The Illusion of Sufficiency: Can Knowledge of Fallacies Alone Cure the Ignorance of Our Age?

First article in the "Illusion of Sufficiency" trilogy

Ramin Saadat
Independent Researcher; Founder, British Institute for Ignorance Studies (BIIS) (biis.uk)

AI Disclosure and Translation Statement:
The original manuscript of this work was authored in Farsi (Persian) by the researcher. The English version was produced through AI-assisted translation and linguistic refinement, utilizing Google Translate and Large Language Models. The author has meticulously reviewed, edited, and verified the final English text to ensure that all theoretical nuances, philosophical arguments, and conceptual frameworks accurately reflect the original Persian content. The author remains fully responsible for the integrity and accuracy of the translation. The original Persian manuscript is available from the author upon reasonable request.

Abstract

Can knowledge of fallacies alone serve as an antidote to the ignorance of our age? This article, the first in a trilogy under the shared title "The Illusion of Sufficiency", answers no. This does not mean such knowledge is useless or misguided. It means that knowledge alone is insufficient, and at times even misleading. The article's argument rests on three grounds.

First, a structural argument. The list of fallacies has never, in practice, been completed, and recognising the form of an argument is not the same as assessing the truth of its premises. However carefully and broadly fallacy training is conducted, it therefore offers closed coverage, not unlimited coverage. It resembles a narrow-spectrum antibiotic, not hand washing.

Second, the empirical evidence. Existing research on the effectiveness of this training gives no consistent result, and the dispute remains open. The article's central argument does not depend on resolving it.

Third, the side costs. The label "fallacy" can become a weapon of identity against an ideological opponent, and in the attention economy, the name of a fallacy, rather than the actual skill of spotting it, is the commodity that spreads more easily.

Keywords: logical fallacies; critical thinking; manufactured ignorance; political polarisation; identity weaponisation; attention economy

---

Introduction and the Place of This Article

This article is the first of three independent but related pieces. Each examines, in turn, a common claim about the sufficiency of "individual-level cognitive tools" (usually classed under "critical thinking" studies) against the ignorance of our age.

The first article addresses logical fallacies. The second addresses cognitive biases. The third addresses a gap, already recognised in the critical-thinking literature itself, between having a skill and being disposed to apply it.

Splitting one broad concept into three articles is not merely a practical choice for the sake of brevity. It reflects the fact that, contrary to common assumption, these three subjects do not sit at the same conceptual level and cannot be treated as three parallel pillars of a single thing.

This split also has historical grounding. The concept of "critical thinking" in its classical sense, a tradition running from Dewey to the present, has always been diffuse and varied, generally involving the systematic assessment of claims, evidence, sources, method and context. Later, the specialist literature in this field, notably from Facione's Delphi Report (1990) through to Ennis, Halpern, and Paul and Elder (Ennis, 1991; Halpern, 1998; Paul & Elder, 2006), divided the concept into two, or in some accounts three, main components.

"Skills", comprising interpretation, analysis, evaluation, inference, explanation and self-regulation; and "dispositions", comprising inclinations such as truth-seeking, open-mindedness and systematicity. Sometimes a third component is added: the ability to transfer a skill to new contexts.

On this division, "fallacies" and "cognitive biases" are not parallel components of critical thinking. They are two content domains to which critical-thinking skills are applied. The first comes from the tradition of logic, the second from the tradition of cognitive psychology. Critical thinking functions here as a methodological umbrella that tries to cover and organise both domains, logic and cognitive psychology, or skill plus disposition.

For this reason, the three articles in this series do not examine three equivalent things. They cover three different facets of a single problem, as follows.

This article, on fallacies, and the next, on biases, show that even where the skill of recognising an error pattern exists, that skill alone is not enough to counter the ignorance of our age.

The third article addresses something the specialist literature treats as a necessary condition for any critical-thinking skill to work: "the disposition to apply skills impartially" (Facione, 1990). Later, in Section 3 of this article, in discussing Kahan's motivated numeracy effect, we will see how having a skill without the disposition to use it impartially can produce the opposite of the intended effect.

In this trilogy, then, we address the three stated components on this basis. At this point, a question may arise: what value does addressing this topic have in our current, complex world?

In answer, and to show why these topics matter, we have evidence at two levels.

At the public level: in recent years, topics related to critical thinking, and in one specific case what is called "misinformation", appear to have gained considerable public importance, sometimes becoming a concern for a large part of society.

For example, in a Pew Research Center survey of over six thousand American adults, half of respondents rated "made-up news" as a "very big problem" for their country, larger than terrorism, climate change or illegal immigration, and 68 per cent said it had severely weakened their trust in government institutions (Pew Research Center, 2019).

We also see that, in recent years, whenever the media and publishing world offers content on "spotting faulty reasoning" and the like, demand is immediate and widespread. For example, the 180-seat capacity of the "Calling Bullshit" course at the University of Washington (on spotting misleading arguments and data) filled in under a minute, tens of thousands of visitors reached its website in the first few days, and it eventually became a bestselling book (Bergstrom & West, 2020).

At the institutional level, the American platform Checkology (which teaches recognition of these same patterns alongside other news-literacy skills) had, by 2025, enrolled close to 50,000 teachers and over 525,000 students across all 50 US states and more than 110 other countries (News Literacy Project, 2025).

The phenomenon is not confined to these two countries, however, and has a far wider reach. For example, a mapping by the European Audiovisual Observatory of 547 media-literacy projects across 28 EU countries in 2016 found that "critical thinking" was the most common category among them (European Audiovisual Observatory, 2016), and UNESCO has published a global curriculum titled "Think Critically, Click Wisely!" (UNESCO, 2021). It should be stressed, though, that a precise assessment of how widespread interest in this genre is in other parts of the world would need separate study.

As the first instalment of the "Illusion of Sufficiency" trilogy, then, and among the three categories set out above in the words of scholars in this field, this article addresses fallacies: the content domain of logic to which critical-thinking skills are applied, not one of those skills itself.

On public demand specifically for the topic of fallacies, we can likewise see evidence at two independent levels.

At the level of gamified interventions: the Bad News game (Roozenbeek and van der Linden), which teaches a list-like structure of misinformation tactics, has been translated into 15 languages and played by over a million people (Roozenbeek & van der Linden, 2019).

At the level of viral spread: the well-known poster from YourLogicalFallacyIs.com, by Jesse Richardson, published under an open licence since 2012 and reshared millions of times. It lists twenty-four common fallacies and introduces itself with the line: "Politicians, the media and others use these cunningly to fool people. Don't be fooled" (Richardson, 2012).

Another example from the same genre is David McCandless's colourful chart in Information Is Beautiful, which lists more than fifty fallacies in six categories and has been translated into several languages, including French, German and Spanish (McCandless, 2012). This chart, too, has not escaped philosophical and logical criticism. Richard Zach, a logician at the University of Calgary, has shown that its definition of "begging the question" does not match the standard philosophical meaning (Weinberg, 2014).

Taken together, all these charts and posters make the same implicit promise: know these patterns and you will be immune to deception. This article goes on to challenge that promise.

Before proceeding to the argument, though, we must first separate three things that public discourse often conflates.

1. Popular fallacy-spotting, based on list-based knowledge of fallacies: this is what the charts and posters above typically show, and it amounts to memorising the name and definition of a few common error patterns, without offering any tool for assessing premises or sources.

2. Scepticism as identity: when distrust of science and knowledge itself becomes a marker of a person's intelligence and independence of mind.

3. Measurable inability to tell truth from falsehood: a person's actual capacity to discriminate true claims from false ones, which can only be measured by fact-checking, and is not necessarily correlated with their level of distrust. In other words, a person can be highly distrustful (item 2) while remaining poor at discriminating truth from falsehood.

This article does not claim that list-based knowledge of fallacies (item 1) is useless or wrong. Its target is the assumption that this knowledge is, on its own, sufficient. That assumption of sufficiency is what gives ground to item 2, scepticism as identity, and item 3 is the measurable empirical consequence of that same false assumption.

The article is accordingly structured as follows.

Section 1 shows why list-based knowledge of fallacies is not, on its own, sufficient.

Section 2 reviews existing empirical evidence on the effect of this training on the actual ability to discriminate true from false, and reports openly on the active dispute between two independent lines of research.

Section 3 examines a separate cost imposed on impartial skill when the label "fallacy" turns into a tool of identity-labelling.

Section 4 shows why platforms' reward structures reinforce the list-like, simplified form this training takes.
The article closes with a conclusion and a list of references.

---

1. Why This Tool Is Not Sufficient on Its Own

The introduction showed that this article does not regard list-based knowledge of fallacies as useless or wrong. What it questions is the assumption of sufficiency. This section tests that assumption in four steps.

First, we see that list-based knowledge of fallacies has never reached a closed, final classification.

Then, a step deeper, we see why: the kind of coverage this knowledge offers is closed by nature, not open.

Next, we will see that even within that closed coverage, spotting a fallacy is not always possible from reading the text alone.

Finally, we see how this same structural gap can become a bottleneck that intensifies manufactured ignorance.

1.1 A List That Was Never Closed

Different traditions have, over time, produced different counts of the number of fallacies and have never reached agreement. This disagreement, however, has roots far older than the internet age.

The first person to produce a list of fallacies was Aristotle. In Sophistical Refutations, he listed thirteen, believing the list complete. Yet that claim to completeness was in doubt from the very start.

Close to two thousand years later, Augustus De Morgan, one of the founders of modern logic, wrote in Formal Logic (1847): "There is no such thing as a classification of the ways in which men may arrive at error; it is to be doubted whether there ever can be" (quoted in Hamblin, 1970, ch. 1).

Half a century later, H. W. B. Joseph, in An Introduction to Logic (1906), put this more precisely: "Truth may have a standard, but error is infinite in its deviations, and can never be reduced to a system" (quoted in Hamblin, 1970).

Charles L. Hamblin, the Australian philosopher and logician, cites both quotations in his influential book Fallacies (1970) as evidence of a long-standing scepticism within the tradition of logic itself, not a marginal objection.

This deep-rooted scepticism has found its way into contemporary literature too. Igor Žagar, in a section titled "The Overabundance and Repetition of Fallacies", shows that the number of identified fallacies has grown and overlapped to the point where they can no longer be coherently distinguished from one another (Žagar, 2021). The Stanford Encyclopedia of Philosophy, in its assessment of informal logic, likewise states that "there is no agreed-upon taxonomy of fallacies" (Groarke, 2021).

In line with this, the number of fallacies on any given list has never settled: Richardson's poster lists 24, McCandless's chart 50 across six categories, Williamson's master list 146 (Williamson, n.d.), and Bennett's compilation, the most comprehensive of today's versions, over 300 (Bennett, 2012). From Aristotle's thirteen to Bennett's three hundred and more, the list has kept growing, and none of these versions has been empirically confirmed to show that knowing exactly that set reduces the rate at which people are deceived.

This disagreement has not escaped scholarly attention, but the usual explanation for it in contemporary literature is not a formal proof. It lies in the structure of the concept itself. The common view is that individual fallacies are identified and named not by any single shared property but by family resemblance (Wittgenstein, 1953), and a concept built on family resemblance cannot, by definition, have a closed, final boundary.

At this point, two claims must be kept separate, so that the article does not go beyond what is documented.

What is documented is that the list of fallacies, from Aristotle to today, has never closed, and every generation has added new items to it.

What is not documented, and this article makes no such claim, is a far stronger proposition: that because of the nature of natural language, one can formally prove this list is infinite. The exception is formal fallacies, where infinity really is a mathematical fact, since invalid combinations in formal logic are uncountable. But the subject of this article is informal fallacies, for which no such proof exists.

A distinction must therefore be drawn between two claims: that the list, in practice and to date, has never been completed (which is documented), and that one can prove the list can logically never be completed (a far stronger claim, for which no proof exists). Conflating the two is itself an example of the unsupported leap this article criticises.

1.2 Unlimited Coverage versus Closed Coverage

The previous section showed that the list of fallacies has never closed. But simply not knowing why the list stays open does not yet explain why this matters to the ordinary user. To see why it matters, we need to compare the kind of knowledge this list offers with another kind of knowledge, and the best place to start is a metaphor commonly used in the misinformation literature.

On this metaphor, public education is a kind of "mental vaccination". Just as a vaccine immunises the body against a virus, familiarity with deception techniques immunises the mind against being misled. This metaphor is so entrenched that it has given rise to a whole research subfield called "inoculation theory".

But researchers who have revisited this field with closer methodological scrutiny have raised serious objections to the metaphor. Ariana Modirrousta-Galian, Philip Higham and Tina Seabrooke, in a piece bluntly titled "Comparing Misinformation to a Virus Is Neither Accurate Nor Useful in Preventing Its Spread", published in 2024 in the British magazine The Skeptic, argue that the comparison is misleading: a virus has a single, uniform biological mechanism, which allows a vaccine to act generically, regardless of the specific exposure, whereas misinformation has no such uniform mechanism (Modirrousta-Galian, Higham, & Seabrooke, 2024).

This point can be put more precisely, taking it further than the authors' own critique. Hand washing works because it has a generic mechanism: soap and water break down the coat proteins of a wide range of pathogens, regardless of which specific virus or bacterium is involved. You do not need to know which pathogen you face in order to wash your hands. One action covers a wide range, and its protocol, duration, type of substance, has been fixed through clinical trials and measurable reductions in infection rates.

This coverage is not complete, however: pathogens without a lipid envelope, such as some enteric viruses, or bacterial spores, are removed mainly by physical dislodgement rather than chemical breakdown of coat proteins. So the claim that a single chemical mechanism covers everything is not strictly true. Hand washing draws on two different mechanisms, chemical breakdown or physical removal, depending on the pathogen. But the point that matters for this article's argument still holds.

In both cases, the effect depends on the physical mechanism of the pathogen itself, not on whether the user already knew its name.

Knowledge of fallacies, by contrast, is pattern-based: to recognise a fallacy, you must already have learned that specific pattern and have it in mind. This is like a narrow-spectrum antibiotic, not hand washing: you must already know exactly which agent you are facing. And, as we saw in 1.1, the list of these "agents" has been drawn up differently by different traditions, which disagree with one another and have never reached consensus.

The core problem, then, lies neither in the endlessness of the fallacy list nor in pattern-based knowledge being inherently the wrong kind of knowledge, since a narrow-spectrum antibiotic is also pattern-based and highly useful when the pathogen is common. The real difference lies in the type of coverage. Once the generic biological mechanism is learned, germ theory, hand washing covers an unlimited and even still-unknown range of pathogens, because its effect depends on the physical structure of the agent, not on whether its name was already known.

The list of fallacies has no such coverage, and this is exactly what we saw in 1.1: every new pattern of misleading that is not already on the list must be separately identified and added to it, not because its compilers were careless, but because this kind of knowledge is, by its nature, as described here. New deception strategies slip past this shield, not because the shield is small, but because its coverage is closed and additive: every new strategy remains invisible to the shield until it is separately identified and added to the list.

1.3 Form versus Premises

Popular fallacy-spotting, as we saw in the previous section, confines itself to recognising the surface form of an argument. But formal correctness does not guarantee the correctness of the content. An argument can be formally flawless and still rest on premises that are incomplete, selectively chosen, or stripped of context.

Informal logic, from Hamblin (1970) to Walton (2008), has criticised exactly this limitation from the start. Based on these critiques, it is wrong to claim that fallacy-spotting has historically only ever assessed an argument's formal validity. What is true and defensible is something different: specialist training in informal logic does assess premises, sources and context. In the popular product, the kind of chart seen in the introduction, these layers are stripped away and reduced to recognising verbal labels. The core problem here is the gap between the two.

A concrete example shows this gap: a company states that "our trial was statistically significant." The form of the argument is flawless; no formal fallacy can be found in it. But the company has hidden nine failed trials and published only the one that succeeded, a case of publication bias repeatedly documented in pharmaceutical and nutrition research.

In such cases, the popular fallacy-spotter, seeing only the argument's surface form and finding it correct, is deceived, whereas spotting this kind of deception requires access to the complete data, not knowledge of an argument pattern's name.

A finer point must be added here, since one could easily object that biased selection, or cherry-picking and suppressed evidence, do themselves appear on many of these same fallacy lists. In the FLICC taxonomy, built to teach how to counter science denial, cherry-picking is one of five main pillars (Cook, 2020). So it cannot be said that this strategy slips past the fallacy list, since the list does have room for it.

The problem lies elsewhere: formal fallacies, and a number of simple informal fallacies, can be spotted from the text alone; read the sentence and you see the pattern. But many other informal fallacies, such as false cause or slippery slope, need background knowledge just as cherry-picking does. Cherry-picking is only the most extreme case on this spectrum, not a wholly separate category.

In other words, to discover whether cherry-picking has occurred, you need to know what has been omitted; you need data outside the text itself, the full distribution of the evidence. Naming this strategy on a list, therefore, does not by itself generate the ability to detect it.

This is exactly the closed-coverage gap we saw in 1.2, but this time it arises not from a pattern's absence from the list but from the list's own inadequacy for items it does contain. The bottleneck is not recognising the pattern; it is access to the omitted context.

1.4 Manufactured Ignorance as an Institutional Product

Robert Proctor, in Golden Holocaust (2011, University of California Press), and Naomi Oreskes and Erik Conway, in Merchants of Doubt (2010, Bloomsbury), have documented that the tobacco industry did not rely on outright fallacy alone, though it did use that too, for instance by invoking fake experts, itself a recognised fallacy. Its central and more durable strategy was something else: commissioning real research, selecting from among genuine results, and exaggerating science's natural uncertainties, the same strategy Oreskes and Conway have traced, beyond tobacco, in the cases of acid rain, the ozone layer and climate-change denial. The internal slogan of the Brown and Williamson Tobacco Corporation, recorded in a 1969 internal memo, exposed during US litigation and quoted in both books, sums this up in one sentence: "Doubt is our product" (Brown & Williamson Tobacco Corporation, 1969).

The important point here is that this strategy recurs across all four phenomena this article addresses.

Commercial false advertising draws on this same logic of data selection.

Political extremism is reinforced by the selective highlighting of real events, not by fabricating them.

Factionalism is stoked by repeating statistics that are accurate but stripped of context.

And corporate ignorance, as we saw in the tobacco case, works on exactly this same pattern.

One can say, then, that what chiefly distinguishes the ignorance of our age from the ordinary, accidental ignorance of individuals is not the type of error but the existence of actors with the resources and organised motive to produce and entrench it, or, as we will see in Section 4, structures such as platforms' reward systems that reproduce the same effect even without anyone's conscious intent. In all four cases, what needs to be identified is not simply the breach of a logical rule; it is the absence of data, context or a source that never appears in the argument's surface form.

At this point, before turning to the empirical evidence in the next section, one methodological point must also be made explicit.

What follows is not directly about the classical fallacy lists described in the introduction, but mainly about similar educational games that teach users a limited set of predefined misleading patterns, impersonation, polarisation, discrediting, and the like. Although these games and fallacy lists differ in content, they share one structural feature: both are closed, pattern-based knowledge. The user must recognise a specific type of trick that they have already learned from a list. For this reason, evidence on the effectiveness of these games also speaks to the whole family of pattern-based interventions, not because the two are identical, but because both fall under the same strategy criticised in 1.2. Beyond this, more direct evidence exists specifically on fallacy training itself, and this follows in the next section.

---

2. A Field Under Dispute: Evidence on Discrimination and Response Bias

The previous section set out the argument at the conceptual and structural level: why list-based knowledge has this kind of coverage, independent of any data. This section now moves from that conceptual level to the empirical evidence, research that has directly tested whether this training actually improves the ability to discriminate true from false, or merely shifts a person's threshold of distrust, and shows that this question, contrary to common assumption, is still the subject of active scientific dispute.

2.1 Response Bias versus Discrimination: The Methodological Critique by Modirrousta-Galian and Higham

One of the most serious lines of research here comes from a reanalysis Modirrousta-Galian and Higham (2023) published in the Journal of Experimental Psychology: General (Modirrousta-Galian & Higham, 2023a). They reanalysed five earlier studies of "mental vaccination" games such as Bad News and Go Viral! using ROC curve analysis, a standard tool in signal detection theory that measures discrimination ability independently of response bias.

The distinction between these two concepts, though it may look technical at first glance, is critical to understanding the finding. Earlier studies reported only the correct-detection rate for fake news, the percentage of fake stories a participant correctly identified, while ignoring any change in the rate of misjudging true news as false. The problem is that these two figures move together: if someone simply becomes more suspicious of everything after playing the game, whether a story is fake or real, their correct-detection rate for fake news rises without their actual ability to

discriminate true from false having improved; they have only become more cautious, not sharper. Telling these two states apart is exactly what ROC analysis does.

When Modirrousta-Galian and Higham calculated both rates together, the key finding was this: these interventions do not improve the ability to discriminate true news from false; what improves is response bias, a person's general tendency to call everything "false", including true news. A later, independent replication using Indian data (Seabrooke, Modirrousta-Galian, & Higham, 2025) confirmed the same pattern, finding no significant improvement in discrimination while the conservative response bias persisted.

2.2 A Reverse Reading and an Active Dispute

This finding, contrary to the impression a single peer-reviewed article usually gives, is not the end of the matter. Simchon, Zipori, Teitelbaum, Lewandowsky and van der Linden (2025, Current Opinion in Psychology), using the same signal detection framework but at a far larger scale, a reanalysis of 33 inoculation experiments with over 37,000 participants in total, using a hierarchical Bayesian model, reached the opposite conclusion: both game-based and video-based interventions, they claim, genuinely improve discrimination without increasing overall response bias.

This study, though, must be read in context: Sander van der Linden, one of its two senior authors and director of Cambridge University's Social Decision-Making Lab, is, together with Jon Roozenbeek, the original designer of the Bad News game. In this reanalysis, then, the intervention's own creator is judging the effectiveness of his own creation. This fact does not, by itself, invalidate the new finding, but it clarifies what we are looking at: not a neutral discovery that has overturned the earlier view, but a live dispute between two research lines with different stakes and interests in the question.

This article takes no position on which of the two reanalyses is correct, not out of neutrality for its own sake, but because, as set out in 1.2 and 1.3, its central argument does not depend on adjudicating between them. Even if the more optimistic assumption is right, and these particular games really do improve discrimination, the structural limitation set out in 1.2 still holds: any such improvement comes through training people to recognise that game's own predefined "badges", not through a generic mechanism like hand washing, and any new deception strategy that does not reproduce these badges will still slip past this closed-coverage shield. In other words, the dispute above concerns the size of one particular intervention's effect, not the kind of coverage this article is criticising.

2.3 Direct Evidence on Fallacy Training: A Two-Sided Finding

So far, the evidence concerned mental-vaccination games, not direct fallacy training. The only randomised trial to test this directly is Hruschka and Appel (2023, PLOS ONE): with 116 participants, a thirty-minute text-based intervention on informal fallacies improved the ability to spot fallacious arguments a week later, and this ability correlated with better discrimination of real from fake news. But the fallacy-trained group performed no better on the news-discrimination task itself than the control group, which had only been trained about fake news. This evidence cuts both ways: it shows that name-based fallacy knowledge can transfer and persist, and it shows that this knowledge, on its own, held no advantage over an alternative, non-list-based form of training. Both halves of this finding should be kept in view.

Further evidence, directly relevant to the explicit/implicit distinction in 1.2, comes from the same critical line of research.

Modirrousta-Galian, Seabrooke, Hanoch, Kelley and Higham (2026, Journal of Experimental Psychology: Applied), building on their earlier work showing inductive learning outperforms Bad News (Modirrousta-Galian & Higham, 2023b), tested in three preregistered experiments (N = 1,135 in total) a method in which participants simply see true and false news items and receive immediate feedback, with no explanation or naming of any pattern: exactly the opposite of the narrow-spectrum antibiotic criticised in 1.2.

The result, though, was not straightforward. In Experiment 1, this method improved discrimination. In Experiment 2, adding reward badges unexpectedly reduced the effect, and in Experiment 3, changing the order of practice, improvement reappeared and was stronger, though it is unclear whether the badges had been removed in this experiment, a point that would need checking against the full text. With that caveat, the finding reinforces the point made in 1.2: explicit, named knowledge is, at least here, part of the problem, not only the solution.

2.4 The Strongest Evidence: Field Data, Not Laboratory Data

The strongest field evidence, as opposed to laboratory evidence, points the other way. In 2025, Nygren and Efimova (2025, PLOS ONE), in a study of 459 Swedish upper-secondary students in real classrooms, neither the Bad News game nor the News Evaluator workshop significantly improved students' ability to assess credible and misleading news; three months on, students also reported no greater use of fact-checking tools. Because this study took place in a real classroom setting over a three-month horizon, it is far stronger evidence for the sceptical half of this article's argument than the laboratory-based ROC reanalyses.

Even so, if we take the position of the critical research line, Modirrousta-Galian, Higham and Seabrooke, with the same caution noted above, as a description of the current state of mainstream training, the implication is this.

Public education, in its current dominant form, probably shifts a person's threshold of distrust rather than their power of discrimination, and pervasive distrust without discrimination is exactly the environment the doubt industry needs. In such an environment, dismissing any inconvenient claim, including valid scientific claims and even official public-health warnings, becomes as easy as dismissing a false one. This is precisely the outcome the doubt industry, in Proctor and Oreskes's terms, has always been after.

Not that people accept false beliefs, but that they trust no belief at all, including true ones. Given the dispute above, though, this scenario should be read as a serious possibility still under investigation, not as a settled conclusion.

For transparency about the research underpinning this section's argument, it should also be added that the evidence on the relationship between analytic ability and impartiality is not uniform. Pennycook and Rand (2019), for instance, showed that in the task of assessing news accuracy, a greater inclination towards analytic thinking goes with better discrimination, not greater polarisation, a finding in tension with what we will see in Section 3, where Kahan and colleagues (Kahan et al., 2017) showed that higher cognitive ability, on issues bound up with identity, does not reduce polarisation but increases it.

These two findings do not invalidate one another. In fact, the effect of analytic ability appears to depend on the type of task rather than following one uniform rule everywhere, a point that narrows and sharpens the scope of this article's thesis.

But even if this sceptical scenario turns out to be wrong and public education genuinely does improve discrimination, a separate question remains: what happens, in practice, to the very label this training gives people, the word "fallacy"?
The next section seeks to answer exactly this question.

---

3. The Label "Fallacy" as a Weapon of Identity

This section is of a different kind: not a link in the inferential chain about effectiveness, but about a separate political cost that is added on even if this training proves partly effective. From a tool's ineffectiveness one cannot infer that, in practice, it does not become a weapon against an ideological opponent. These are independent claims.

Taber and Lodge (2006), in two experiments on evaluating political arguments, affirmative action and gun control, showed that people rate arguments aligned with their prior attitudes as stronger and produce more counter-arguments against opposing ones; the result is increasing polarisation, especially among people with greater political knowledge. In 2017, Kahan and colleagues (Kahan et al., 2017), in an experiment on interpreting statistical data about gun control, showed that people with higher numeracy, on issues bound up with identity, not only fail to become less polarised, as one might expect, but become more polarised still. In other words, greater cognitive ability is put to work defending one's group position rather than impartial truth-seeking.

This finding, though, must be reported with a clear caveat: an attempt by Ballarini and Sloman (2017) to replicate Kahan's motivated numeracy effect did not succeed, while a further replication by Connor, Sullivan, Alfano and Tintarev (2024), using a European sample, did reproduce the effect but found no evidence that people with higher numeracy became more polarised. This methodological doubt must be stated openly and left in the text, and this article's claim does not depend on the full strength of this one study.

Setting aside the contested details just noted, what this literature suggests overall is that, in social use, the label "fallacy" is often not a tool of evaluation but a tool for dismissing an opponent.

This weaponisation complements the finding of Section 2 (particularly 2.4): if public education does not raise discrimination power but does raise the threshold of distrust, we must now add that once that threshold has risen, it may be used as a weapon in the hands of someone with an identity-driven motive, against any source incompatible with that motive, rather than on the basis of the source's actual quality.

Section 2, then, showed this wave is empirically contested, and Section 3 has shown that, even setting that dispute aside, the label "fallacy" carries a separate political cost. A further question must now be answered: why did this wave, in precisely its list-like, simplified form, become so widespread in the first place? And does explaining that spread itself give any reason to think its content is wrong?
This is the subject of the next section.

---

4. The Attention Economy

The previous section showed how the label "fallacy" becomes a weapon in the political arena. One question remains unanswered, though: of all the possible forms education could take, why did this wave end up exactly in its current list-like, poster-friendly, shareable form?

A simple claim sometimes made is that, because this wave is a product of the attention economy, it must therefore be ineffective or wrong. That argument is itself a genetic fallacy, and committing it in an article about fallacies would not be acceptable.

The correct formulation, in this article's view, is that the attention economy explains why this wave took its list-like, simplified form, a structural mechanism that complements, rather than competes with, the knowing actors of 1.4, not that its content is wrong.

The platform's reward structure, not the educational value of either option, decides the algorithmic choice between "24 fallacies you need to know" (shareable, fast, visual, a colourful chart understood at a glance) and "how to trace the source of a claim" (slower, requiring hands-on practice, with no immediate visual payoff). The skill that actually works is often tedious and visually unappealing; the skill that spreads online is often whichever one raises the share rate, not the one that raises effectiveness. These two paths have entirely different criteria for success, and the platform rewards only one of them.

This explanation, by itself, delivers no verdict on effectiveness; it only shows why, in the attention market, the ineffective form always outpaces the effective one, and for exactly this reason, the training that actually works is usually not the one getting the most views. That obscurity is no evidence of its ineffectiveness.

Whether this same attention-market logic repeats in the domain of cognitive biases, whether the name of a bias, for the same reason, is likewise easier to state and share than the skill of spotting it in the moment of action, is a question the second article in this trilogy will take up.

Conclusion

This article, the first in the "Illusion of Sufficiency" trilogy, tested only one content domain: fallacies, which come from the tradition of logic.

The claim tested throughout was always a limited one: not that this knowledge is useless, but that the skill of recognising an error pattern, even where it exists, is not on its own enough to counter the ignorance of our age. This article's argument is, in the end, mechanistic rather than empirical. Knowledge of fallacies has closed, additive coverage, not the unlimited coverage of a generic mechanism, and even where the list does name a deception strategy, spotting it requires access to the omitted context, not merely recognition of the pattern. For this reason, neither the scientific dispute in Section 2 nor any new finding that emerges hereafter can overturn this argument, since the argument never depended, from the outset, on the result of any single experiment.

Two further costs, independent of this mechanism, were added to it: the label "fallacy" can turn from a tool of evaluation into a tool for dismissing an opponent, and the list-like, poster-friendly form this knowledge has taken is a product of the attention market, not of its educational value.

The real boundary, then, is not between critical thinking and its uselessness, but between a skill that traces a claim back to its source and the mechanism that produced it, and a skill that merely recognises the name of a familiar pattern, without sufficient practice or feedback.

Someone who has memorised the list of fallacies but does not know how to trace a claim back to its source is like a sailor holding only a compass, with neither a helm nor a chart. The compass is not useless; it can show direction. But on its own, it will not carry anyone from the open ocean to the shore.

References

Ballarini, C., & Sloman, S. (2017). Reasons and the "motivated numeracy effect." Proceedings of the 39th Annual Meeting of the Cognitive Science Society.

Bennett, B. (2012). Logically fallacious: The ultimate collection of over 300 logical fallacies. eBookIt.com.

Bergstrom, C. T., & West, J. D. (2020). Calling bullshit: The art of scepticism in a data-driven world. Random House.

Brown & Williamson Tobacco Corporation. (1969). Smoking and health proposal [Internal memorandum, Bates no. 680561778–1786]. UCSF Industry Documents Library. https://www.industrydocuments.ucsf.edu/tobacco/

Connor, P., Sullivan, E., Alfano, M., & Tintarev, N. (2024). Motivated numeracy and active reasoning in a Western European sample. Behavioural Public Policy, 8(1), 24–46.

De Morgan, A. (1847). Formal logic: Or, the calculus of inference, necessary and probable (p. 237). Taylor & Walton. (As cited in Hamblin, 1970)

Ennis, R. H. (1991). Critical thinking: A streamlined conception. Teaching Philosophy, 14(1), 5–24. https://doi.org/10.5840/teachphil19911412

European Audiovisual Observatory. (2016). Mapping of media literacy practices and actions in EU-28. Council of Europe.

Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report). American Philosophical Association. https://doi.org/10.1037/e506962012-001

Groarke, L. (2021). Informal logic. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/logic-informal/

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. American Psychologist, 53(4), 449–455. https://doi.org/10.1037/0003-066X.53.4.449

Hamblin, C. L. (1970). Fallacies. Methuen.

Hruschka, T. M. J., & Appel, M. (2023). Learning about informal fallacies and the detection of fake news: An experimental intervention. PLOS ONE, 18(3), e0283238. https://doi.org/10.1371/journal.pone.0283238

Joseph, H. W. B. (1906). An introduction to logic (p. 569). Clarendon Press. (As cited in Hamblin, 1970)

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. Behavioural Public Policy, 1(1), 54–86. https://doi.org/10.1017/bpp.2016.2

McCandless, D. (2012). Rhetological fallacies [Infographic]. Information Is Beautiful. https://informationisbeautiful.net/visualizations/rhetological-fallacies/

Modirrousta-Galian, A., & Higham, P. A. (2023a). Gamified inoculation interventions do not improve discrimination between true and fake news: Reanalyzing existing research with receiver

operating characteristic analysis. Journal of Experimental Psychology: General, 152(9), 2411–2437. https://doi.org/10.1037/xge0001395

Modirrousta-Galian, A., & Higham, P. A. (2023b). Effects of inductive learning and gamification on news veracity discernment. Journal of Experimental Psychology: Applied, 29(3), 599–619. https://doi.org/10.1037/xap0000458

Modirrousta-Galian, A., Higham, P. A., & Seabrooke, T. (2024, September 1). Comparing misinformation to a virus is neither accurate nor useful in preventing its spread. The Skeptic. https://www.skeptic.org.uk/2024/09/comparing-misinformation-to-a-virus-is-neither-accurate-nor-useful-in-preventing-its-spread/

Modirrousta-Galian, A., Seabrooke, T., Hanoch, Y., Kelley, N. J., & Higham, P. A. (2026). An inductive learning intervention to improve news veracity discernment. Journal of Experimental Psychology: Applied. Advance online publication. https://doi.org/10.1037/xap0000580

News Literacy Project. (2025, November). Checkology: Educator booklet [Brochure]. https://newslit.org

Nygren, T., & Efimova, E. (2025). Investigating the long-term impact of misinformation interventions in upper secondary education. PLOS ONE, 20(7), e0326928. https://doi.org/10.1371/journal.pone.0326928

Oreskes, N., & Conway, E. M. (2010). Merchants of doubt: How a handful of scientists obscured the truth on issues from tobacco smoke to global warming. Bloomsbury Press.

Paul, R., & Elder, L. (2006). The miniature guide to critical thinking: Concepts and tools. Foundation for Critical Thinking.

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. Cognition, 188, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

Pew Research Center. (2019, June 5). Many Americans say made-up news is a critical problem that needs to be fixed. https://www.pewresearch.org/journalism/2019/06/05/many-americans-say-made-up-news-is-a-critical-problem-that-needs-to-be-fixed/

Proctor, R. N. (2011). Golden holocaust: Origins of the cigarette catastrophe and the case for abolition. University of California Press.

Richardson, J. (2012). Thou shalt not commit logical fallacies [Poster]. YourLogicalFallacyIs.com. https://yourlogicalfallacyis.com/pdf/FallaciesPoster24x36.pdf

Roozenbeek, J., & van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. Palgrave Communications, 5, Article 65. https://doi.org/10.1057/s41599-019-0279-9

Seabrooke, T., Modirrousta-Galian, A., & Higham, P. A. (2025). Re-examining the Bad News game: No evidence of improved discrimination of Indian true and fake news headlines. Psychonomic Bulletin & Review. https://doi.org/10.3758/s13423-025-02827-x

Simchon, A., Zipori, T., Teitelbaum, L., Lewandowsky, S., & van der Linden, S. (2025). A signal detection theory meta-analysis of psychological inoculation against misinformation. Current Opinion in Psychology, 67, Article 102194. https://doi.org/10.1016/j.copsyc.2025.102194

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. American Journal of Political Science, 50(3), 755–769. https://doi.org/10.1111/j.1540-5907.2006.00214.x

UNESCO. (2021). Media and information literate citizens: Think critically, click wisely! UNESCO.

Walton, D. (2008). Informal logic: A pragmatic approach (2nd ed.). Cambridge University Press.

Weinberg, J. (2014, December 12). Colorful fallacies chart. Daily Nous. https://dailynous.com/2014/12/12/colorful-fallacies-chart/

Williamson, O. M. (n.d.). Master list of logical fallacies. University of Texas at El Paso. https://utminers.utep.edu/omwilliamson/engl1311/fallacies.htm

Wittgenstein, L. (1953). Philosophical investigations (G. E. M. Anscombe, Trans.). Blackwell.

Žagar, I. Ž. (2021). Four critical essays on argumentation. Pedagoški inštitut.
