gershoff & Grogan-Kaylor 2016). Whether this reflects transmission or adults rationalising their own upbringing, the result is the same: a saying and a practice can sustain each other across generations. Science, by contrast, revises itself in public, through replications (Macnamara & Maitra 2019), corrections, even minor ones (Macnamara et al. 2018), and adversarial collaborations (Finkel et al. 2015, Killingsworth et al. 2023). Several of the verdicts in this paper rest on such revisions, whereas a proverb has no means of correcting itself. Its wording stays the same whatever the evidence shows.

Implications for Proverb Studies and Linguistics

The findings bear on proverb studies and linguistics in three ways. First, they give empirical content to Norrick's (1985) distinction between a proverb's literal reading and its standard proverbial interpretation. Four sayings that fail literally survive under their conventional reading (Table 1), so the proverbial interpretation is not merely a figurative gloss on the words: it often carries a narrower and more defensible claim than the words themselves. Second, the four mechanisms describe changes in proverb meaning that are linguistic in kind. Truncation and compression shorten a sentence and favour the balanced, parallel form that, together with rhyme and alliteration, marks it as proverbial (Arora 1984). Universalisation turns a first-person statement into a generic one and recontextualisation extends a sentence metaphorically from boilers, magnets or dogs to people. Each process makes a

saying more proverbial and, at the same time, removes a condition on which its truth depends. Third, the results support treating proverbs as generic sentences (Krifka et al. 1995). In conversation, as Yankah (1984) observed, their exceptions are tolerated and their conditions supplied by context. Offered as general advice, however, they leave the listener to supply conditions that the evidence shows to matter. For philology, the cases also suggest a use for dated attestations. Comparing a saying's earliest recorded or original form with its modern one, as with Gray's "where" or Nietzsche's "me", shows not only how its wording changed but which qualifications it lost.

Limitations

The sample is small, purposive and English only, and it probably over-represents sayings that fare poorly. The distribution of verdicts should not be read as an estimate for English proverbs in general. Evidence strength varies, and verdicts rated C could change with a single new study. Our rating scale also counts studies rather than weighing their design, so a large observational study can earn the same grade as an experiment. Much of the evidence comes from North American and European samples and is observational. The verdicts and conventional readings, although independently coded and checked against a dictionary, remain interpretive. The four mechanisms were identified from the same cases they are used to explain and account for only some of the sayings, so they should be treated as hypotheses to be tested on new material. A rule-based sample drawn from a proverb dictionary would allow firmer conclusions.

Because our sample is limited to English, comparison across languages is a promising direction for future research, since other languages may preserve conditions that English has lost. The Japanese *hayaoki wa sanmon no toku* ("early rising brings three small coins") promises early risers only a small gain, a claim closer to the evidence than Franklin's triple promise. *Hyakubun wa ikken ni shikazu* ("a hundred hearings are not worth one seeing"), a saying of Chinese origin, makes a comparative claim that is easier to defend than "Seeing is believing".

Conclusions

Placed beside the evidence, familiar English sayings are neither simply right nor simply wrong. Read in their conventional sense, most contain real insight; read literally, many fail; and under either reading, the largest group holds only under conditions the sayings leave unsaid. Earlier work showed that people judge a proverb differently once its conditions are specified (Furnham 1987) and that proverbs are better judged applicable or inapplicable than true or false (Gibbs & Beitel 1995). The cases suggest where the conditions go. Some are cut away, as Gray's "where" was; some disappear when one person's boast becomes a rule for everyone; some are left behind when a sentence moves from boilers, magnets or dogs to human beings; and some never fit into a short, balanced sentence at all. What remains is easy to say and easy to remember, and research on processing fluency

suggests why it stays believed. For proverb studies, the findings show two things. A saying's usual meaning is often narrower, and easier to defend, than its literal words. Comparing a saying's earliest recorded form with its modern one can also show which conditions it has lost. The four routes come from these cases only, so they still need to be tested on a larger, systematically chosen sample.

Nothing here implies that proverbs should be abandoned. Their value as consolation, humour and shared reference does not depend on their accuracy. This research offers no final verdict on any saying. We offer instead an invitation: to ask of each familiar saying not only what it means, but whether, and when, it is true.

Acknowledgments

The authors used Claude (Anthropic) to assist with formatting the manuscript and checking that the references conform to the journal's requirements.

References

Arora SL (1984) The perception of proverbiality. *Proverbium* 1: 1–38.

Aune D, Giovannucci E, Boffetta P, Fadnes LT, Keum N, Norat T, et al. (2017) Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality: A systematic review and dose–response meta-analysis of prospective studies. *International Journal of Epidemiology* 46(3): 1029–1056.

Bahrami B, Olsen K, Latham PE, Roepstorff A, Rees G, Frith CD (2010) Optimally interacting minds. *Science* 329(5995): 1081–1085.

Bushman BJ (2002) Does venting anger feed or extinguish the flame? Catharsis, rumination, distraction, anger, and aggressive responding. *Personality and Social Psychology Bulletin* 28(6): 724–731.

Davis MA, Bynum JPW, Sirovich BE (2015) Association between apple consumption and physician visits: Appealing the conventional wisdom that an apple a day keeps the doctor away. *JAMA Internal Medicine* 175(5): 777–783.

Dechêne A, Stahl C, Hansen J, Wänke M (2010) The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review* 14(2): 238–257.

Ericsson KA, Krampe RT, Tesch-Römer C (1993) The role of deliberate practice in the acquisition of expert performance. *Psychological Review* 100(3): 363–406.

Farrell G, Pease K (1993) *Once Bitten, Twice Bitten: Repeat Victimisation and Its Implications for Crime Prevention*. Crime Prevention Unit Paper 46. London: Home Office Police Department.

Finkel EJ, Norton MI, Reis HT, Ariely D, Caprariello PA, Eastwick PW, et al. (2015) When does familiarity promote versus undermine interpersonal attraction? A proposed integrative model from erstwhile adversaries. *Perspectives on Psychological Science* 10(1): 3–19.

Frattaroli J (2006) Experimental disclosure and its moderators: A meta-analysis. *Psychological Bulletin* 132(6): 823–865.

Frazier P, Tennen H, Gavian M, Park C, Tomich P, Tashiro T (2009) Does self-reported posttraumatic growth reflect genuine positive change? *Psychological Science* 20(7): 912–919.

Furnham A (1987) The proverbial truth: Contextually reconciling and the truthfulness of antonymous proverbs. *Journal of Language and Social Psychology* 6(1): 49–55.
Gale C, Martyn C (1998) Larks and owls and health, wealth, and wisdom. *BMJ* 317(7174): 1675–1677.
Gershoff ET (2002) Corporal punishment by parents and associated child behaviors and experiences: A meta-analytic and theoretical review. *Psychological Bulletin* 128(4): 539–579.
Gershoff ET, Grogan-Kaylor A (2016) Spanking and child outcomes: Old controversies and new meta-analyses. *Journal of Family Psychology* 30(4): 453–469.
Gibbs RW, Beitel D (1995) What proverb understanding reveals about how people think. *Psychological Bulletin* 118(1): 133–154.
Gigerenzer G, Garcia-Retamero R (2017) Cassandra's regret: The psychology of not wanting to know. *Psychological Review* 124(2): 179–196.
Gray T (1747) *An Ode on a Distant Prospect of Eton College*. London: R. Dodsley.
Hasher L, Goldstein D, Toppino T (1977) Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior* 16(1): 107–112.
Hertwig R, Engel C (2016) Homo ignorans: Deliberately choosing not to know. *Perspectives on Psychological Science* 11(3): 359–372.
Horwitz TB, Balbona JV, Paulich KN, Keller MC (2023) Evidence of correlations between human partners based on systematic reviews and meta-analyses of 22 traits and UK Biobank analysis of 133 traits. *Nature Human Behaviour* 7(9): 1568–1583.
Jiang LC, Hancock JT (2013) Absence makes the communication grow fonder: Geographic separation, interpersonal media, and intimacy in dating relationships. *Journal of Communication* 63(3): 556–577.
Kahneman D, Deaton A (2010) High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences of the United States of America* 107(38): 16489–16493.
Karau SJ, Williams KD (1993) Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology* 65(4): 681–706.
Keats J (1820) *Lamia, Isabella, The Eve of St. Agnes, and Other Poems*. London: Taylor and Hessey.
Killingsworth MA (2021) Experienced well-being rises with income, even above $75,000 per year. *Proceedings of the National Academy of Sciences of the United States of America* 118(4): e2016976118.
Killingsworth MA, Kahneman D, Mellers B (2023) Income and emotional well-being: A conflict resolved. *Proceedings of the National Academy of Sciences of the United States of America* 120(10): e2208661120.
Kjærvik SL, Bushman BJ (2024) A meta-analytic review of anger management activities that increase or decrease arousal: What fuels or douses rage? *Clinical Psychology Review* 109: 102414.
Knutson KL, von Schantz M (2018) Associations between chronotype, morbidity and mortality in the UK Biobank cohort. *Chronobiology International* 35(8): 1045–1053.
Krifka M, Pelletier FJ, Carlson GN, ter Meulen A, Chierchia G, Link G (1995) Genericity: An introduction. In GN Carlson, FJ Pelletier (eds.), *The Generic Book*, 1–124. Chicago: University of Chicago Press.
Larzelere RE, Kuhn BR (2005) Comparing child outcomes of physical punishment and alternative disciplinary tactics: A meta-analysis. *Clinical Child and Family Psychology Review* 8(1): 1–37.

Loftus EF, Palmer JC (1974) Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior* 13(5): 585–589.

Macnamara BN, Maitra M (2019) The role of deliberate practice in expert performance: Revisiting Ericsson, Krampe & Tesch-Römer (1993). *Royal Society Open Science* 6(8): 190327.

Macnamara BN, Hambrick DZ, Oswald FL (2014) Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science* 25(8): 1608–1618.

Macnamara BN, Hambrick DZ, Oswald FL (2018) Corrigendum: Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science* 29(7): 1202–1204.

McGlone MS, Tofighbakhsh J (2000) Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science* 11(5): 424–428.

Montoya RM, Horton RS, Kirchner J (2008) Is actual similarity necessary for attraction? A meta-analysis of actual and perceived similarity. *Journal of Social and Personal Relationships* 25(6): 889–922.

Myers PH (1857) *The Prisoner of the Border: A Tale of 1838*. New York: Derby & Jackson.

National Weather Service (n.d.) *Lightning myths*. Available at: https://www.weather.gov/safety/lightning-myths. (Accessed 2 October 2026.)

Nietzsche F (1990) *Twilight of the Idols and The Anti-Christ*. Translated by RJ Hollingdale. London: Penguin. (Original work published 1889.)

Norrick NR (1985) *How Proverbs Mean: Semantic Studies in English Proverbs*. Berlin: Mouton.

Norton MI, Frost JH, Ariely D (2007) Less is more: The lure of ambiguity, or why familiarity breeds contempt. *Journal of Personality and Social Psychology* 92(1): 97–105.

Park DC, Lodi-Smith J, Drew L, Haber S, Hebrank A, Bischof GN, et al. (2014) The impact of sustained engagement on cognitive function in older adults: The Synapse Project. *Psychological Science* 25(1): 103–112.

Payne MJ (2017) "The milk of birds": A proverbial phrase, ancient and modern, and its link to nature. *Athens Journal of Philology* 4(2): 95–110.

Phillips JP (1866) A Pembrokeshire proverb. *Notes and Queries* s3-IX: 153.

Reber R, Schwarz N, Winkielman P (2004) Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review* 8(4): 364–382.

Reis HT, Maniaci MR, Caprariello PA, Eastwick PW, Finkel EJ (2011) Familiarity does indeed promote attraction in live interaction. *Journal of Personality and Social Psychology* 101(3): 557–570.

Salthouse TA (2009) When does age-related cognitive decline begin? *Neurobiology of Aging* 30(4): 507–514.

Seery MD, Holman EA, Silver RC (2010) Whatever does not kill us: Cumulative lifetime adversity, vulnerability, and resilience. *Journal of Personality and Social Psychology* 99(6): 1025–1041.

Simons DJ, Chabris CF (1999) Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception* 28(9): 1059–1074.

Skurnik I, Yoon C, Park DC, Schwarz N (2005) How warnings about false claims become recommendations. *Journal of Consumer Research* 31(4): 713–724.

Speake J (Ed.) (2015) *The Oxford Dictionary of Proverbs*. 6th Edition. Oxford: Oxford University Press.

Stafford L, Merolla AJ (2007) Idealization, reunions, and stability in long-distance dating relationships. *Journal of Social and Personal Relationships* 24(1): 37–54.
Sweeny K, Melnyk D, Miller W, Shepperd JA (2010) Information avoidance: Who, what, when, and why. *Review of General Psychology* 14(4): 340–353.
Tversky A, Kahneman D (1971) Belief in the law of small numbers. *Psychological Bulletin* 76(2): 105–110.
Yankah K (1984) Do proverbs contradict? *Folklore Forum* 17(1): 2–19.