



Confirmation bias studies: towards a scientific theory in the humanities

Thomas Rist¹ 

Received: 18 October 2021 / Accepted: 1 June 2023
© The Author(s) 2023

Abstract

This article argues that a global crisis of interpretation can and should be confronted by humanities programmes in UK and similar universities. It contends that raising the standards of proof for theoretical models of interpretation in the humanities will help reverse erosions of trust undermining democratic institutions and expertise. To this end, it considers how financial challenges facing UK universities mould the teaching of theory in the humanities and the knowledge this teaching gives rise to. The article considers how standards of proof from the social sciences can interrogate theory in these conditions, developing it and increasing its assurance. The essay illustrates this claim through a series of sample theories and literary works: Roland Barthes' 'Death of the Author', the Orientalism derived of Edward Said, Chinua Achebe's novel *Things Fall Apart* and Jung Chang's *Wild Swans: Three Daughters of China*. From these examples the essay draws larger conclusions, the biggest of which is a new subfield of study for the humanities: Confirmation Bias Studies. The article is structured as follows. In 'Part 1: Introduction', the article considers the openness of literary theory to confirmation bias, which is considered historically, in cognitive processes, especially those processes in the age of the internet, and in educational processes in the humanities. In 'Part 2: Procedure', the article explores the challenges of applying a confirmation bias approach to literary theory as a means of interpretation. In 'Part 3: Conclusion', the article summarises the key strategies for overcoming confirmation bias in theoretical approaches to the humanities discovered in the article.

Keywords Humanities · Theory · Confirmation Bias · Public Discourse · Truth · Method

✉ Thomas Rist
t.rist@abdn.ac.uk

¹ Department of English, King's College, University of Aberdeen, Aberdeen, UK

Introduction

Confirmation bias studies does not yet exist in the humanities, but in theory it should. Theory, especially literary theory, is an evolving discipline, defined by its willingness to take on new thoughts, and so, paradoxically, by its lack of definition. To earlier theories like structuralism, poststructuralism and feminism, more recently have been added postmodernism, postcolonialism, gender studies, trauma theory and queer theory among other theoretical models. More theory—a lot more of it—seems likely to follow. In one definition, literary theory is ‘increasingly hybrid’, with an ‘endlessly fecundating terminology’.¹ A consequence is the ‘resurgence of scepticism about the possibility of verifiable knowledge’, to which considerations of confirmation bias provide a response.²

There is a place, therefore, for study of confirmation bias in the theoretical humanities. Yet presently this study exists on no humanities curriculum, and it is only here, in this article, that the study receives a name. Nevertheless, ‘Confirmation Bias Studies’ has been stirring for aeons. It begins with the ancients, with the contention of Plato, Aristotle and others that a world of fact exists that is independent of our ideas and judgements, a contention implying some ideas are correct while others are not.³ For some, the debate reached its zenith in the ‘fake news’ crisis of the Trump presidency, in which we learnt of ‘alternative facts’—a crisis to which, in his inauguration speech, President Biden replied directly: ‘There is truth, and there are lies—lies told for power, and for profit ... And each of us has a duty and a responsibility as citizens ... to defend the truth, to defeat the lies’.⁴ Yet one would be an optimist to think a contention that began before Plato will end with Biden. In the internet’s global village, Plato’s old enemy, opinion, is alive and kicking. If you ‘like’ an idea, and then others like it, before long it’s around the world, apparently as good as true.

An aspect of ‘likes’ of the kind is that they please. One ‘likes’ it because one likes it, not because one considers it to have a basis in fact or to be more than an opinion.⁵ The criteria on which public discourses of this kind are judged is desire. *I desire to believe this*. Or, *I desire to please this friend*. Or, simply, *I desire to have a say in the public discourse*. With a public discourse based on these criteria, anything and everything becomes say-able; and if one searches far enough, on the net, one finds this everything and anything. It behoves the academy to do better.

This is where confirmation bias theory, from departments of Psychology, comes in. Defined as ‘the seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand’, confirmation bias has hyperbolically been called ‘the single problematic aspect of human reasoning that

¹ See Wolfreys et al. (2014, p. x).

² See McSweeney, ‘Part 2’ (2021b, p. 842).

³ On this history, see Rorty (2009); for an influential modern defence of this philosophical position, see Nagel (1997).

⁴ On ‘fake news’ and the “‘post-truth’/‘post factual’” politics characterised by the diminishing importance of anchoring utterances in verifiable facts’, see McSweeney, ‘Part 1’ (2021a, p. 1065).

⁵ See Dhir et al. (2019, pp. 544–545), Lesley et al. (2017, p. 153).

deserves attention above all others'.⁶ Confirmation bias has been identified in fields as diverse as strategy formation, enhanced commitment to unprofitable courses of action, risk assessment, software testing, criminal investigations, medical diagnoses, judicial decisions, bullet comparisons, skull sexing, animal behaviour research, psychiatric confinement examinations and visual perception.⁷ Closer to home, it is found in political debate, academic discussion and, most importantly, literary criticism and interpretation.⁸ This focus of confirmation bias theory on interpretation gives it an important place in literary theory and in the humanities. Like these disciplines, confirmation bias study is concerned with interpretation, but it approaches the topic from the point of view of the prevalence of misinterpretation. At the heart of the problem is a human tendency to seek and endorse viewpoints and information confirming one's existing, subjective beliefs.⁹

The far-reaching implications for literary and wider cultural scholarship are only beginning to be understood. In 2014 Maurice Lee observed there is 'no sustained work' in literary studies on evidentiary superabundance: a crisis in literary interpretation caused not by a lack of evidence, but by its inexhaustible profusion.¹⁰ As in the wider world, the internet plays a significant role in the crisis: 'targeted searches of immense databases combined with evidentiary inclusiveness and confirmation bias allow literary critics to discover support for an amazing range of claims'.¹¹ Lee compared this to playing tennis with the net down, which invites the question: why play the game at all? Yet this fundamental challenge to the value of literary (and wider cultural) studies is as nothing when placed beside the wider crisis of truth at large. Since superabundance characterises academic and non-academic search engines alike, the problems of modern literary interpretation are the problems of our age.¹²

Two answers to this problem one encounters in cultural debates are censorship and peer-review.¹³ Peer-review, it is argued, guarantees quality interpretation because the reviewers are experts. Censorship, it is also argued, guarantees truth by weeding out falsehoods. Yet both approaches suffer from the same weakness and are open to the same challenge. Both depend on a higher-order interpreter, or 'specialised class', who may or may not be reliable.¹⁴ One hopes, of course, that an academic reviewer will be better informed and motivated than the censor for a totalitarian state, but it is not guaranteed. Academics can be partisan and—as importantly—everyone knows it.¹⁵ Especially in the humanities, too, some academics

⁶ See Nickerson (1998, p. 175).

⁷ McSweeney, 'Part 1' (2021a, p. 1065).

⁸ McSweeney, 'Part 1' (2021a, p. 1065).

⁹ McSweeney, 'Part 1' (2021a, p. 1065).

¹⁰ See Lee (2012, pp. 87–94).

¹¹ Lee (2012, p. 163).

¹² On social media creating confirmation bias today, see Lokot and Diakopolous (2016, pp. 682–699), Dubois and Blank (2018, pp. 729–745), McSweeney, 'Part 1' (2021a, p. 1065).

¹³ See Brady (2010, pp. 10–11), Unerman (2020, pp. 1–26), King et al. (2018, pp. 843–855), Nickerson (2020, pp. 224–225).

¹⁴ See Herman (2000, p. 101).

¹⁵ Unerman (2020, pp. 1–26), King et al. (2018, pp. 843–855), McSweeney, 'Part 1' (2021a, p. 1064).

reject the existence of objective truths, reasonably raising the question of why they should be trusted to review objectively. The question arises even when these academics review honestly.¹⁶

Despite President Biden's resolve to restore it in public life, trust alone is not the answer. There must be good reason to trust.¹⁷ For literary study in the age of superabundance, that reason comes from the application of confirmation bias theory. Reading literary theories for their capacity to misinterpret, rather than to interpret, provides unusual evidence of their value. The argument that a theory proves its worth by the interpretations it enables is circular.¹⁸ If the theory is wrong, the interpretation arising from it will likely contain errors. For this reason, it is argued, all literary interpretation requires faith.¹⁹ Yet testing a literary theory for confirmation bias *before* putting it to interpretative use helps validate the theory, allowing us to apply it with greater faith thereafter. Just as one wouldn't expect to take a covid vaccine without the assurance of validating procedures beforehand, so one shouldn't expect to apply a theoretical model to interpretation without a proper initial test.²⁰

It's worth emphasising here how our age of superabundance creates the need for a new standard of proof. Scientific standards of proof lean towards two critical but adverse models, those of T.S. Kuhn and Karl Popper.²¹ Kuhn noted confirmation-seeking is fundamental in scientific practice, but Popper observed the method is also likely to produce pseudoscience.²² Finding examples which confirm your hypothesis can produce vaccinations against Covid-19 and adherents to Daesh or Q-Anon. It is not enough, therefore, to seek new examples substantiating your hypothesis. For a theory to be valid, one needs to know not only how true it is, but also how false.²³

This is old hat in the social sciences, but despite their interdisciplinary turns, it is mostly absent in the theoretical humanities; and, therefore, in the bright young things it sends out to the wider world. One reason for this is the traditional aim of the humanities to explain primary sources.²⁴ Explaining a source means interpreting it, so theoretical models are valued for their capacity to interpret rather than for their capacity *not* to interpret.²⁵ The bias operates at several levels. If it's at all concerned with a theory, the one-hour undergraduate lecture must show students the theory's value. In practice, this means first explaining the theory and then applying it in interpretation, so that the undergraduate audience can see its usefulness.²⁶ That leaves little or no time to address its deficiencies; and being as interested as their teachers

¹⁶ McSweeney, 'Part 2' (2021b, p. 842).

¹⁷ On distinctions between good and bad reasoning, see Mercier (2012, pp. 244–245).

¹⁸ See Simpson (2003, pp. 215, 221).

¹⁹ Simpson (2003, p. 229).

²⁰ On the importance of scientific tests, in this respect, see Zu et al. (2022, p. 1).

²¹ For introduction, see Noturno (1984, pp. 273–289).

²² See Rosende (2009, pp. 135–154), Robergs (2017, pp. 1–11)

²³ Robergs, 4. On good and bad theory in Popper, see also Boyer (2009, p. 247).

²⁴ See Jay (2014, p. 142).

²⁵ See Jay (2014, pp. 114–142).

²⁶ See Byrne (2011, pp. 117–118).

in interpreting texts, or artworks or historical documents, students are likely to resist as obfuscation or irrelevance critiques of theories they are barely accustomed to.²⁷

At the postgraduate level, the problem is different.²⁸ Postgraduates tend to be undergraduates who rose to the top, so some capacity for critical thinking inheres. Yet as in the undergraduate scenario, where the single hour of the lecture (or tutorial) limits a theoretical enquiry, the pressure of income-streams on universities means doctoral students often lack time for their deepest research.²⁹ In UK universities graduate students (especially those paying higher fees, from abroad) are a main source of revenue. Humanities especially tend to be cash-strapped, meaning they are under-staffed and want as many doctoral students as possible.³⁰ Churning out contributions to knowledge (the definition of the PhD) becomes a desideratum, and the easiest model for it goes like this: here's a theory, here's a text (or artwork or document), here's a new interpretation.³¹ Operating under financial and temporal constraints, what this process aimed at interpretation does not do is produce many theses on misinterpretation. And since new theories encourage fresh interpretations, there is the additional pressure to side-line older theories in favour of newer ones that are less tried.³² Sniffing out misinterpretation is a part of most doctoral projects, but rarely is it the main part, not least because misinterpretation in the Humanities has come widely to be seen as 'a misnomer for an opinion delinquent enough to differ from our own'.³³ Moreover, in the humanities only the mainly theoretical or philosophical doctorate might attend to large-scale misinterpretation; yet as we have seen, there's no such thing as yet as Confirmation Bias Studies in the humanities.

Procedure

How, though, might such study work? In 2014, Lee noted literary critics weren't yet under a mandate to meet scientific standards of proof, but he saw the moment coming and as he observed, sampling is the answer.³⁴ The trick in this process is to find the universal claim in a hypothesis and then to test it for shortcomings; if counter-examples to the hypothesis can be found, then the hypothesis needs to be modified. As a simple example, consider the so-called death of the author. Clearly some authors (including the present one) are alive, ergo the theoretical hypothesis needs at least some qualification and modification. If authors aren't all dead in the

²⁷ Byrne (2011, p. 118).

²⁸ See Mattison (2012, pp. 5–10).

²⁹ See Levin (2005, pp. 17–28).

³⁰ See Collini (2012, p. 31).

³¹ My characterisation of the theoretical doctorate develops Carter's description of humanities doctorates as 'creating and occupying a gap in existing knowledge, making an original contribution that is accepted by the community ... transforming its author from novice to licenced practitioner'. See Carter (2011, p. 730).

³² On the 'capitalist presentism' driving this tendency, see McGlazer (2020, p. 23).

³³ Lang (2005, p. 159).

³⁴ Lee (2014, p. 163).

same way, then in what sense are they dead? What does it mean to be dead in this sense and (by extension) what might it mean for an author to be alive? These and deeper questions proliferate. If, more substantively, authors are just products of their societies, then Hitler would seem reducible to the 1918 Armistice—to which many will reply: really?³⁵ It was part of the magisterial genius of Roland Barthes that his essay proclaiming the death of the author gave only sweeping readings of texts and authors, most of which look questionable today.³⁶ Yet to question Barthes on these terms is only to do the everyday work of literary criticism: challenging interpretations. The Confirmation Bias Studies approach to ‘The Death of the Author’, by contrast, challenges the fundamental hypothesis: since, biologically, some authors are alive (and since Hitler seems irreducible to the Armistice) the hypothesis of the death of the author needs modification.³⁷ The Death of Some Authors is not as poetic a title as Barthes provided, but considered for confirmation bias, it probably has more mileage.

Tilting at Barthes is fun and in view of my previous comments about truth (and Hitler) it also has social use. Yet for all its influence, ‘The Death of the Author’ is a small essay based on smaller evidence: barely worth considering from the confirmation bias viewpoint. A much bigger example (but one just as influential today) might be Edward Said’s *Orientalism*. Here we have a full-scale study devoted to a defined hypothesis: western texts fundamentally misrepresent the east as a strategy of domination. Unlike Barthes, whose attention to texts in ‘The Death of the Author’ is cursory, Said supported his claim with detailed analyses of central western texts and many others have followed suit. *Orientalism* as first hypothesised by Said, therefore, can be said to have attained the level of proof required by T.S. Kuhn, since repeated confirmations of the hypothesis have been found. This is the case even if some of these confirmations are more persuasive than others. Yet confirmation bias remains a problem and to be fully validated the theory must also pass the criteria for proof of Popper; or, if it cannot be so validated, it needs to be rendered more precise through modification. As with the challenge of biological and undead authors to Barthes, what’s needed is counter examples.

I stress at this point that it is not this essay’s aim to disprove *Orientalism*. Far more substantial critiques of *Orientalism* already exist than this essay could aspire to and (depending on the author in question) more and less inverted versions of *Orientalism* already exist in *Occidentalism*.³⁸ The aim of the essay here is to propose a process for Confirmation Bias Studies. Like Barthes’ ‘Death of the Author’, *Orientalism* serves in the process as an illustrative theory. In so far as it shows shortcomings, what follows not so much disproves the theory as reveals areas to be developed.

Objectors might want to claim at this point that there are no counter examples to Said’s thesis, arguing that is why it is so powerful. They might add, too, that any cultural interpretation found counter to Said can itself be revealed to be

³⁵ Lang (2005, p. xii and pp. 15–16).

³⁶ See Ryan (2022, pp. 83–84). Notably, Barthes’ essay was ‘never [sic.] meant to be a traditional literary or scholarly essay’. See Logie (2013, p. 494).

³⁷ Lang (2005, p. 159).

³⁸ For a sample of the variety, see Salhi (2019), and Chen (2002).

a misinterpretation, leaving his thesis intact. This may be true. Yet the claim that something *can* be done is not the same as that it *ought* to be done; and unless it passes the falsifiability test, the thesis will remain open to doubt: one of the many theories circulating in the world of alternative facts, as likely to inflame opposition as to be found persuasive.³⁹ For scholarship, too, failing the falsifiability test is serious. A theory found to be unfalsifiable cannot be improved, even in the limited, Popperian sense of improvement ‘criticising our guesses’.⁴⁰ Scholarship following such unfalsified theory is therefore doomed to repetition.

This essay can only sketch out what a Confirmation Bias Studies response to this problem might look like. Only a consideration of every possible counter example to a prevailing thesis would finally reveal what qualifications the thesis would need to obtain Popper’s criteria of truth; even though for Popper (dealing in inductive science) a single counter example was enough. In this sense, one needs not individual case studies, but a continual testing of theory (established, emerging, or new) for confirmation bias. If that sounds like a lot of work, take comfort. If getting the full picture is going to keep us busy, we can make consistent, if incremental, gains along the way. Knowledge can develop. The humanities can address the need for truth in the public sphere with new commitment and energy. Yet methodology needs emphasis. Since counter examples are necessary in falsifiability tests, to begin work one needs a clear example of the theory under consideration. Clarifications, therefore, are likely to be an initial part of the procedure. In the case of Orientalism, this requires a little work.

Orientalism depends on the view that (in Said’s words) ‘European culture was able to manage—and even produce—the Orient politically, sociologically, militarily, scientifically and imaginatively during the post-Enlightenment period’.⁴¹ There would seem to be slippage here between ‘manage’ and ‘produce’, which points to stronger and weaker versions of the thesis. The claim that Europe ‘managed’ large parts of the Orient is uncontentious, but that it ‘produced’ that world—a much stronger claim—is remarkable and has stood out. This slippage in Said’s thesis, I think, continued throughout his work, but it is the stronger claim that is interesting, and which has been widely taken as representative. Interpreting Said here, for example, Shehla Burney writes: ‘In other words, Said argues that Orientalism is a built in system or method by which the West not only socially constructed and actually produced the Orient, but controlled and managed it through the tropes, images, and representations of literature, art, visual media, film and travel writing, among other aspects of cultural and political appropriation’.⁴² This is both the more interesting and stronger of Said’s claims and—while I do not think Said ever clearly

³⁹ For a recent discussion of Hume’s Law (‘you cannot derive an “ought” from an “is”’) see Radcliffe (2022, p. 44).

⁴⁰ Boyer (2009, p. 246).

⁴¹ Said (1978, p. 11).

⁴² Burney (2012, p. 23).

distinguished the two⁴³—it is this remarkable claim that has had most purchase and merits attention.

Having clarified the thesis to be tested, the next step in a confirmation bias study will be to find counter examples. In the case of Orientalism, that means finding examples of oriental culture isolated from western influence. For a thesis like Said's, a counter example requires isolation from western influence to qualify. That puts a premium on oriental works predating western influence, a premium on scholars who can read these works in their original language, and an onus on western scholars unable to read them to listen to those who can.⁴⁴ Evidence from these sources and scholars are precisely relevant to the claim that western influence 'produced' the Orient. If the pre-western sources reveal similar behaviours and attitudes in oriental countries to those found after the arrival of the west, the sources confirm a bias in the Orientalist thesis. Modifications to the thesis will therefore need to be made, for example regarding indigenous and local structures of power and morality; and these modifications will themselves eventually need testing for confirmation bias, since scholars today, including non-westerners, are rarely if ever entirely isolated from western influence. Clearly, there is a lot of work from scholars who can read pre-western, oriental languages for us to attend to.

Counter examples to Said's thesis (in its stronger form) might also be found in other kinds of oriental isolation from the west. Since isolation is key, many works post-dating east–west engagement will likely not qualify for a confirmation bias study. With western assumptions, for example about moral behaviour, they do not provide the truly counter example.

Yet greater and lesser degrees of isolation can exist and eastern artefacts post-dating western colonisation can be more and less isolated from the experience: more and less valuable, therefore, as relevant evidence in confirmation bias studies.⁴⁵ Though long post-dating African westernisation, for example, Chinua Achebe's novel *Things Fall Apart* purports to tell an African history before and after westernisation. Though scholars might argue the novel's presentation of pre-colonised Africa is damaged by western influences on Achebe, few would argue these so entirely harm his presentation of pre-colonial Africa that it has no merit.⁴⁶ For confirmation bias studies, the lesson is that partial counter examples too have a place in the testing procedure, albeit in these cases conclusions may be more provisional.

In such examples, the case for their isolation—as perhaps for any sample—needs to be made, which means assessing their eastern components against their western ones. Only through this procedure can distinctions necessary for conclusions about confirmation bias be drawn between their eastern and western elements. In the case of Achebe, for example, the validity of his portrait of the pre-colonial Igbo depends on his intimate, first-person and inherited knowledge of the Igbo through

⁴³ Vacillation between the stronger and weaker theses is implicit, for example, in Said's later claim that Orientalism concerns 'overlapping domains'. See Said (1985, p. 90).

⁴⁴ For example, Warscheid (2018, pp. 1–10).

⁴⁵ Hence the recent call for 'African language independence' from 'the hegemony of European languages in African Literature'. See Cantalupo (2016, p. 1).

⁴⁶ See Morrison (2014, pp. 177–192), Tiffin (1995, pp. 97–98).

membership.⁴⁷ Yet his membership is also of the Igbo after colonisation, taking in experiences, for instance, like his study of English literature at University College in Nigeria.⁴⁸ Evaluating post-colonial portraits of pre-colonial cultures means surrounding analysis of the sample texts with a ledger explaining the strengths and weaknesses of the evidence. Its provisional nature will then be clear, but that does not amount to ‘the impossibility of objectivity and impartiality in the human sciences’.⁴⁹ Rather, it ensures the evidence is understood in its proper context: neither over-estimated nor underestimated.

Studying isolated examples of this kind may also bring insights that are otherwise unobtainable. As a narrative about China by a native Chinese, for instance, Jung Chang’s *Wild Swans: Three Daughters of China* makes a claim to stand largely outside an Orientalist framework. Isolating Europe and America from its main narrative, stories of personal and local experience dominate in the work. The driving, though not exclusive, motive is the experience of the three daughters, to appropriate which would invite orientalist charges.⁵⁰ As with the veracity of Achebe’s portrait of the pre-western Igbo, the strength of Chang’s portrait in *Wild Swans* is its seeming reliance on first-person experience and the oral and cultural evidence gathered from forbears. From the point of view of Confirmation Bias Studies, therefore, *Wild Swans* provides the prima facie counter example required to test the Orientalist thesis.

Much as a reading of *Things Fall Apart* observes the evidence of gradual Igbo disruption caused by the west, a full-scale confirmation bias study of *Wild Swans* will detail the forms of unfolding disruption. In Chang’s story, though, the failings catalogued, largely emerging in isolation from western influence, signal oriental failings firstly, rather than western ones. Each example of Chinese abuse or atrocity in *Wild Swans* stands as a counter example to the thesis that Europe (or America) produced oriental China; and the hundreds of examples in the work (or in sampled works like it) each imply a need to modify Said’s stronger thesis. This is despite the role of colonisation and Marx in nineteenth and twentieth century Chinese history.⁵¹ Chinese agency is paramount in *Wild Swans*. Although its author is of China, the work in this aligns with wider ‘China-centred’ histories, which through research into Chinese archives, have brought to attention Chinese agency in historical processes of change.⁵²

As the present plight of the Uighurs illustrates, the proposed, literary work of cataloguing and detailing is important, yet there are wider points to observe here too. According to the Global Slavery Index, in 2016 there were 3.8 million people in China living in conditions of modern slavery.⁵³ From its opening sentence, older and newer forms of slavery are the abiding topic of *Wild Swans*:

⁴⁷ AbdelRahman (2005, p. 179).

⁴⁸ AbdelRahman (2005, pp. 178–180).

⁴⁹ AbdelRahman (2005, p. 189).

⁵⁰ On orientalist charges, see Spivak (1995, p. 24).

⁵¹ For discussion, see Hutters (2005).

⁵² See, notably, Cohen (1984, 2003).

⁵³ See ‘Global Slavery Index’, at China | Global Slavery Index (accessed, 31/10/22).

At the age of fifteen my grandmother became the concubine of a warlord general, the police chief of a tenuous national government of China. The year was 1924 and China was in chaos. Much of it, including Manchuria, where my grandmother lived, was ruled by warlords. The liaison was arranged by her father, a police chief official in the provincial town of Yixian in southwest Manchuria, about a hundred miles north of the Great wall and 250 miles northeast of Peking.

Like most towns in China, Yixian was built like a fortress. It was encircled by walls thirty feet high and twelve feet thick dating from the Tang dynasty (AD 618-907), surmounted by battlements, dotted with sixteen forts at regular intervals, and wide enough to ride a horse quite easily along the top. There were four gates into the city, one at each point of the compass, with outer protecting gates, and the fortifications were surrounded by a deep moat.⁵⁴

The slavery of concubinage, which is soon revealed as a commonplace deal made for advantage by the girl's father, stands out here in a series of images of Chinese and Manchurian isolation dating back to the Tang dynasty. The implication is that this slavery is natively Chinese, though the originally western name 'Peking' and the dating *anno domini* offer subtle modifications.

Whether this view is entirely correct is the task of historical criticism to discern and here the process of confirmation bias studies stands out. Isolating and so distinguishing the properly Chinese elements of the girl's concubinage from features externally derived allows the girl's representative and personal concubinage to be understood. Said, who claimed Orientalism is an ally of women's studies, in principle approves this procedure.⁵⁵ In the case of Manchuria, non-Chinese elements include the differently oriental influences of Russia and Japan: by no means isolated from the west (especially Russia), yet modulating its influence with their own oriental features. The slavery and misery of the 15-year-old girl sold into concubinage, Chang's grandmother, opens oriental histories in which the west plays a distant third fiddle. This is the origin of older and newer forms of enslavement in Chang's narrative, as Chang's two following daughters of China—her mother and then Chang herself—grind out lives defined either by, or by reaction to, concubinage as the original sin.

Other forms of enslavement loom along the way. The puppet state of Manchukuo which followed the Japanese invasion of 1931 stands out both as an external influence on Manchuria's isolation and for its brutality. The Chinese Kuomintang (backed by America) is also brutal, inviting both condemnation and critical procedure. The torture of Chang's mother by the Kuomintang, for example, entails putting her through a mock execution by firing squad—not an especially Chinese means of intimidation—but also one in which Chinese agents and agency stand out. In a passage like this (*Wild Swans*, p. 137) the isolating procedure of confirmation bias studies gives a basis for the apportioning of responsibility, providing some affirmation of the strong Orientalist thesis and, also, correctives to it.

⁵⁴ Chang (1993, p. 27).

⁵⁵ Said (1985, p. 106).

Sampled scenes of this kind provide a very large terrain for literary study along confirmation bias lines, but they also invite theoretical considerations important for procedure. There is a case for western disruption of Chinese lifestyles in the Kuomintang passages of *Wild Swans*, but transparently the passages are part of a very much bigger story. In its evaluative procedure, therefore, Confirmation Bias Studies will want to indicate the frequency of the kinds of passages it samples vis a vis passages and themes of other kinds.

The relations of parts to the whole raise bigger questions too. The forgoing analysis holds oriental (Chinese, and more briefly, Japanese) culture largely responsible for enslavements and abuse in *Wild Swans* and America somewhat so. A larger analysis of *Wild Swans* would need, too, fully to consider the relation of Marxism and Enlightenment ideas of progress to the Cultural Revolution, addressing the agency and responsibility of each.⁵⁶ Yet at stake here are also literary issues of time and representation. It is perfectly possible to imagine a version of *Wild Swans* focusing exclusively on the Kuomintang era to bring out its (and America's) abuses, just as it is possible to imagine a *Wild Swans* that does not open with an isolated Chinese history dating back to the Tang dynasty; and just as it would be possible to write a *Wild Swans* with the Enlightenment and Marx to the fore. Literature constructs sequences of cause and effect that need recognition, especially if a confirmation bias approach to it is to be persuasive.⁵⁷ Chang writes: 'Manchuria was the key battleground in the civil war, and what happened in Jinzhou was becoming more and more critical to the outcome of the whole struggle for China'.⁵⁸ This is, in effect, the historical justification for opening the work in the earlier Manchuria of the warlords and in the province's Tang origins. Reading works for confirmation bias in theories will mean recognising the way they set events within temporal parameters of cause and effect. It will also mean testing those parameters against other historical accounts. In a case like *Wild Swans*, these accounts will be somewhat self-selecting. Since *Wild Swans* is banned in China, official or state sponsored accounts of Chinese history disqualify themselves as historical tests of the work since they do not recognise it.

No account of the ledger of evidence needed for assessing works like *Wild Swans* or *Things Fall Apart* is complete without comment on English, the medium where east meets west.⁵⁹ Despite Achebe's defence of English in colonial representation, the confirmation bias approach (like any truth-seeking approach) should here too stress the provisional nature of its findings.⁶⁰ To be stressed too is that cultural translation need not work against the country described in English, especially if the describer is from that country. Critiques of *Wild Swans* tend not to fault its broad

⁵⁶ Chen comments relevantly here: 'it would not be accurate to say that Chinese political and intellectual culture is nothing more than an outpost of mindlessly replicated Western thought. However Western these 'Chinese' ideas may be in their origins, it is undeniable that their mere utterance in a non-Western context inevitably creates a modification of their form and content'. See Chen (2002, p. 2).

⁵⁷ See de Guevara and Kostic (2017, p. 13), Fernandez (2018, p. 80). More generally, see, for example, Andrews (2007); on cause and effect in *Wild Swans*, see Li and Li (2021, pp. 54–77), though the comparative (rather than historical) approach limits the article's conclusions.

⁵⁸ Chang (1993, p. 128).

⁵⁹ On language creating worldviews, see Underhill (2011).

⁶⁰ On Achebe's defence of English and questions arising, see Lynn (2017, pp. 77–95).

picture of Chinese (and from 1931, Japanese) atrocities performed on Chinese people. Rather, they criticise a tale of the Cultural Revolution too sharply dividing the evils perpetrated by an elite cadre (the so-called Gang of Four) from the actions of the wider populace. *Wild Swans* implies Chang and most students avoided violence during the revolution, even dividing Red Guards into a peaceful majority and a small minority ‘actually involved in cruelty or violence’.⁶¹ It is here, claims Shuyu Kong, that Chang’s English readership most exerts its influence on the narrative of China:

the account is plausible and seductive precisely because we can imagine ourselves adopting a similarly detached, sceptical attitude in episodes of mass hysteria. We like to think that we, too, are sensitive individuals, would avoid violence and would be among the first to challenge it, or at least would avoid participating wherever possible. This is, after all, how rational people behave.⁶²

The rational ‘we’, here, are the English speakers who are Chang’s target readership. For many, *Wild Swans* is a bridge to a country largely unknown, raising questions of the relation of geography and distance to responsibility. Yet Kong’s point is that the meeting of the oriental author and the western readership produces a happier picture of the Chinese than is due. Since the author is a native of China, this is neither the idealisation nor demonization of the east by western authors posited by Said. It is, rather, a third thing, which we might call ‘angelification’, in which identification by the eastern author with the west enhances the moral presentation of the east. Considering English works on the orient by its natives, Confirmation Bias Studies will need to keep the phenomenon in mind.

What bearing have these academic points on our global crisis of interpretation? As consequences of a Confirmation Bias procedure, each point reveals an aspect of Orientalism needing development. They reveal the efficacy of the confirmation bias method and so the capacity of the humanities to grapple with hypothetical and theoretical truths. In the age of superabundance, they reveal the academy setting a systematic example of how to tell fake theoretical news from true; and they suggest a procedure by which the humanities can prepare students for superabundance, as a model of truth in the public sphere.

Just as there is no bias without a category of truth, the fore-mentioned terms (idealisation, demonization, angelification) base standards of judgement on the putative existence of facts; what Said called, ‘the world of reality’.⁶³ Confirmation Bias Studies seeks this putative world through patient testing of the theories about it. As an academic field, it is largely untrod in the humanities and so ripe for development. Yet it is also the means by which the academe can set an interpretative example, presenting itself to the world of superabundance as a model for telling theoretical wheat from chaff. Were the next generation of humanities graduates to be trained in confirmation bias theory, they would be more habituated to making critical distinctions

⁶¹ Kong (quoting *Wild Swans*): (1999, p. 245).

⁶² Kong (1999, p. 249).

⁶³ Said (1985, p. 100). See, more recently, Fischer and Klazar (2020, p. 6), Brown (2016, pp. 1–2), Ryan (2022, pp. 84–87).

and less susceptible to fake news. Public discourse and our democratic institutions would benefit.

Conclusion: nine features of confirmation bias study for the humanities of the future

To summarise, the confirmation bias study proposed here for the humanities has eight distinct features. Its principal address is to theories and to theory. Its focus is on misinterpretation, rather than interpretation. Its aim is to develop theory through modification. Its means is the identification and analysis of counter examples. Its conclusions will explain how the examples modify the theory. The conclusions will be provisional. It is a contribution from the humanities to public discourse in the age of superabundance. It maintains a category of truth.

Since it answers human tendencies to seek and interpret evidence in ways partial to existing beliefs, a ninth and last feature will be reflection on oneself as the reasoning analyser. In this essay, I have argued for a crisis of truth in the public sphere that is reflected in the academy and to which the humanities can respond. I have illustrated a procedure for this response through consideration of the ‘Death of the Author’ and Orientalism as sample theories, giving main attention to Said’s stronger theory of the relations of east and west, which has garnered attention in both spheres. I have done so as a middle-aged man who spent his first 9 years in Canada, who has lived in the UK, Europe and Israel, travelled in Egypt and Turkey, and resides in Scotland. From around the eighteenth century, the Rist family were farmers in Suffolk, but the name Rist is much more common in Germany and Estonia than in England, so an earlier immigrant history seems likely. On my mother’s side, my grandmother was English but with family origins in France, while my Jewish grandfather’s family had come from Poland. How far these identities shed light on this essay is for me unclear. As a middle-class academic born into an academic family, I engage actively with academic (including political) topics, especially in the humanities. As someone who would have been executed under Nazi race laws, and whose Jewish forbears died by those laws, I am alive to antisemitism, racism, dictatorships and the distortions of truth on which they depend. No doubt this last has a bearing on this essay, for it inculcates the view that truth, in culture, is a matter of life and death; and that in the age of superabundance, maintaining truth requires of us a coherent process of reason.

Acknowledgements Not applicable.

Author contributions I, Thomas Rist, am the sole author of this article.

Funding No funding was received by the author in relation to this article.

Data availability Not applicable.

Declarations

Conflicts of interest There are no conflict of interest.

Ethical approval No committee was involved in this research or therefore approved it. All research was performed in accordance with relevant guidelines and regulations.

Informed consent This article does not contain any studies with human participants performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- AbdelRahman F (2005) Said and Achebe: writers at the crossroads of culture. In: *Journal of Comparative Poetics. Edward Said and critical decolonisation*, vol 25. pp 177–192
- Andrews M (2007) *Shaping history: narratives of political change*. Cambridge University Press, Cambridge
- Boyer A (2009) Open rationality: making guesses about nature, society and justice. In: Parusnikova Z, Cohen R (eds) *Rethinking Popper. Boston studies in the philosophy of science*, vol 272. Springer, Dordrecht, pp 245–55
- Brady A-M (2010) *Propaganda and thought work in contemporary China*. Rowman & Littlefield Publishers, Lanham
- Brown T (2016) Evidence, expertise, and facts in a “post-truth” society. *Br Med J*. <https://doi.org/10.1136/bmj.i6467>
- Burney S (2012) Orientalism: the making of the other. In: *Counterpoints: studies in criticality. Pedagogy of the other: Edward Said, postcolonial theory and strategies for critique*, vol 417. Peter Lang, New York, pp 23–39
- Byrne K (2011) From theory to practice: literary theory in the classroom. In: Bradford R (ed) *Teaching theory*. Palgrave Macmillan, New York
- Cantalupo C (2016) Africa antitranslation. *Res Afr Lit* 47(3):1–14
- Carter S (2011) Doctorate as genre: supporting thesis writing across campus. *High Educ Res Dev* 30(6):725–736
- Chakravorty Spivak G (1995) Can the subaltern speak? In: Ashcroft B, Griffiths G, Tiffin H (eds) *The post-colonial studies reader*. Routledge, London, pp 24–8
- Chang J (1991) *Wild swans: three daughters of China*. Harper Collins, London. Reprint 1993
- Chen X (2002) *Occidentalism: a theory of counter-discourse in post-Mao China*, 2nd edn. Rowan & Littlefield, Lanham
- Cohen P (1984) *Discovering history in China: American historical writing on the recent Chinese past*. Columbia University Press, New York
- Cohen P (2003) *China unbound: evolving perspectives on the Chinese past*. Routledge, London
- Collini S (2012) *What are universities for?* Penguin, London
- De Guevara BB, Kostic R (2017) Knowledge production in/about conflict and intervention: finding “facts”, telling “truth.” *J Interv State Build* 11(2):1–20
- Dhir A, Khalil A, Kaur P, Rajala R (2019) Rationale for “liking” on social networking sites. *Soc Sci Comput Rev* 37(4):529–550
- Dubois E, Blank G (2018) The echo chamber is overstated: the moderating effects of political interest and diverse media. *Inf Commun Soc* 21:729–745

- Fernandez J (2018) Story makes history, theory makes story: developing Rusen's Historik in logical and semiotic directions. *Hist Theory* 57(1):75–103
- Fischer R, Klazar E (2020) Facts, truth and post-truth: access to cognitively and socially just information. *Int J Inf Divers Incl* 4(3/4):5–19
- Herman ES (2000) The propaganda model: a retrospective. *Journal Stud* 1(1):101–112
- Huters T (2005) Bringing the world home: appropriating the west in late Qing and early Republican China. University of Hawai'i Press, Honolulu
- Jay P (2014) The humanities 'crisis' and the future of literary studies. Palgrave Macmillan, New York
- King E et al (2018) Systematic subjectivity: how subtle biases infect the scholarship review process. *J Manag* 44(3):843–855
- Kong S (1999) Swan and spider eater in problematic memoirs of cultural revolution. *Positions East Asia Cult Crit* 7(1):239–52
- Lang B (2005) Post-holocaust: interpretation, misinterpretation and the claims of history. Indiana University Press, Bloomington
- Lee M (2012) Evidence, coincidence and superabundant information. *Vic Stud* 54(1):87–94
- Lee M (2014) Falsifiability, confirmation bias, and textual promiscuity. *J Ninet Century Am* 2(1):162–71
- Lesley LK, Emrich O, Gupta S, Norton MI (2017) 'Does "liking" lead to loving?': the impact of joining a brand's social network on marketing outcomes. *Am Mark Assoc* 54(1):144–55
- Levin J (2005) The business culture of the community college: students as consumers, students as commodities. *New Dir High Educ* 129:11–26
- Li L, Li X (2021) Who "let all this happen"?: shifts of responsibilities in representing the cultural revolution in Jung Chang's *Wild Swans*. *Lang Lit* 30(1):54–77
- Logie J (2013) 1967: The birth of "The Death of the Author." *Coll Engl* 75(5):493–512
- Lokot T, Diakopolous N (2016) New bots; automating news and information dissemination on twitter. *Digit Journal* 4:682–699
- Lynn T (2017) Chinua Achebe and the politics of narration: envisioning language, African histories and modernities. Palgrave MacMillan, Cham
- Mattison J (2012) Literary theory in the postgraduate classroom: it's role and challenges. In: Karlsson L (ed) *Lärlärdom: högskolepedagogisk konferens*. Kristianstad University Press, Kristianstad
- McGlazer R (2020) Old schools: modernism, education and the critique of progress. Fordham University Press, New York
- McSweeney B (2021a) Fooling ourselves and others: confirmation bias and the trustworthiness of qualitative research, part 1. *J Organ Change Manag* 34(5):1063–1075
- McSweeney B (2021b) Fooling ourselves and others: confirmation bias and the trustworthiness of qualitative research, part 2. *J Organ Change Manag* 34(5):848–859
- Mercier H (2012) Reason is for arguing: understanding the success and failures of deliberation. *Polit Psychol* 33(2):243–258
- Morrison J (2014) Chinua Achebe. Manchester University Press, Manchester
- Nagel T (1997) The last word. Oxford University Press, Oxford
- Nickerson R (1998) Confirmation bias: a ubiquitous phenomenon in many guises. *Rev Gen Psychol* 2(2):175–222
- Nickerson R (2020) Argumentation: the art of persuasion. Cambridge University Press, Cambridge
- Noturno MA (1984) The Popper/Kuhn debate: truth and the two faces of relativism. *Psychol Med* 14:273–289
- Radcliffe E (2022) Hume on the nature of morality. Cambridge University Press, Cambridge
- Roberts R (2017) Lessons from Popper for science, paradigm shifts, scientific revolutions and exercise psychology. *BMJ Open Sports and Exercise Medicine* 3/1:1–11
- Rorty R (1979) Philosophy and the mirror of nature, 13th edn. Princeton University Press, Princeton. Reprint 2009
- Rosende D (2009) Popper on refutability: some philosophical and historical questions. In: Parusnikova Z, Cohen R (eds) *Rethinking popper*. Boston studies in the philosophy of science, vol 272. Springer, Dordrecht, pp 135–54
- Ryan M-L (2022) Media, genres, facts and truths: revisiting basic categories of narrative diversification. *Neohelicon* 49:75–88
- Said E (1978) Orientalism. Routledge and Kegan Paul, London
- Said E (1985) Orientalism reconsidered. *Cult Crit* 1:89–10
- Salhi ZS (2019) Occidentalism: literary representations of the Maghrebi experience of the east–west encounter, Edinburgh studies in modern Arabic literature. Edinburgh University Press, Edinburgh

- Simpson J (2003) Faith and hermeneutics: pragmatism vs pragmatism. *J Mediev Early Mod Stud* 33(2):215–239
- Spivak GC (1995) Can the subaltern speak? In: Ashcroft B, Griffiths G, Tiffin H (eds) *The post-colonial studies reader*. Routledge, London, pp 24–28
- Svensson G, Wood G (2007) Are university students really customers? When illusion may lead to delusion for all! *Int J Educ Manag* 21(1):17–28
- Tiffin H (1995) Post-colonial literatures and counter-discourse. In: Ashcroft B, Griffiths G, Tiffin H (eds) *The post-colonial studies reader*. Routledge, London, pp 95–98
- Underhill J (2011) *Creating worldviews: metaphor, ideology and language*. Edinburgh University Press, Edinburgh
- Unerman J (2020) Risks from self-referential peer-review echo chambers developing in research fields; 2018 keynote address presented at the British accounting review 50th anniversary celebrations, British accounting and finance association annual conference, London. *Br Account Rev* 52(5):1–26
- Warscheid I (2018) The Islamic literature of the precolonial Sahara: sources and approaches. *Hist Compass* 16(5):1–10
- Wolfreys J, Womack K, Robbins R (2014) *Key concepts in literary theory*, 3rd edn. Edinburgh University Press, Edinburgh
- Zu S et al (2022) Exposure effects or confirmation bias? Examining reciprocal dynamics of misinformation, misperceptions, and attitudes towards Covid-19 vaccines. *Health Commun*. <https://doi.org/10.1080/10410236.2022.2059802>