



Social Epistemology

A Journal of Knowledge, Culture and Policy

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tsep20

Gatekeeping in Science: Lessons from the Case of Psychology and Neuro-Linguistic Programming

Katherine Dormandy & Bruce Grimley

To cite this article: Katherine Dormandy & Bruce Grimley (03 Apr 2024): Gatekeeping in Science: Lessons from the Case of Psychology and Neuro-Linguistic Programming, Social Epistemology, DOI: [10.1080/02691728.2024.2326828](https://doi.org/10.1080/02691728.2024.2326828)

To link to this article: <https://doi.org/10.1080/02691728.2024.2326828>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 03 Apr 2024.



Submit your article to this journal [↗](#)



Article views: 3467



View related articles [↗](#)



View Crossmark data [↗](#)

Gatekeeping in Science: Lessons from the Case of Psychology and Neuro-Linguistic Programming

Katherine Dormandy ^a and Bruce Grimley^b

^aDepartment of Christian Philosophy, University of Innsbruck, Innsbruck, Austria; ^bPsychology, Universidad Central de Nicaragua, Managua, Nicaragua

ABSTRACT

Gatekeeping, or determining membership of your group, is crucial to science: the moniker ‘scientific’ is a stamp of epistemic quality or even authority. But gatekeeping in science is fraught with dangers. Gatekeepers must exclude bad science, science fraud and pseudoscience, while including the disagreeing viewpoints on which science thrives. This is a difficult tightrope, not least because gatekeeping is a human matter and can be influenced by biases such as groupthink. After spelling out these general tensions around gatekeeping in science, we shed light on them with a case study from psychology. This concerns whether academic psychologists rightly or wrongly classify the applied-psychology framework of NLP (‘neuro-linguistic programming’) as unscientific and even pseudoscientific. This example of gatekeeping is particularly instructive because both the NLP community and the psychology community, we argue, make legitimate but also illegitimate moves. This case gives rise to several general insights about gatekeeping in science more generally.

ARTICLE HISTORY

Received 23 February 2024
Accepted 1 March 2024

KEYWORDS

Gatekeeping; pseudoscience; neuro-linguistic programming; philosophy of psychology

1. Introduction

Gatekeeping, or determining membership of your group, is crucial to science. For science is an epistemic gold standard: calling methods or results ‘scientific’ indicates epistemic quality or even authority, and false pretenders to this quality abound. But gatekeeping is fraught with dangers. Gatekeepers in science must *exclude* scientific fraud, bad science and pseudoscience, while still welcoming the disagreement on which scientific progress thrives. This is a narrow tightrope. One reason is that fruitful scientific disagreement often concerns not just hypotheses but more foundational issues such as methodology and even basic ontology. Another is that gatekeeping brings in human factors such as careerism, groupthink and other biases.

We investigate these tensions with a case study from psychology. It concerns whether NLP (‘neuro-linguistic programming’), a phenomenologically oriented framework for psychotherapy, coaching and body-mind health interventions, is rightly or wrongly classed as unscientific. A formidable literature in academic psychology says ‘unscientific’: NLP is supposedly discredited (Heap 1989b; Nocross, Koocher and Garofalo 2006), ethically dubious (Greif 2022), empirically unconfirmed (Passmore and Rowson 2019), and for these and other reasons, pseudoscientific (Briner 2016; Greif 2022; Kanning 2019; Roderique-Davies 2009; Witkowski 2010).

Prominent NLP-representatives reject this verdict. They concede that NLP needs professional oversight and theoretical and empirical improvement but reject the labels ‘unscientific’ and

CONTACT Katherine Dormandy  katherine.dormandy@uibk.ac.at  Department of Christian Philosophy, University of Innsbruck, Karl-Rahner Platz 1, Innsbruck 6020, Austria

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

especially ‘pseudoscientific’. They further accuse psychology of *inappropriate gatekeeping* against NLP (Arroll and Henwood 2017b; de Rijk, Gray and Bourke 2022; Parker 2022a; Wake, Gray and Bourke 2013). They argue that psychology is biased against NLP, that it straw-mans NLP, and that its empirical criticisms rely on flawed techniques, poor meta-analyses and cherry-picking (Einspruch and Forman 1985; Gray et al. 2013).

This case of gatekeeping is difficult to adjudicate because both sides have flaws. In exploring it, we seek insight into NLP’s scientific status and into the extent to which academic psychology’s gatekeeping against NLP abides by, or falls short of, scientific standards. We conclude with general lessons about gatekeeping in science generally.

This discussion bears indirectly on the demarcation problem in the philosophy of science. This concerns how to distinguish the practices, methods, theoretical commitments and results (for short: framework) of science from those of activities that might superficially resemble science – including bad science, science fraud and pseudoscience (Mukerji and Ernst 2022); NLP has been placed in each category.

We’ll use phrases like ‘psychology says’ and ‘NLP says’ to mean ‘many in academic psychology say’ and ‘many NLP-representatives say’ respectively. Section 2 characterizes gatekeeping in science. Section 3 describes NLP. Section 4 gives the psychology community’s case against NLP, and Section 5 gives NLP’s case against the psychology community. Section 6 draws conclusions about the NLP case and gatekeeping in science generally.

2. Gatekeeping in Science

Gatekeeping’s role, ostensibly, is to preserve something your group holds dear – its standards, values, reputation, the physical or emotional safety of its members or just its exclusivity. Gatekeeping can do this directly, by barring entry to potential threats, or indirectly, by setting norms of appropriateness that become habitual.

Whether a case of gatekeeping is good, bad or neutral depends on at least three factors. First is the *value* of what it is protecting. Gatekeeping can be neutral: think of charging an average ticket price to a cinema. It can also be bad: think of schoolyard cliques or racial segregation. Or it can be good, e.g. creating safe spaces for vulnerable populations. Second, the evaluation of gatekeeping depends on whether the *gatekeeping methods are fitting*. Fittingness varies between contexts. An intuitive example, however, is that beating up outsiders is typically not a fitting way to keep them from joining your ethics committee. Third, the evaluation of gatekeeping depends on whether it *succeeds*: whether (i) it preserves what it seeks to preserve (ii) without including things that deserve to be kept out or excluding things that deserve to be let in. Successful gatekeeping avoids the Scylla of *excessive leniency* and the Charybdis of *excessive tightness*.

We’ll focus on gatekeeping in science: determining which frameworks to accept or reject as scientific. We use the term ‘science’ normatively: not everything done under the aegis of science is good. But science as we understand it represents certain ideals of inquiry.¹ It is these ideals that gatekeeping in science seeks to preserve, and we assume that gatekeeping in science thus seeks to preserve something valuable. These ideals include seeking understanding, striving for accuracy or empirical adequacy (or, for anti-realists, practical efficacy); substantiating results with arguments and other evidence; subjecting results and methods to critical feedback by peers, and engaging in ongoing epistemic self-reflection (Tetens 2013). They include systematicity (Hoyningen-Huene 2013), i.e. organized categories and methods, and they presuppose a commitment to openness, i.e. to accepting whatever results a responsible inquiry delivers (McIntyre 2019). Science’s advances in understanding and technology suggest that these ideals deserve preservation against those who would undeservingly cash in on science’s credentials. We’ll ask whether psychology preserves them in gatekeeping against NLP.

We assume, second, that science has *fitting gatekeeping methods*; these respect the abovementioned ideals of science. They promote or at least do not hinder the search for understanding,

accuracy or empirical adequacy, substantiation of results, critical peer feedback, epistemic self-reflection, systematicity and open-endedness. Examples include peer review rather than prestige bias, striving to understand a view before rejecting it and reflecting on whether your motives are scientific. That science *has* fitting gatekeeping methods, however, does not entail that scientists always *use* them. We'll ask whether psychology's gatekeeping against NLP is fitting.

Third, is gatekeeping in science *successful*? Sometimes yes, but sometimes no. It can be *excessively lenient* when background beliefs from the surrounding culture infiltrate the scientific framework. Think of the sexist ideology that influenced the doctor Edward Clarke to argue, in respected scientific venues (with an out-of-context application of thermodynamics), that higher education would damage women's fertility (Oreskes 2019, 76–80). Gatekeeping in science can also be *excessively tight*. This happens, for instance, when a dissenter is excluded despite pinpointing genuine flaws in the scientific framework. Barbara McClintock's research on transposition and mobile genetic material was belittled for decades until it won her a Nobel prize (Keller 1983). To be fair to overzealous gatekeepers, most scientific heterodoxies do fall flat. But excessively tight gatekeeping can also come from bias; McClintock's case is plausibly due at least partly to sexism (Keller 1983).

Excessively lenient or tight gatekeeping in science is dangerous. Science thrives on disagreement – not just over hypotheses, but over methodology and even sometimes basic ontology. Admitting the wrong foundational ideas can derail science for centuries, but so can excluding fruitful new paradigms. Gatekeeping in science needs continued vigilance.

Gatekeeping in science seeks to keep out perversions of science including science fraud, bad science and pseudoscience (Mukerji and Ernst 2022). The boundaries between them are fuzzy, but here is a description of prototypical cases. The *science fraudster* acts as if they abide by scientific ideals but covertly and deliberately violates them (McIntyre 2019). Someone doing *bad science*, in contrast, *does* abide by scientific ideals but does so poorly – because of incompetence, insufficient effort or insufficient resources. To draw an analogy with playing chess (Mukerji and Ernst 2002), the science fraudster distracts her opponent and switches his pieces around and then pretends to win honestly, whereas the bad scientist, playing in good faith, moves her queen into a vulnerable position without realizing it.

Pseudoscience is often assimilated to bad science or science fraud (e.g. Greif 2022) but differs from both. Like the bad scientist, the pseudoscientist has methods and results that are scientifically substandard and thus apt to be inaccurate or practically defective. But the bad scientist at least *aims* for accuracy or practical efficacy. Not so the pseudoscientist: she might or might not happen to think that her contribution is accurate (Boudry 2022; McIntyre 2019), but accuracy is not her primary concern. Her concern rather is to influence people in some area having nothing to do with science. Consider the lead-paint salesman, who cares less about the safety of his product than about selling it. This is dishonesty, but it differs from science fraud. Whereas the fraudster cheats *while participating in science*, the pseudoscientist cannot really be said to cheat because she *does not actually participate in science*. She is rather participating in an entirely different type of discourse, for example that of *influencing people by getting them to think that she is participating in science*. Pseudoscientists are thus a kind of *bullshitter* (Frankfurt 2005; Ladyman 2013). One frequent pseudoscientific move is to spuriously claim scientific status; another is to hoodwink people about what constitutes science to begin with. The lead-paint salesman, aiming for profit, claims falsely that his remedy has been researched or spuriously downplays scientific research into the dangers of lead paint. Returning to the chess analogy, the pseudoscientist does not play chess, but 'pigeon chess' (Mukerji and Ernst 2022, 393–394) – the pigeon knocks the pieces over and relieves itself on the board, claiming to have won the game.

Another possibility is *proto-science* (Mukerji and Ernst 2022, 394). A proto-scientific framework has received as yet little research, perhaps because it is young or because its proponents lack resources but work as responsibly as they can. Its claims are at best promising hypotheses. Many NLP-ers regard NLP as a proto-science, but many psychologists think it is irremediably bad science, or even science fraud or pseudoscience.

Before evaluating psychology's gatekeeping against NLP, we must describe NLP.

3. NLP

There is controversy about what NLP is, not least within NLP itself (Sturt 2012; Tosey and Mathison 2009, 3). Its three co-founders (Richard Bandler, John Grinder and Frank Pucelik) practice different versions with little mutual exchange. In a qualitative study, Grimley (2016) discovered 14 definitions of NLP, ranging from 'Whatever works' to 'A model from cognitive psychology' (2016, 168–169).

One division in the NLP community concerns attitudes toward science. Those we'll call *ascientific NLP-ers* are practitioners, content to outsource questions of NLP's scientific status to researchers. They found their practices on positive personal and clinical experience of NLP interventions and claim nothing beyond these experiences. Another group of NLP-ers, however, manifests disrespect for science, for example claiming more scientific status for NLP-interventions than is warranted – typically for commercial reasons and independently of a concern for accuracy. We may thus call them *NLP-bullshitters* (and often, as we'll see, peddlers of pseudoscience). Finally, what we'll call *science-minded NLP-ers* research or promote research into NLP. They continue to articulate a theoretical base for NLP (de Rijk, Gray and Bourke 2022, Linder-Pelz and Hall 2007, 12–17; Tosey and Mathison 2009; Wake, Gray and Bourke 2013), perform studies of NLP-derived techniques (Arroll et al. 2017a; Parker 2022b) and advocate science-mindedness in the NLP community (de Rijk and Parker 2022; de Rijk et al. 2019; Grimley 2013, 2019, 2020; Linder-Pelz 2010; Sturt 2012).

Given these complications we'll describe NLP as generally as possible. Its founding interest was to understand how skilled people achieve success and whether their behavioral strategies could be emulated, operationalized and taught. The challenge was that skilled people often cannot *describe* these strategies (Bandler and Grinder 1979, 6–7): a dancer cannot explain how he pirouettes. Rather, in pirouetting he deploys *tacit knowledge* (Bostic St Clair and Grinder 2001, 41), akin to philosophers' *knowing-how*. In exploring how to make this knowledge explicit, NLP aimed to get under the hood of expert performance (Tosey and Mathison 2009, 115). Its founders thus describe NLP as 'The study of the structure of subjective experience' (Dilts et al. 1980).

NLP sought to do this by modelling skilled performers. To *model* an exemplar is to emulate her behavior in a particular context and to systematize her strategies for achieving success (Dilts 1998). The resulting *model* describes both the exemplar's patterning and the modeler's patterning based on her (Bandler and Grinder 1979; Bostic St Clair and Grinder 2001; Burgess 2014; DeLozier and Grinder 1987; Dilts 1998; Grinder and Pucelik 2013).

Expert performance in any domain can be modelled. But for their first foray into modelling the NLP founders chose psychotherapy, modelling the eminent psychotherapists Fritz Perls, Virginia Satir and Milton Erickson. Although each is known for particular theoretical approaches, NLP was less interested in these than in their tacit knowledge, identifying what the exemplars *did* in communicating with clients (Bandler and Grinder 1979, 6–7; Dilts et al. 1980). This translated into a perceived antagonism to psychological theory. Early NLP texts even called psychological theory 'psychotheology', involving 'different religious belief systems with very powerful evangelists working from all of these different orientations' (Bandler and Grinder 1979, 6). These statements predictably irritated psychologists (e.g. Ouellette in Grimley 2015a, 124).

NLP developed two models for psychotherapy and other behavioral-change work. The 'meta-model', based on Perls and Satir, describes how language can help people understand and re-structure their *maps of reality* (Bandler and Grinder 1975b; Grinder and Bandler 1976; Korzybski 1994, 58). The 'Milton model', based on Erickson, identifies how practitioners can use language to help people shift their unconscious sensory experience without needing to map that shift into conscious understanding (Bandler and Grinder 1975a; Grinder, DeLozier and Bandler 1977). Despite distancing itself from psychological theory, these two flagship NLP models were liberally supplemented with borrowings from psychology and linguistics, some wholesale and others in altered form. Examples include Pavlovian conditioning (which inspired NLP's notion of 'anchoring') (Pavlov 1927), Chomsky's transformational grammar (1957), the TOTE model of Miller et al. (Miller, Galanter and Pribram 1960) and Ericksonian hypnosis. But modelling, as opposed to theory, was always primary in NLP.

Modelling's primary role in NLP was to emulate exemplars generally, with psychotherapy being the first application. But it acquired a secondary role as a tool *within* client interventions. NLP practitioners regard the client as a unique exemplar who has 'expertly' got himself into a behavioral muddle. Modelling that behavior helps practitioners co-engineer bespoke alternative strategies for him.

Because most exemplar patterning stems from tacit knowledge, NLP needed a framework for systematizing such knowledge. It thus developed a general – and complex – account of how human experience and thought (i.e. human phenomenology) is structured and how communication of these succeeds or fails (Bostic St Clair and Grinder 2001). When we receive information from the world, our cognitive faculties transform it on two levels, one sensory and the other linguistic. The sensory level transforms information using our *representational systems* – visual, auditory, kinesthetic, olfactory or gustatory. The linguistic level transforms it through language and other cognitively 'higher' functions. These two transformations interact: sensory input affects the language we use, and language constrains how we perceive and frame sensory input. Thus information gets distorted and deleted to produce generalizations, which are linked together to create meaning.² This meaning can conceal or make salient certain behavioral options. (For example, if you associate seeing bees with feeling pain, you may fear bees and thus springtime walks.) Thus a person's map, and the tacit knowledge it contains, makes the difference between excellent and mediocre performance.

Because modelling aims to make explicit the exemplar's tacit knowledge, which the exemplar can usually not articulate, it is helpful for a practitioner to *calibrate* to the exemplar. This means observing her minute behavioral cues, such as pulse rate, skin tone, direction of gaze and language while suspending pre-judgments. Such cues may yield insight into the exemplar's current subjective experiences, and thus tacit knowledge, in ways that both she and even the practitioner may be consciously unaware of.

Behavioral-change interventions in NLP focus on helping clients make their maps more conducive to excellent performance, by helping them change the structure of their experiences and language. Innovative NLP-ers have used the above ideas and methods to develop concrete techniques for behavioral-change interventions (see Bodenhamer and Hall 1999). One representative example is the RTM protocol. The practitioner invites the client to re-experience a phenomenological state from the past, then break that pattern and create a dissociated state in the present by observing themselves in the past from the third-person perspective, on an imagined cinema screen to create distance. In this labile state, the client can undergo re-consolidation, as his unconscious mind adjusts his sensory representations and associated meanings into a new map – enabling him to experience and consolidate an emotional state that better meets his needs (Sturt 2022). Empirical criticism in psychology focuses on NLP-derived techniques such as this (sections 4 and 5.3).

Because of its complexity and the skill needed to calibrate to clients, NLP requires extensive training to understand and practice.

4. The Case Against NLP

Some claim that psychology should not take NLP seriously because psychologists supposedly do not. NLP is seldom taught in universities and features in no standard psychology textbooks (Heap 2008, 5), and some regard it as discredited (Witkowski 2010, 65).³ But this argument from authority can be dismissed. It does not defend the move from 'is' to 'should'. The fact that scientists *are* gatekeeping against NLP does not itself imply that they *should*. We are asking whether this gatekeeping is *justified*.

Let's consider some possibly more promising arguments against NLP. Some are sound, others are not; we'll simply report them here and evaluate them in Section 5.

4.1. Ethical Criticisms

NLP has been subject to harsh ethical criticism, even recommending that practitioners be criminally indicted (O'Donohue and Ferguson 2006; Witkowski 2010). First, some NLP-practitioners are accused

of *manipulation, sometimes reckless, of clients or audiences*. Grant narrates how an NLP trainer, when asked by a student why his course started 20 minutes late, publicly shamed the latter by crawling on the floor repeating sarcastically, ‘your Lordship I am so sorry, please forgive me’ (Grant 2019, 50; cf. Tosey and Mathison 2009, 146).

Second are accusations of *guruism*, promoting cult-like groups around practitioners (Grant 2019). Some NLP-ers actively encourage this and are listed on a global website for ‘top gurus’, headed by NLP co-founder Richard Bandler (Global Gurus 2024).

A third criticism is *pernicious consumerism*, focus on profit at the expense of ethical and scientific integrity (Grimley 2016; Heap 1989a). For example, it is not unusual to see ‘accredited’ and ‘certified’ NLP practitioner courses offered at the ‘discounted’ price of £26.46, from centers approved by the Institute of Leadership and Management (Centre of Excellence 2023). Even an NLP practitioner certificate from a recognized NLP certification body, The American Board of NLP (ABNLP), can be obtained in only 5 days (Consoul coaching 2023).

Fourth, despite the NLP community’s attempts to establish oversight bodies and other accountability structures such as NLP Leadership Summit (2023), ANLP (2023), IN (2023), INLPTA (2023), IANLP (2023) and ABNLP, these are just window dressing, *lacking a clear or enforceable code of ethics* to police alleged abuses (Grimley 2016; Tosey and Mathison 2009).

Two further ethical criticisms draw on epistemic considerations: fifth, *practicing with NLP is unethical*, because NLP is allegedly theoretically confused, empirically unsupported and pseudoscientific (Elich, Thompson and Miller 1985; Grant 2001; Heap 1989a, 1989b, 2008; Kotera, Sheffield and Van Gordon 2018; Sharpley 1984, 1987; Witkowski 2010). Sixth, NLP-ers are accused of ‘a severe violation of basic ethical scientific principles’ – specifically, the ‘withholding of negative research results’ (Greif 2022, 764); in other words, *scientific fraud*.

We’ll evaluate the purely ethical criticisms in Section 5.1, and the epistemically based ones in Sections 5.2, 5.3 and 6.1.

4.2. Theoretical and Empirical Criticisms

One theoretical criticism is that NLP’s *concepts are imprecisely defined*, such as the notion of representational system (Greif 2022; Witkowski 2010, 64) or ‘neuro-linguistic programming’ itself (Heap 1989a; Sturt 2012). A related theoretical criticism, second, is that it is *unclear what constitutes NLP itself* (de Rijk et al. 2019; Grimley 2016; Kanning 2019; Sturt 2012): it is accused of being a jumble of incoherent ideas and practices, with few if any theoretical interrelationships (Greif 2018, 2022; Heap 1989a).

Some argue, third, that NLP *misconstrues* the psychological and linguistic views that it borrows from, including the notions of modelling (Greif 2018, 380), Pavlovian conditioning, Chomsky’s transformational grammar (Greif 2018, 381, 2022, 766), Miller et al.’s TOTE model (Greif 2022, 768) and Ericksonian hypnosis (Greif 2022, 768).

NLP faces empirical criticism too. First, NLP proponents are said to *make extreme and implausible claims about what NLP techniques can achieve, and about the extent of NLP’s empirical support* (Tosey and Mathison 2009). An example is its co-founder’s claim, on the basis of personal anecdotes, that ‘we can reliably get rid of a phobia in 10 minutes every time’ (Bandler 2008, xix).

Second is the claim that *empirical evidence fails to support NLP*. Some say it actively disconfirms NLP, i.e. lowers the probability that NLP is effective (Grant 2001; Sharpley 1984, 87). But others are more moderate, saying instead that the extant empirical evidence, so far, simply fails to confirm it (Grant 2001; Greif 2018; Sharpley 1987; Sturt et al. 2012).

A third criticism focuses specifically on empirical studies supporting NLP, saying that *they are of low quality* (Kotera, Sheffield and Van Gordon 2018) – either because they are methodologically poor (Briner 2016; Greif 2022), or because they are individual case studies or trials without controls, instead of the randomized control trials (RCTs) seen by many as the gold standard in psychology.

A fourth empirical criticism focuses on three claims that are attributed to NLP, saying that *they are false or disproven*. Although these claims are not central to NLP (see [Section 5.3](#)), they occupy such a large proportion of the critical literature that some detail is needed. First, recall that representational systems for processing sensory information play a role in NLP's account. Some early NLP literature seems to assert:

PRS: Each person has a primary representational system that they use in preference to others.

For example, one might uptake information or construct maps mainly visually, auditorily and so forth (Grinder and Bandler 1976, 9–15).

Second, recall that NLP practitioners seek insight into people's subjective experience by reading minute behavioral cues. These include, some NLP literature asserts, people's non-visual gaze patterns, which NLP calls 'eye-accessing cues': how people move their eyes when not focusing visually on anything particular. Thus the second controversial claim attributed to NLP:

EAC: Eye-accessing cues indicate which representational system a person is accessing at a given moment.

For example, looking up is a clue that you are accessing the visual representational system (i.e. picturing something), and looking below and to the right is a clue that you are accessing the kinesthetic representational system (e.g. feeling something).

Third, Wiseman et al. (2012) attribute to NLP a third claim:

EAC-Lying: Eye-accessing cues can indicate whether a person is lying.

More specifically, that looking up-left indicates truth-telling, whereas looking up-right indicates lying.

A substantial literature criticizes these attributed claims on empirical grounds (Elich, Thompson, and Miller 1985; Heap 1988; Sharpley 1984; Witkowski 2010). Now, science has room for controversy, but the criticism of these claims goes farther, saying that affirming them makes NLP bad science or even pseudoscience. However, it is far from clear that NLP even endorses them, as we'll see below.

4.3. Pseudoscience?

Many critics claim that NLP is not just bad science but pseudoscience. Many of these criticisms fall short as given, but we can extract a list of behaviors that, if NLP does engage in them, would be pseudoscientific.

One target for pseudoscience accusations is the supposed imprecision of many NLP concepts and the domain of NLP itself. Some think that imprecision suffices for pseudoscience (Greif 2022). But we demur. Imprecise concepts and domain-demarkation, by themselves, are not necessarily violations of scientific standards. Much science starts with an unclear, proto-scientific stage, which hopefully recedes as empirical research and reflection inform each other. If clarity does not improve over time, the label 'bad science' may become appropriate.

But even if imprecision is not inherently pseudoscientific, it can be exploited in pseudoscientific ways. You might respond to criticism by defining your concepts or research domain on the fly and accusing the critic of misconstruing it (Boudry 2022). And conceptual clarity needs trained eyes to recognize, making it easy to hoodwink non-academics with scientific-sounding jargon. This *exploiting* of unclarity, we take it, is what may concern Greif (2022). But in this case we need an argument that NLP-ers, particularly among the science-minded, are guilty of it. They may or may not be; critics have not shown us.

Second, NLP has been called pseudoscientific on etiological grounds. Greif (2018) suggests that NLP developed from retroactive explanations of observations and is thus improperly theoretically constructed. But we suggest that post-hoc theory development does not suffice for pseudoscience. It does not even suffice for bad science. Most scientific frameworks arise from a combination of observation and hypothesizing.⁴ Which takes priority when is context-dependent. What matters is

just that, once we have a hypothesis, we should be ready to test predictions from it to the extent that it allows.

But a nearby form of behavior is pseudoscientific, namely a *continued pattern of post-hoc explanation* without making or testing predictions (Boudry 2022). Whether NLP does *this* is what bears on its scientific respectability. Critics must show us.

Third, NLP has been called pseudoscientific because of its founders' rejection of psychological theory and 'psychotheology' in favor of 'what works' (Greif 2018; Kanning 2019) – this refers to NLP's aim to model what excellent psychotherapists do instead of what they say, reflecting an interest in their tacit knowledge instead of traditional psychological theorizing. Does this make NLP pseudoscientific? Again, not necessarily. First, some argue that what works, rather than accuracy as such, is in fact what science seeks (van Fraassen 2008).⁵ Criticizing NLP on these grounds alone is thus merely declaring allegiance with a realist philosophy of science. Second, for NLP's attitude to psychology to be pseudoscientific, it would have to meet two conditions: (i) it must violate ideals of scientific inquiry too badly to count as scientific, and (ii) NLP-practitioners must nonetheless purport to be doing science without concern for accuracy. But it is not clear that either clause holds. Take (i): it is not clear that NLP's attitude to traditional psychology *does* violate scientific ideals. After all, science must be porous enough to sometimes consider new paradigms. Most proposed new paradigms are unfruitful or amount to bad science. So more discussion of NLP's status is needed here, but it does not clearly violate clause (i).

But suppose for argument's sake that NLP's attitude did violate scientific ideals. This would not suffice to make NLP pseudoscientific, for we must still establish the second conjunct – that NLP *claims* to be doing science without concern for accuracy. Does it? It depends on whom you ask. Science-minded NLP-ers definitely claim to be doing (or aspiring to do) science and do so in good faith concerned for accuracy. They therefore open themselves to charges of bad science but not pseudoscience. What about the ascientific NLP-ers, who leave NLP's scientific status to others? As long as they limit their claims to the anecdotal (e.g. 'in my clinical experience NLP interventions tend to satisfy my clients and bring good results'), they are not stepping into the realm of science and are thus not candidates for pseudoscience.

Matters differ for the NLP-bullshitters. They sometimes do claim scientific status for NLP techniques where such status has not yet been established (e.g. Bandler 2017; see Section 4.2 above) and do so unconcerned for accuracy, thus meeting clause (ii). Because of this they also meet clause (i), since claiming more scientific status than you are entitled to is an egregious violation of scientific norms. Thus criticizing psychological theory does not suffice to make you a pseudoscientist, but making un evidenced and exaggerated claims to scientific status without regard to accuracy does. Science-minded NLP-ers and ascientific NLP-ers are not pseudoscientists but NLP-bullshitters are.

In conclusion, of the criticisms we considered, only the following count as accusations of pseudoscience: first, exploiting unclarity by obfuscating when criticized, second, regularly constructing post-hoc explanations without testing them, third, bullshitting about science: unwarrantedly claiming to be doing science (or providing a science-based service) without concern for accuracy. Whether any NLP-ers are guilty of the first two violations, we don't have space to establish here. It is uncontroversial, in contrast, that a subset of the NLP community, the bullshitters, are guilty of the third pseudoscience accusation. This is likely a major source of the negative whiff that NLP has in psychology.

Let's see how science-minded NLP-ers might respond.

5. Bad Gatekeeping?

5.1. Concessions: NLP Must Do Better

Science-minded NLP-ers concede many abovementioned concerns (Gray 2022; Grimley 2016, 2019; Linder-Pelz 2010; Tosey and Mathison 2009; Wake, Gray and Bourke 2013). They acknowledge that

NLP is often misused to manipulate and promote guruism, that profit-motivation can distort quality or integrity, that NLP lacks meaningful accountability structures (de Rijk and Parker 2022; Grimley 2019) and that the factionalization of the NLP community make improvements unlikely soon (Grimley 2016). NLP concepts need sharpening and the field itself better definition, indeed standardization (Gray 2022; Grimley 2019). *Some* empirical work supporting NLP (by no means all!) is substandard. NLP practitioners have made extreme and implausible claims about what NLP can achieve and its empirical support (Tosey and Mathison 2009). Science-minded NLP-ers recognize that these problems rightly impact NLP's perception in psychology, threatening to relegate NLP to pop psychology at best, bad science or pseudoscience at worst. Moreover, after 48 years of NLP (taking Bandler and Grinder 1975b as the start) this situation has not changed significantly. Science-minded NLP-ers are so proportionally few that they are arguably not representative of NLP (Grimley 2016, 2019; Tosey and Mathison 2009). In 2009, Tosey and Mathison concluded that NLP is at a crossroads (188–195), a prognosis that de Rijk and Parker repeat thirteen years later (2022, 241).

But do *all* the criticisms from psychology apply? Even the science-minded NLP-ers who vocally criticize the NLP community strongly resist the claim that NLP *as such* is bad science, let alone pseudoscience. Many are converging around the view that NLP is *proto-scientific* – it *could* devolve into pseudoscience if these concerns are not addressed, but it has the potential to become science proper. Moreover, science-minded NLP-ers *accuse psychology of bad gatekeeping* (Arroll and Henwood 2017b; de Rijk, Gray and Bourke 2022; Parker 2022a; Wake, Gray and Bourke 2013). We'll turn to these concerns.

5.2. Bias

Psychology exhibits bias against NLP; we'll outline three types.

One is *rhetorically loaded negative language and associations*. Here are some representative examples. In the journal of the Trade Association of German Psychologists, Kanning (2019, our translation) describes NLP with such terms as 'myth' (11), 'absurd' (15) and entertaining 'fantasies of omnipotence' (14). In a LinkedIn post, Devlin claims to establish, in just 435 words, that NLP is 'neurobollocks' because it has 'zero scientific backing' (2023). Such language is even common in peer-reviewed literature. In a paper widely cited among NLP-critics, Witkowski (2010) calls NLP a 'pseudoscientific farce' (64) and a 'cruel deception' (64), concluding that 'my analysis leads *undeniably* to the statement that NLP represents *pseudoscientific rubbish*, which should be *moth-balled forever*' (64, italics added); Passmore and Rowson (2019, 60) cite this verbatim as if such hyperbole were acceptable in academic settings. Greif (2018) calls NLP's founders pseudoscientists *before* arguing that they are (379). Passmore and Rowson (2019) frame NLP in association with the kooky-sounding 'angel treatment', 'chrysal healing' and 'dolphin mental health therapy' (69), while Greif (2018) frames it with the emotive *non sequitur* of the massacre in the religious Jonestown sect (377). These critics do not even pretend to neutrality.

One might object that derisive language, even in academic contexts, can be appropriate when deserved; think of Holocaust denial. Perhaps so. We are not advocating a blanket 'neutral tone-of-voice' norm. We are rather arguing that such language is certainly not appropriate if the target has not been shown conclusively to merit it – where the bar for conclusiveness is high. For such language can *generate a perception of targets meriting it even if they have not been shown to*. Scientists who opt for emotive rhetoric must use it with caution; otherwise the epistemic trust crucial to science can devolve into groupthink and damage scientific integrity. Why the rhetoric described here is a case in point is elaborated in Sections 5.3 and 5.4.

Second, bias can be documented in *academic and professional offerings being rejected only due to association with NLP*. Here are representative examples. Grimley's presentation of his coaching approach in (Shams 2022) was stonewalled by peer reviewers for mentioning NLP as one of many inspirations. One reviewer reasoned that NLP is a modality that 'many psychologists would consider has little concrete evidence of efficacy'. This justification is *prima facie* understandable, given the

abovementioned concerns, but when Grimley removed NLP and re-framed the chapter ‘pluralist coaching’ with few other changes, it was accepted. Evidence seems to have played no role here.

Science-minded NLP-ers want to provide this ‘concrete evidence of efficacy’, but many who try report repeated rejection. Why? Because it is NLP that they are studying. For example, Arroll et al. (2017a) submitted an RCT-study exploring the effectiveness of an NLP phobia-cure pattern for *The International Journal of Psychiatry in Medicine*; it was rejected with the comment: ‘the case for why “NLP” should warrant our attention after 40 years of failing to produce any evidence is not established . . . I would strip away any reference to “NLP” and focus purely on calling the intervention what it actually is – a visualisation technique’ (Arroll and Henwood 2017b, 25). Psychology’s calls for ‘concrete evidence of efficacy’ for NLP thus do not seem to be made in good faith. It also puts NLP in a double-bind: gatekept against on empirical grounds but barred against providing empirical evidence to prove its merit.

A similar disinterest in evidence appears in the UK’s National Institute of Clinical Excellence (NICE) draft guideline in December 2020 advised against using NLP for myalgic encephalomyelitis or chronic fatigue syndrome: ‘do not offer people with ME/CFS – therapies derived from osteopathy, life coaching and *neurolinguistic programming (for example the Lightning Process)*’ (italics added); this is despite published studies, including RCTs, demonstrating the promise of the 3-day Lightning Process program (e.g. Crawley et al. 2018; Parker, Aston and de Rijk 2021; Fauske et al. 2021).⁶ Rejecting studies *because they support NLP*, and not because of their quality, is simply bias.

In this climate it is standard for NLP-derived approaches to be given other names, further camouflaging any potential that NLP might be shown to have.⁷ A priori rejection of NLP research may be a self-fulfilling prophecy: the more tightly closed the gate is to NLP, the harder it is to argue that it deserves a place inside.

One might object that bias can be important for gatekeeping. Scientists cannot thoroughly explore every heterodox idea; imagine spending time reviewing grant proposals on the flat-earth hypothesis. And views such as Holocaust denial deserve only derision. Gatekeeping prevents time- and money-wasting, and bias can be a useful heuristic, once sufficient evidence has mounted against a view, to aid quicker gatekeeping. In response, we agree that *gatekeeping* is important. But bias, exhibited in any of the ways we pinpoint, is not the right way to gatekeep. Respecting your evidence is. We have sufficient evidence that flat earthers are mistaken and Holocaust deniers not only mistaken but hate-driven. Do we have sufficient evidence against NLP? As we saw, many psychologists seem to think so. But we’ll see that matters are more complicated. Both theoretical and empirical criticisms of NLP have significant problems.

5.3. *Straw-Manning and Ignorance of NLP*

The psychology literature straw-mans NLP considerably. To *straw-man* a view is to describe it inaccurately, often oversimplifying, in a way that is easier to argue against than the view itself, and then to argue against your inaccurate version instead of against the original view. It is a kind of bait and switch.

Before describing how psychology straw-mans NLP, let’s anticipate an objection from psychologists: surely this ‘straw-manning’ answer is exactly what one would expect from pseudoscientists. When criticized about their unclarity, they turn the tables with complaints about being misunderstood. We may agree that this answer is characteristic of pseudoscience. But it does not follow that anyone who makes it is a pseudoscientist. Straw-manning occurs, and we must judge cases individually.

We’ll illustrate three types of straw-manning of NLP in psychology literature. First is what we’ll call *false exemplification*. Critics often describe NLP as if the claims PRS, EAC and EAC-Lying are its theoretical bedrock, singling them out for critique (e.g. Heap 1989a, 1989b; Passmore and Rowson 2019, 58–59; Sharpley 1984, 1987). But the topics they concern are marginal to NLP itself (see Section 3); key NLP literature has noted this often (de Rijk, Gray and Bourke 2022; Einspruch and Forman 1985; Grimley 2015b; Wake, Gray and Bourke 2013), raising suspicions that critics have not done their homework. Critics’ bibliographies reinforce this suspicion: (Greif 2018) for instance contains

only one foundational NLP text (Bandler and Grinder 1975b) amid 30 references. (Kanning 2019) cites only two key NLP texts, relying otherwise on 10 citations from Cooper's *NLP for Dummies* (2009). Passmore and Rowson's bibliography (2019) includes only three foundational NLP texts among 65 total references, omitting from their lengthy discussion of EAC reference to Dilts's definitive work on eye-accessing cues (1983).

Not only are the topics of these three claims marginal to NLP, *NLP does not endorse them*. This brings us to the second type of straw-manning: *attributing to NLP easy-to-criticize claims that it does not endorse*. Start with PRS, the claim that each person has a primary representational system. Granted, early NLP texts were unclear. Some seemed to endorse it (Grinder and Bandler 1976, 9); others seemed to make the more nuanced claim that people are apt to prioritize particular representational systems not across the board, but in certain contexts (e.g. kinesthetic when jogging, visual when painting) (Grinder and Bandler 1976, 26). But somewhat more recently, science-minded NLP-ers, as well as NLP co-founder John Grinder, have explicitly and conclusively denied PRS (Grimley 2020; Grinder and Pucelik 2013, 214).

NLP does not endorse EAC either, the claim that eye-accessing cues indicate which representational system a person is accessing. There is a claim in the vicinity that many (though not all) NLP-ers do endorse (Dilts 1983):

EAC-Qualified: Eye-accessing cues are a defeasible clue about which representational system a person is accessing at a given moment, which must be evaluated in combination with many other behavioral cues; different people may have different eye-accessing patterns.

EAC-Qualified is much more nuanced than EAC. But critics attribute EAC to NLP (Passmore and Rowson 2019, 61). NLP supposedly holds that eye-accessing cues are infallible clues, identical between individuals and can be read independently of other behavioral cues. This is a straw-man.

What about EAC-Lying, the claim that moving your eyes up-right indicates lying? NLP has always decisively rejected it. Instead, NLP says this: looking up (in any direction) is a defeasible clue that you are accessing the visual representational system, i.e. picturing something. Looking up-*left* is a defeasible clue that your picture is *remembered* (e.g. the color of your door), whereas up-*right* is a defeasible clue that it is *constructed* (e.g. a six-legged Sphinx). These claims, though subject to empirical confirmation,⁸ are worlds apart from EAC-Lying, which says that someone looking up-right is apt to be *deliberately misleading hearers*.

Why do psychologists attribute EAC-Lying to NLP? The main source is (Wiseman et al. 2012). Their reasons? First, 'two well known YouTube videos encouraging lie detectors to adopt this approach have received 30,000 and 60,000 views respectively' (introduction). No links are given to these 'well known' videos, and no connection is made between popular YouTube videos and NLP sources. Second, Wiseman et al. acknowledge that in fact 'the originators of NLP didn't view "constructed" thoughts as lies' (introduction) but they maintain nonetheless that 'this notion has become commonplace, leading many NLP practitioners to claim that it is possible to gain a useful insight into whether someone is lying from their eye-movements' (Wiseman et al. 2012, italics added). Which practitioners are these? We are not told. In the conclusion, Wiseman et al.'s claim changes without comment. There they remove the qualification 'many', generalizing further that the EAC-Lying claim is 'made by NLP practitioners'. But which NLP practitioners specifically do the authors have in mind? They cite only one (R. Gray 1991). It turns out, however, that Gray does not endorse EAC-Lying either. Instead he says (13):

when a client is asked a concrete question – 'Where were you last night?' – eye movement up or over to the right might suggest that he or she is constructing a response, not recalling one. This in itself may indicate valuable lines for further investigation.

Gray does not say 'lie'. Nor does he conclude that the client is speaking falsely, let alone deliberately. 'Construct' here just means 'visualize a picture you have not seen before'. His only conclusion is that more investigation 'may' be worthwhile. Wiseman thus falsely attribute EAC-Lying to Gray and to

'NLP practitioners' generally on the sole basis of this false attribution. They have constructed a straw-man.

Third: *key NLP concepts, and their theoretical context, are misunderstood and falsely represented*. For example, Greif (2018, 380; cf. 2022) accuses NLP of misunderstanding Pavlovian conditioning and distorting Bandura's behavioral modelling 'beyond recognition'. We demur: NLP explains how its notion of anchoring differs from Pavlovian conditioning (Grimley 2013, 87–89) and that NLP-modelling never intended to use Bandura's theory (Grimley 2013, 130). Greif further criticizes NLP for misconstruing' (G. A. Miller, Galanter and Pribram 1960)'s TOTE unit of goal-oriented human action as a general problem-solving method. But this is also mistaken. Grimley (2013, 114–115) notes that TOTE just is a simple relationship between sensory units and that NLP sees itself as extending this to illustrate the structure of behavior (Grimley 2013, 114–115). Science often progresses by using old ideas in new contexts (Hofstadter and Sander 2013, Chapter 8).

Further, Greif criticizes NLP's use of Ericksonian hypnosis, saying it 'reduces in NLP to a fast handshake induction method' (Greif 2022, 768); similarly, Kanning (2019) reduces NLP to 'a few simple psycho-tricks' (2023, no page, our translation). Neither substantiates these claims. Erickson himself had a more positive verdict. In his preface to NLP's book on his hypnotic technique, he called the NLP model 'a much better explanation of how I work than I myself can give' (Bandler and Grinder, 1975a, viii).

On to Chomsky's transformational grammar (TG). Greif expresses puzzlement about how the linguist, NLP co-founder John Grinder, 'could so completely distort' Chomsky 'by interpreting his theory of linguistic syntax psychologically' (Greif 2022, 766). The distortion is Greif's. TG has long been subject to psychological interpretations. An early one was (G. A. Miller 1956), suggesting that TG has the psychological consequence that people take longer to respond to passive-voice sentences than active-voice ones.⁹ Moreover, the classic NLP text drawing on TG says that NLP, although 'inspired by' it, has 'adapted the model . . . for our objectives in therapy' (Bandler and Grinder 1975b, 40). NLP never intended to adopt TG wholesale.

Psychology thus straw-mans NLP: falsely exemplifying it with claims which it does not even endorse and misunderstanding NLP's concepts and theoretical context.

5.4. Empirical Matters

Many empirical criticisms of NLP have problems too. We'll highlight six concerns.

What we'll call the *mistaken-construct* concern arises from the abovementioned false attributions to NLP and from mischaracterization of NLP constructs such as rapport (Wake, Gray and Bourke 2013, 201), pattern recognition (Einspruch and Forman 1985, 592), the meta-model (Einspruch and Forman 1985, 593) and of how NLP therapy works (Einspruch and Forman 1985, 592–3). Of course, the NLP critic can object here too that any pseudo-scientist complains about their supposedly vague account being misunderstood. But we take the foregoing discussion (5.3 and 5.4) to indicate a general carelessness in understanding NLP and refer readers for detail to (Einspruch and Forman 1985) and (Gray et al. 2013, 194–208).

The *poor experiment* concern notes that key studies regarded as disconfirming NLP techniques were poorly designed or executed. For example, often experimenters were not trained NLP practitioners, so techniques were inappropriately performed (Einspruch and Forman 1985; Gray et al. 2013). And often context was ignored, for example with researchers locking people's heads in restraints to gauge their eye movements and thus failing to understand that eye movements, for NLP, are embedded in natural conversational dynamics (Einspruch and Forman 1985, 592). Negative results from such studies are unsurprising, but they are not about NLP.

Another concern, applying to meta-analyses, is *uncharitable interpretation of NLP-supporting research*. For example, Passmore and Rowson (2019) gloss Sturt's (2012) statement that 'there is currently insufficient evidence to recommend use of NLP for any individual health outcome' (e763) as 'damaging evidence' (60). This is certainly one interpretation. But there are others. One is that better research is needed. This seems to be Sturt's own conclusion, since she¹⁰ is now co-leading research

into the NLP-based Reconsolidation of Traumatic Memories project, funded at King's College London by the Forces in Mind Trust (de Rijk et al. 2023; Sturt et al. 2022, 2022b).

Similarly, Greif (2022) says that (Zaharia, Reiner and Schütz 2015)'s results supporting a certain NLP technique 'could hardly be more devastating' (763). The reason is unclear. He argues that (Zaharia et al.)'s mean Cohen D effect size of .51, a medium size, comes from their leaving in the study by Pourmansour (1997), which Greif claims should have been removed as a high individual value. But is unclear why Greif thinks this. First, it can be appropriate to leave in high values, since extreme values can reflect true variations in distribution (Osborne and Overbay 2004). Second, although it is possible that the Pourmansour study is a high individual value, Zaharia et al. kept it in because their statistical analyses gave no indication that it would introduce a publication bias. To exclude publication bias they used an Egger analysis, a funnel plot and a sensitivity analysis which omits each study once and calculates the overall effect size again to account for large negative or positive effects in any one study. So why does Greif single out the Pourmansour study as needing to be removed, when Zaharia et al.'s statistical analysis did not involve a publication bias? Greif does not say. To substantiate his claim he needs an argument. Lacking one, it would seem that Zaharia et al.'s Cohen coefficient of .51 stands.¹¹ However, even if Greif were right that the Pourmansour study should have been removed, we would be left with a Cohen coefficient of .39, still a good enough score to fall in the average range. Either way Greif's 'devastating' gloss thoroughly misrepresents their results.

But Greif adds insult to injury. He accuses NLP of a 'severe violation of basic ethical scientific principles' constituted by 'withholding of negative research results' (2022, 764) – in other words, scientific fraud. It turns out that his sole basis for these harsh generalizing words about NLP is *his analysis of Zaharia et al.'s results*. We just saw that he leaves his analysis – and thus this potentially libelous accusation – unsubstantiated. This is potentially a serious ethical matter. But for argument's sake, let's suppose counterfactually that Greif's gloss of Zaharia et al.'s results as 'devastating' were apt. This would still not justify his accusation of scientific fraud. Greif has not excluded the possibility of a simple mistake, either Zaharia et al.'s or his own. Further, even if this accusation were apt, Greif's generalization to all of NLP is unsubstantiated. This is just one example of uncharitable, indeed biased and possibly unethical, interpretations of NLP-supporting results.

A related concern is *cherry-picking negative results about NLP and overlooking positive ones*. For example, Passmore and Rowson (2019, 60) conclude that 'no robust evidence exists to support its [NLP's] use within health settings'. Really? They draw this universal conclusion from a single meta-analysis (Sturt 2012), overlooking potentially promising results of applying NLP to allergic reactions (Witt 2008) and neuro-immunological conditions such as chronic fatigue syndrome (reviewed in Parker and Phil 2022b).

The first four concerns yield a fifth, which (Gray et al. 2013, 202) memorably call *intellectual telephone*, after the game where a message becomes distorted when passed on. Psychology's narrative about NLP is shaped by empirical studies and meta-analyses. When these works exhibit the problems just discussed, the narrative becomes increasingly negatively skewed. Later research uncritically assumes that earlier research was conceptually and methodologically sound. A garbled message results. A recent example is (Passmore and Rowson 2019). Despite extensive discussion of problems in NLP-critical research (Einspruch and Forman 1985; Gray et al. 2013), they uncritically accept earlier negative studies about NLP. They uncritically affirm Wiseman et al.'s study on EAC-Lying (61; see above 5.3). They overlook important limitations to (Ahmed 2013), such as the author's declining to state where he obtained his NLP certificate (a major worry given the proliferation of cheap-and-easy practitioner certificates), and having only one supposedly trained observer (himself!) rather than many, therefore being unable to generate an inter-observer statistical coefficient for the reliability of his observations. These are just representative examples of intellectual telephone.

Sixth is the issue of *qualitative vs. quantitative research*. Several promising NLP applications are being researched quantitatively (when funding can be found despite anti-NLP bias, see Section 5.3), with promising initial results (Gray and Bourke 2015; Gray, Budden-Potts and Bourke 2017; Gray,

Davison and Bourke 2021; Gray et al. 2020; Parker 2022b; Parker, Aston and Finch 2018). To date, however, much NLP-supporting research is qualitative. Some critics dismiss it on these grounds alone (e.g. Heap 1987, 105). Granted, it is common in psychology to value quantitative over qualitative research (Ahmad et al. 2019, 2828), but there are prominent calls to recognize its weaknesses amid the complexity of human cognition and behavior, and to emphasize mixed-method research (e.g. EAP 2021). This is especially so for modalities, such as NLP, hypnosis or Freudian psychotherapy, that involve phenomenology or the unconscious. It is thus increasingly untenable in today's research environment to dismiss qualitative NLP-favoring research *only* because it is qualitative. But beyond this, it is downright mysterious to *dismiss NLP research for using qualitative methods while endorsing them elsewhere*. This is what NLP-critic Briner (2016) does. He criticizes the NLP-supporting meta-analysis of Zaharia et al. (2015) for lacking RCTs (falsely, in fact, as they discuss six!), but his own area of interest, evidence-based management, calls the 'blind adoption' of 'evidence-based management ... that prizes randomized control trials and meta-analyses above all other kinds of research evidence' a 'danger' (Briner, Denyer and Rousseau 2009, 20). This selectivity appears at best unprincipled, at worst hypocritical.

The concerns arising from false constructs, experimental poverty, uncharitable reading, cherry-picking, intellectual telephone and unprincipled dismissal of qualitative research are further reasons to worry about the quality of psychology's gatekeeping against NLP.

6. Lessons

Let's draw some conclusions about the NLP case and about gatekeeping in science generally.

6.1. NLP and Psychology

NLP faces many charges that even (science-minded) NLP-ers concede. It is often misused, including by self-styled gurus, for manipulative ends, profit motivation can distort quality and integrity, and it lacks accountability structures. NLP's concepts need sharpening, the field needs definition and some poor empirical work has been done on NLP (though none of these criticisms applies in the ways many critics assert). The pseudoscience accusation sticks to *the NLP bullshitters* but not to ascientific NLP-ers or the science-minded.

Other pseudoscience charges against NLP cannot be established without further discussion: NLP-ers – particularly the science-minded – have not been shown to exploit unclarity by obfuscating or to regularly construct post-hoc explanations without testing them. Also unestablished is the potentially libelous charge of scientific fraud against science-minded NLP-ers. Despite the ball-and-chain of NLP bullshitters, there is good reason, including promising initial studies, to categorize NLP as a proto-science with significant potential as a behavioral-change modality. This means recognizing, first, that NLP needs more research, both theoretical and empirical, but second, that it shows enough promise to make such research a good bet. NLP's theoretical account, and its efficacy as a behavioral-change modality, are thus presently hypotheses with significant initial promise but meriting, and requiring, further investigation.

Let's return to an ethical worry voiced in Section 4.1: surely it is wrong to establish training programs and license practitioners if your modality can be labelled as nothing better than proto-scientific. This worry is often apt – but not always. Practices often do not develop in the textbook fashion of first having their hypotheses tested with sufficiently large and diverse samples and *then* applied. Few would get off the ground – including cognitive behavioral therapy (CBT), which many call the gold standard of behavioral-change interventions because of its robust testing. Even CBT, however, is far from evidentially secure. David et al. (2018) note that CBT, as a progressive research program, does not presently claim evidence-based status, due to weak controls among other issues. In defense of using CBT, they cite the meta-analysis by (Cuijpers et al. 2016), with its stronger controls, as an accurate estimate of CBT's efficacy. Yet even Cuijpers et al. are hesitant, noting significant limitations in their research. Another example is

Miller and Rollnick's 'motivational interviewing' (2023), which began with anecdotal evidence, only later receiving scientific attention. Given the difficulty of securing research funding (especially, but not only, as a clinician without a university affiliation), small-scale, anecdotally supported clinical practice is often one of the only early testing grounds available for behavioral-change interventions. Hence the mere fact of applying a modality before it has been tested to a certain standard is not inherently unethical. To remain ethical, however, it is important that practices developing in these bottom-up ways be honest about their state of knowledge. And as long a practice is merely proto-scientific, it should take measures toward greater scientific robustness. This is what science-minded NLP-ers are doing.

Will they succeed? One thing this depends on is gatekeeping in psychology. This has involved unceasing and insufficiently justified bias, straw-manning and numerous problems with empirical work. Even given NLP's problems, this is not exemplary gatekeeping. Let's use our account of good gatekeeping (Section 2) to see why.

Value of ideals that the gatekeeping seeks to preserve. Ideals of scientific inquiry are valuable, and psychologists who gatekeep against NLP claim to want to preserve them. Charity demands taking these gatekeepers at their word. But we must also remember that gatekeeping involves group dynamics, and we saw evidence of their influence above, not least in the bias, negative interpretations, cherry picking and double standards vis-à-vis NLP research. It is thus legitimate to ask whether there are other, unscientific and unspoken, ideals entering in for some gatekeepers. These might be bad ideals, e.g. careerism. But they might also be good, e.g. distaste for the unethical trappings of NLP-bullshitters. Whether good or bad, however, any non-scientific ideals should be declared, not brushed under the carpet of science.

Fitting gatekeeping method. Are psychology's gatekeeping methods against NLP fitting to scientific ideals? Many are not. The search for *understanding* is hindered through biased language, blacklisting NLP research, straw-manning NLP and the problems in empirical studies. These things, being distortions, also hinder the search for *accuracy*, and since they skew our picture of the total evidence, also damage *empirical adequacy*. The ideal of *substantiating your results* is violated, given that pejorative rhetoric and association with NLP replace arguments and the flaws in empirical studies. The ideals of *critical peer feedback* and *epistemic self-reflection* are violated by the bias against NLP. The ideals of *systematicity* and *open-endedness* are violated by bias, straw-manning and cherry-picking. We may conclude that many of the methods that psychology uses to gatekeep against NLP are *not* fitting to the ideals of science but violate them significantly.

Successful gatekeeping? Because psychology's gatekeeping against NLP violates scientific ideals, it falls at the first hurdle for success, which is *preserving scientific values*. We charitably granted that psychologists aim to do this but saw that their methods instead violate them. What about the second marker of successful gatekeeping, appropriately balancing leniency and tightness? Since psychology aims to keep NLP out, the main question is whether it has been *too tight*: is NLP excluded despite having the potential to contribute to psychology? We suggest yes. It is arguably a proto-science, with the potential – given research funding and respect – to become a science. But it also has potential to dissolve into pop psychology or pseudoscience. One risk factor is the NLP-bullshitters. Another, however, is psychology itself, whose bad gatekeeping could be a self-fulfilling prophecy – pressuring the science-minded to dissociate from the NLP label and motivating ascientific NLP-ers to dissociate from a science they have reason not to trust.

6.2. Gatekeeping in Science Generally

Now for some lessons about gatekeeping in science.

One is that social dynamics matter: *poor gatekeeping damages science's credibility among non-scientists and would-be scientists*. This has an important corollary for the term 'pseudoscience'. It works as an epithet because science is respected. If science loses respect, for example by bad gatekeeping, then this term may lose its punch, as it has for one NLP-er (Derks 2023).

Another lesson concerns a tendency, discernible in psychology's gatekeeping against NLP, that we may call *science fundamentalism* (Dormandy [forthcoming](#)). Where fundamentalism is the tendency to accept a framework uncritically, with strong confidence, and resist counterevidence, science fundamentalism is the tendency to take a fundamentalist attitude toward *science* – either in general, toward your discipline or toward some consensus position or methodology. We see this in the dichotomous 'us-versus-them' attitude promoted by psychology's unprofessional rhetoric. We see it in the rigidity of psychologists who dismiss NLP theory (without understanding it) for existing catty-corner to its own. And we see in it the hypocrisy of dismissing NLP studies for being qualitative while approving of such studies elsewhere. Psychology's gatekeeping against NLP arguably exhibits science fundamentalism.

A third lesson concerns a priori exclusion: when do you count as having enough evidence against would-be members of your scientific community to exclude them a priori, without considering new evidence? Many psychologists clearly think they do for NLP, and NLP bullshitters make this sadly understandable. But scrutinizing the evidence for this judgment call, psychology's knee-jerk response looks overhasty. The lesson is that discerning whether science is entitled to exclude a framework a priori is sometimes harder than it seems. We must avoid time-wasting in science, but we must also be open to insights from left field.

Finally, we need to think more about trust in science (Dormandy [2020](#), 18). Scientists work by trustingly building on each others' results (Frost-Arnold [2013](#)). But collegial trust can tip into groupthink (Janis [1972](#); Oreskes [2019](#), chap. 2). This arguably happened in psychology's gatekeeping against NLP. Unprofessional rhetoric, one sign of groupthink, is easy to spot once you are sensitized to it. But identifying substandard conceptual and meta-analyses takes more work, sometimes prohibitive when exploring outside your field. If trusting others' results – a *sine qua non* for science – could go so awry as to produce generations of skewed meta-analyses about NLP, we might need to re-think conditions for rational trust in science generally.

7. Conclusion

We have examined a complex case of gatekeeping in science and discovered faults on all sides. Psychology's gatekeeping against NLP is understandable due to the public presentation of some NLP-ers, but it is bad gatekeeping. It rather reflects flat-out bias, straw-manning and highly problematic empirical analyses. For this reason alone, psychology's gatekeeping is bad – it violates the ideals it seeks to protect. This is not a vindication of NLP, however. Gatekeeping can be done badly, but bad gatekeeping can be serendipitous – the thing being kept out might really deserve to be kept out. Our arguments suggest that NLP is theoretically much more sophisticated and empirically better supported than its critics maintain. At best, however, we think that they support classifying NLP as a proto-science. It is not a full-fledged science yet, but it has the strong potential to be. If only the NLP-bullshitters would let it (de Rijk and Parker [2022](#)), and if only science would let it.

We extracted some lessons about gatekeeping in science. One is that poor gatekeeping in science damages science's credibility, making pseudoscience sound less bad. Another is that science fundamentalism is alive and well and we should beware of it. We learned that whether someone should be excluded a priori is often harder to tell than we may think. Finally, science is utterly dependent on trust among scientists, and that trust must be earned.

Notes

1. This is not to be understood as an attempted definition of science, just a characterization of some of its important features.
2. To describe this nesting of units of representation, NLP draws on the TOTE model (Miller, Galanter and Pribram [1960](#)); see (Dilts et al. [1980](#), 26–40).

3. Witkowski cites (Nocross, Koocher and Garofalo 2006)'s Delta pole, but neglects to mention the authors' caution in interpreting their results, including possible selection bias (Nocross, Koocher and Garofalo 2006, 518).
4. Miller and Rollnick's 'motivational interviewing' technique (W. Miller and Rollnick 2023) for instance started primarily with observations which it theorized around.
5. Thanks to a reviewer for *Social Epistemology*.
6. Despite (Fauske et al. 2021)'s its being published by an MDPI journal, it documents promising results with an expert authorship lineup and we detected no methodological problems.
7. Examples include Clean Language (Cairns-Lee, Lawley and Tosey 2021, 2022), RTM (Gray and Bourke 2015; Gray, Budden-Potts and Bourke 2017; Gray, Davison and Bourke 2021; Gray et al. 2020; Tylee et al. 2017), 7Cs (Grimley 2021a, 2021b), the Lightning Process (Parker, Aston and Finch 2018), and Social Panorama (Derks, Oetsch and Walker 2016; Derks and Alexandru 2016).
8. See (Gray 2022, 206) for the current state of evidence about EAC.
9. Thanks to John Grinder.
10. Passmore and Rowson (2019, 59) misgender Jackie Sturt with 'Sturt and his colleagues', another indication of unfamiliarity with their topic.
11. Thanks to Cătălin Zaharia.

Acknowledgments

Thanks for helpful comments go to Scott Hill, Jane Mathison, Phil Parker, Cătălin Zaharia and the anonymous reviewers for *Social Epistemology*. Thanks for helpful discussion go to Michael Dormandy, John Grinder and the participants of the 2nd Interdisciplinary Colloquium on Personality Growth at Fürstenried Palace, Munich. Thanks also to Jonathan Matheson for editing this special issue. Katherine Dormandy's work was generously funded by the Psychological Cross-Training Grant ID 62700 from the John Templeton Foundation. The John Templeton Foundation may not hold the views expressed in this article.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the John Templeton Foundation [62700].

Notes on contributors

Katherine Dormandy is University Professor of Philosophy at the University of Innsbruck. She has also held positions at the Humboldt University of Berlin, the University of Leipzig, the University of Saarbrücken and the Munich School of Philosophy, and did her graduate work at Oxford University. She is the University of Innsbruck's Ombudsperson for Good Scientific Practice in the Humanities and Law and heads the Innsbruck Center for Philosophy of Religion. Her research interests include epistemology (traditional, formal and feminist), moral psychology philosophy of science, philosophy of psychology and philosophy of religion.

Bruce Grimley is a Chartered Psychologist in 2 divisions of the British Psychological Society. He is independently employed (Achieving Livest Ltd) and currently evaluates post-graduate work at UCN. He regularly writes on the topic of NLP, having authored 2 books and 9 chapters showing how NLP draws upon psychology. His speciality as a psychologist of some 30 years in standing is one to one work, whether coaching, counselling or psychotherapy. His PhD research asked the compelling question: "What is NLP?" with answers being published in the International Coaching Psychology Review.

ORCID

Katherine Dormandy  <http://orcid.org/0000-0003-3947-6594>

References

- Ahmad, S., S. Wasim, S. Irfan, Z. Farheen, S. Gogoi, and A. Srivastava. 2019. "Qualitative v/s Quantitative Research." *Journal of Evidence Based Medicine and Healthcare* 6 (43): 2828–2832. <https://doi.org/10.18410/jebmh/2019/587>.
- Ahmed, K. Z. 2013. "Lying Eyes: The Truth About NLP Eye Patterns and Their Relationship with Academic Performance in Business and Management Studies (MBA)." *International Journal of Business & Management* 8 (23). <https://doi.org/10.5539/ijbm.v8n23p67>.
- ANLP (Association of NLP). Accessed December 12, 2023. <https://anlp.org/>.
- Arroll, B., and S. Henwood. 2017b. "NLP Research, Equipose and Reviewer Prejudice." *Rapport* 54: 24–26. Accessed March 6, 2024. [https://www.achieving-lives.co.uk/files/Files/suzanne%20paper%20on%20fear%20of%20heights%20\(2\).pdf](https://www.achieving-lives.co.uk/files/Files/suzanne%20paper%20on%20fear%20of%20heights%20(2).pdf).
- Arroll, B., S. Henwood, F. I. Sundram, D. W. Kingsford, V. Mount, S. P. Humm, and H. B. Wallace. 2017a. "A Brief Treatment for Fear of Heights: A Randomized Controlled Trial of a Novel Imaginal Intervention." *The International Journal of Psychiatry in Medicine* 52 (1): 21–33. <https://doi.org/10.1177/0091217417703285>.
- Bandler, R. 2017. "What is NLP?" Accessed December 18, 2023. <https://www.youtube.com/watch?v=WWMiC9VwqaA&t=3s>.
- Bandler, R., and J. Grinder. 1975a. *Patterns of the Hypnotic Techniques of Milton H. Erickson, MD*. Vol. 1. Capitola: Meta Publications.
- Bandler, R., and J. Grinder. 1975b. *The Structure of Magic: A Book About Language and Therapy*. Palo Alto: Science and Behavioural Books.
- Bandler, R., and J. Grinder. 1979. *Frogs into Princes: Neuro Linguistic Programming*. Moab: Real People Press.
- Bodenhamer, R. G., and L. M. Hall. 1999. *The User's Manual for the Brain*. Vol. 1. Carmarthen: Crown House Publishing.
- Bostic St Clair, C., and J. Grinder. 2001. *Whispering in the Wind*. Scotts Valley: J and C Enterprises.
- Boudry, Maarten. 2022. "Diagnosing Pseudoscience – By Getting Rid of the Demarcation Problem." *Journal for General Philosophy of Science* 53 (2): 83–101. <https://doi.org/10.1007/s10838-021-09572-4>.
- Briner, R. 2016. LinkedIn Discussion. Accessed December 12, 2023. [ow.ly/jX7x50hWxYh](https://www.linkedin.com/feed/update/urn:li:activity:6666666666666666666).
- Briner, R. B., D. Denyer, and D. M. Rousseau. 2009. "Evidence-Based Management: Concept Clean-Up Time?" *The Academy of Management Perspectives* 23 (4): 19–31. <https://doi.org/10.5465/AMP.2009.45590138>.
- Burgess, F. 2014. *The Bumper Book of Modelling. NLP Modelling Made Simple*. Clitheroe: Kilmonivaig Publishing.
- Cairns-Lee, H., J. Lawley, and P. Tosey. 2021. "Enhancing Researcher Reflexivity About the Influence of Leading Questions in Interviews." *Evaluation Practices* 58 (1): 164–188. <https://doi.org/10.1177/00218863211037446>.
- Cairns-Lee, H., J. Lawley, and P. Tosey, eds. 2022. *Clean Language Interviewing: Principles and Applications for Researchers and Practitioners*. Emerald Publishing. <https://emeraldinsight.com/page/detail/clean-language-interviewing/?k=9781801173315>.
- Centre of Excellence. 2023. Accessed December 13, 2023. <https://www.centreforexcellence.com/>.
- Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.
- Consoul coaching. 2023. Accessed December 13, 2023. <https://consoulcoaching.com/course-dates/>.
- Cooper, L. 2009. *NLP in the Profession for Dummies*. Weinheim: Wiley-VCH.
- Crawley, E. M., D. M. Gaunt, K. Garfield, W. Hollingworth, J. A. C. Sterne, L. Beasant, S. M. Collin, N. Mills, and A. A. Montgomery. 2018. "Clinical and Cost-Effectiveness of the Lightning Process in Addition to Specialist Medical Care for Paediatric Chronic Fatigue Syndrome: Randomised Controlled Trial." *Archives of Disease in Childhood* 103 (2): 155–164. <https://doi.org/10.1136/archdischild-2017-313375>.
- Cuijpers, P., I. A. Cristea, E. Karyotaki, M. Reijnders, and M. J. Huibers. 2016. "How Effective Are Cognitive Behavior Therapies for Major Depression and Anxiety Disorders? A Meta-Analytic Update of the Evidence." *World Psychiatry* 15 (3): 245–258. <https://doi.org/10.1002/wps.20346>.
- David, D., I. Cristea, and S. G. Hofmann. 2018. "Why Cognitive Behavioral Therapy Is the Current Gold Standard of Psychotherapy in Psychiatry." Accessed October 23, 2018. <https://doi.org/10.3389/fpsy.2018.00004>.
- DeLozier, J., and J. Grinder. 1987. *Turtles All the Way Down*. Scotts Valley: Grinder and Associates.
- de Rijk, L., G. Campbell, N. Greenberg, and D. Murphy. 2023. "Exploring the Feasibility of Providing Reconsolidation of Traumatic Memories (RTM) Therapy Training to Clinicians in Ukraine – A Service Evaluation." *European Journal of Trauma & Dissociation* 7 (2): 100326. <https://doi.org/10.1016/j.ejtd.2023.100326>.
- de Rijk, L., L. Derks, B. Grimley, and J. Hollander. 2019. "The Evidence for NLP." *International Coaching Psychology Review* 14 (1): 5–30. <https://doi.org/10.53841/bpsicpr.2019.14.1.5>.
- de Rijk, L., R. Gray, and F. Bourke, eds. 2022. *Neurolinguistic Programming in Clinical Settings: Theory and Evidence-Based Practice*. London: Routledge.
- de Rijk, Lisa, and Phil Parker. 2022. "The Future of NLP Research." In *Neurolinguistic Programming in Clinical Settings: Theory and Evidence-Based Practice*, edited by Lisa de Rijk, Richard Gray, and Frank Bourke, 236–244. London: Routledge.
- Derks, L. 2023. "How I Became A Pseudophile." *INZICHT* 79. <https://www.msp-academy.com/wp-content/uploads/2023/10/How-I-became-a-pseudophile-PDF-final.pdf>.
- Derks, L. A. C., and A. I. Alexandru. 2016. "On the Emergence of Mental Space in Psychology: Interview with Lucas Albert Charles Derks." *Europe's Journal of Psychology* 12 (2): 304–308. <https://doi.org/10.5964/ejop.v12i2.1186>.

- Derks, L. A. C., W. O. Oetsch, and W. Walker. 2016. "Relationships are Constructed from Generalized Unconscious Social Images Kept in Steady Locations in Mental Space." *Journal of Experiential Psychotherapy* 19 (1): 73.
- Devlin, J. 2023. "LinkedIn Post." Accessed December 18, 2023. https://www.linkedin.com/search/results/content/?keywords=joseph%20devlin&origin=CLUSTER_EXPANSION&sid=QSu.
- Dilts, R. 1983. *Roots of Neuro-Linguistic Programming*. Capitola: Meta Publications.
- Dilts, R. 1998. *Modeling with NLP*. Capitola: Meta Publications.
- Dilts, R., J. Grinder, R. Bandler, and J. DeLozier. 1980. *Neuro-Linguistic Programming. Volume 1. The Study of the Structure of Subjective Experience*. Capitola: Meta Publications.
- Dormandy, Katherine. 2020. "Introduction: An Overview of Trust and Some Key Epistemological Applications." In *Trust in Epistemology*, edited by Katherine Dormandy, 1–40. New York: Routledge.
- Dormandy, Katherine. Forthcoming. "Science Fundamentalism." In *The Epistemology of Expert Judgment*, edited by Peter Brössel, Anna-Maria Asunta Eder, and Thomas Grundmann. New York: Routledge.
- EAP. 2021. "Position Paper on the Proper Nature and Policy Applications of Psychotherapy Research. European Association of Psychotherapy Science and Research Committee Working Group." Accessed December 21, 2023. <https://www.europsyche.org/app/uploads/2021/04/EAP-Research-Statement-March-13-2021-adopted.pdf>.
- Einspruch, E. L., and B. D. Forman. 1985. "Observations Concerning Research Literature on Neuro-Linguistic Programming." *Journal of Counselling Psychology* 32 (4): 589–596. <https://doi.org/10.1037/0022-0167.32.4.589>.
- Elich, M., R. W. Thompson, and L. Miller. 1985. "Mental Imagery as Revealed by Eye Movements and Spoken Predicates: A Test of Neurolinguistic Programming." *Journal of Counseling Psychology* 12 (4): 622–625. <https://doi.org/10.1037/0022-0167.32.4.622>.
- Fauske, L., Ø. S. Bruland, A. A. Dahl, A. Myklebostad, and S. E. Rem. 2021. "Does the Lightning Process Training Programme Reduce Chronic Fatigue in Adolescent and Young Adult Cancer Survivors? A Mixed-Methods Pilot Study." *Cancers*. Accessed December 17, 2023. <https://www.mdpi.com/2072-6694/13/16/4076/htm>.
- Frankfurt, Harry. 2005. *On Bullshit*. Princeton: Princeton University Press.
- Frost-Arnold, Karen. 2013. "Moral Trust and Scientific Collaboration." *Studies in History and Philosophy of Science* 44 (3): 301–310. <https://doi.org/10.1016/j.shpsa.2013.04.002>.
- Global Gurus top. 2024. Accessed February 6, 2024. <https://globalgurus.org/nlp-gurus-top-30/>.
- Grant, A. M. 2001. "Neuro-Associative Conditioning." *Australian Psychologist* 36 (3): 232–238. <https://doi.org/10.1080/00050060108259660>.
- Grant, A. M. 2019. "A Personal Perspective on Neuro-Linguistic Programming: Reflecting on the Tension Between Personal Experience and Evidence-Based Practice." *International Coaching Psychology Review* 14 (1): 45–56. <https://doi.org/10.53841/bpsicpr.2019.14.1.45>.
- Gray, Richard. 1991. "Tools for the Trade: Neuro-Linguistic Programming and the Art of Communication." *Federal Probation* 55 (11): 11–16.
- Gray, Richard. 2022. "Indirect Evidence Update." In *Neurolinguistic Programming in Clinical Settings Theory and Evidence-Based Practice*, edited by L. de Rijk, R. Gray, and F. Bourke, 185–224. London: Routledge.
- Gray, R., and F. Bourke. 2015. "Remediation of Intrusive Symptoms of PTSD in Fewer Than Five Sessions: A 30- Person Pre-Pilot Study of the RTM Protocol™." *Journal of Military Veteran and Family Health* 1 (2): 85–92. <https://doi.org/10.3138/jmvfh.2996>.
- Gray, R., D. Budden-Potts, and F. Bourke. 2017. "Reconsolidation of Traumatic Memories for PTSD: A Randomized Controlled Trial of 74 Male Veterans." *Psychotherapy Research* 14 (5): 621–639. <https://doi.org/10.1080/10503307.2017.1408973>.
- Gray, R. M., D. Budden-Potts, R. J. Schwall, and F. F. Bourke. 2020. "An Open- Label, Randomized Controlled Trial of the Reconsolidation of Traumatic Memories Protocol (RTM™) in Military Women." *Psychological Trauma: Theory, Research, Practice, & Policy* 13 (6): 641–651. <https://doi.org/10.1037/tra0000986>.
- Gray, R. M., A. Davison, and F. Bourke. 2021. "Reconsolidation of Traumatic Memories, the RTM Protocol™: Albuquerque Trainee Results." *PsyArxiv Preprints*. <https://doi.org/10.17605/OSF.IO/PFQG4>.
- Gray, Richard M., Richard F. Liotta, Lisa Wake, and Joe Cheal. 2013. "Research and the History of Methodological Flaws." In *The Clinical Effectiveness of Neuro-Linguistic Programming: A Critical Appraisal*, edited by Lisa Wake, Richard M. Gray, and Frank S. Bourke, 194–216. New York: Routledge.
- Greif, S. 2018. "How to Recognize Pseudo-Scientific Theories and Why They are Problematic in Coaching—For Example NLP." *Organisationsberatung, Supervision, Coaching* 25 (3): 371–387. <https://doi.org/10.1007/s11613-018-0568-y>.
- Greif, S. 2022. "Pseudoscience and Charlatany in Coaching." In *International Handbook of Evidence Based Coaching: Theory, Research and Practice*, edited by S. Greif, H. Möller, W. Scholl, J. Passmore, and F. Müller, 755–772. Cham: Springer Nature.
- Grimley, B. 2013. *Theory and Practice of NLP Coaching. A Psychological Approach*. London: Sage.
- Grimley, B. 2015a. "What is NLP?" PhD Diss., University of Central Nicaragua. Accessed December 18, 2023. <https://ow.ly/XQqcA>.
- Grimley, B. 2015b. "NLP: Misunderstood by Psychologists." *Sports and Exercise Psychology Review* 11 (2): 66–72. <https://doi.org/10.53841/bpssepr.2015.11.2.66>.

- Grimley, B. 2016. "What is NLP? The Development of a Grounded Theory of Neuro-Linguistic Programming, (NLP), Within an Action Research Journey. Implications for the Use of NLP in Coaching Psychology." *International Coaching Psychology Review* 11 (2): 166–178. <https://doi.org/10.53841/bpsicpr.2016.11.2.166>.
- Grimley, B. 2019. "The Need for Neuro-Linguistic Programming to Develop Greater Construct Validity." *International Coaching Psychology Review* 14 (1): 31–44. <https://doi.org/10.53841/bpsicpr.2019.14.1.31>.
- Grimley, B. 2020. *The 7Cs of Coaching: A Personal Journey Through the World of NLP and Coaching Psychology*. London: Routledge.
- Grimley, B. 2021a. "Sailing the 7C's of Courage - Away from the Practice of NLP and Towards a Novel and Cohesive Theory of Change. A Single Case History." *Journal of Experiential Psychotherapy* 24 (2(94)): 3–11.
- Grimley, B. 2021b. "Sailing the 7C's of Courage. A Framework for One-to-One Work Emerging from Neuro-Linguistic Programming (NLP) - A Single Case History." *Journal of Experiential Psychotherapy* 24 (1(93)): 3–17.
- Grinder, J., and R. Bandler. 1976. *The Structure of Magic 2. A Book About Communication and Change*. Palo Alto: Science and Behaviour Books.
- Grinder, J., J. DeLozier, and R. Bandler. 1977. *Patterns of the Hypnotic Techniques of Milton H. Erickson. MD. Vol. 2*. Capitola: Meta Publications.
- Grinder, J., and F. Pucelik. 2013. *The Origins of Neurolinguistic Programming*. Carmarthen: Crown House Publishing.
- Heap, M. 1989a. "Neurolinguistic Programming, A British Perspective." *Hypnos: Swedish Journal of Hypnosis in Psychotherapy and Psychosomatic Medicine* 15: 4–13. Accessed March 6, 2024. <https://www.aske-skeptics.org.uk/nlp3.pdf>.
- Heap, M. 1989b. "Neurolinguistic Programming: What is the Evidence?" In *Hypnosis: The Fourth European Congress at Oxford*, edited by D. Waxman, D. Pedersen, I. Wilkie, and P. Mellett, 118–124. London: Whurr Publishers.
- Heap, M. 2008. "The Validity of Some Early Claims of Neuro-Linguistic Programming." *Sceptical Intelligencer* 11: 6–13. Accessed March 6, 2024. <https://www.aske-skeptics.org.uk/nlp5.pdf>.
- Hofstadter, Douglas, and Emmanuel Sander. 2013. *Surfaces and Essences: Analogy as the Fuel and Fire of Thinking*. New York: Basic Books.
- Houningen-Huene, Paul. 2013. *Systematicity: The Nature of Science*. Oxford: Oxford University Press.
- IANLP (International Association for NLP). Accessed December 12, 2023. <https://www.ia-nlp.org/>.
- IN (International Association of NLP institutes). Accessed December 12, 2023. <https://www.nlp-institutes.net/>.
- INLPTA (International NLP Trainers Association). Accessed December 12, 2023. <https://www.inlpta.org/>.
- Janis, I. L. 1972. *Victims of Groupthink*. Boston: Houghton Mifflin.
- Kanning, P. 2019. "NLP zwischen Anspruch und Wirklichkeit." *Report Psychologie* 44 (1): 110–118.
- Keller, Evelyn Fox. 1983. *A Feeling for the Organism: The Life and Work of Barbara McClintock*. San Francisco: Freeman.
- Korzybski, A. 1994. *Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics*. 5th ed. New York: Institute of General Semantics.
- Kotera, Y., D. Sheffield, and W. Van Gordon. 2018. "The Applications of Neuro-Linguistic Programming in Organisational Settings: A Systematic Review of Psychological Outcomes." *Human Resource Development Quarterly* 30 (1): 101–116. <https://doi.org/10.1002/hrdq.21334>.
- Ladyman, J. 2013. "Toward a Demarcation of Science from Pseudoscience." In *Philosophy of Pseudoscience: Reconsidering the Demarcation Problem*, edited by M. Pigliucci and M. Boudry, 45–60. Chicago: The University of Chicago Press.
- Linder-Pelz, S. 2010. *NLP Coaching: An Evidence-Based Approach for Coaches, Leaders and Individuals*. London: Kogan Page.
- Linder-Pelz, S., and M. Hall. 2007. "The Theoretical Roots of NLP-based Coaching." *The Coaching Psychologist* 3 (1): 12–17.
- McIntyre, Lee. 2019. *The Scientific Attitude: Defending Science from Denial, Fraud, and Pseudoscience*. Cambridge, MA: MIT Press.
- Miller, G. A. 1956. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63 (2): 81–97. <https://doi.org/10.1037/h0043158>.
- Miller, G. A., E. Galanter, and K. H. Pribram. 1960. *Plans and the Structure of Behaviour*. New York: Holt, Rhinehart & Winston.
- Miller, W., and S. Rollnick. 2023. *Motivational Interviewing*. New York: Routledge.
- Mukerji, Nikil, and Edzard Ernst. 2022. "Why Homeopathy Is Pseudoscience." *Synthese* 200 (5): 394. <https://doi.org/10.1007/s11229-022-03882-w>.
- NLP Leadership Summit. Accessed December 12, 2023. <https://nplleadershipsummit.org/>.
- Nocross, J. C., G. P. Koocher, and A. Garofalo. 2006. "Discredited Psychological Treatments and Tests: A Delphi Poll." *Professional Psychology: Research and Practice* 37 (5): 515–522. <https://doi.org/10.1037/0735-7028.37.5.515>.
- O'Donohue, W., and K. E. Ferguson. 2006. "Evidence-Based Practice in Psychology and Behavior Analysis." *The Behaviour Analyst Today* 7 (3): 335–350. <https://doi.org/10.1037/h0100155>.
- Oreskes, Naomi. 2019. *Why Trust Science?*. Princeton: Princeton University Press.
- Osborne, Jason W., and Amy Overbay. 2004. "The Power of Outliers (and Why Researchers Should Always Check for Them)." *Practical Assessment, Research & Evaluation* 9 (6): 1–8.
- Parker, Phil. 2022a. "Is NLP Based on Science or is it Pseudoscience?" Accessed December 21, 2023. <https://www.philparker.org/is-nlp-based-on-science-or-is-it-pseudoscience/>.

- Parker, Phil. 2022b. "Psychoneuroimmunology – Research on Lightning Process." In *The Clinical Effectiveness of Neuro-Linguistic Programming: A Critical Appraisal*, edited by Lisa Wake, Richard M. Gray, and Frank S. Bourke, 138–161. New York: Routledge.
- Parker, P., J. Aston, and L. de Rijk. 2021. "A Systematic Review of the Evidence Base for the Lightning Process." *Explore* 17 (4): 372–379. <https://doi.org/10.1016/j.explore.2020.07.014>.
- Parker, P., J. Aston, and F. Finch. 2018. "Understanding the Lightning Process Approach to CFS/ME; A Review of the Disease Process and the Approach." *Journal of Experiential Psychotherapy* 21 (2): 82.
- Passmore, J., and T. Rowson. 2019. "Neuro-Linguistic Programming: A Review of NLP Research and the Application of NLP in Coaching." *International Coaching Psychology Review* 14 (1): 57–69. <https://doi.org/10.53841/bpsicpr.2019.14.1.57>.
- Pavlov, I. P. 1927. *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford: Oxford University Press.
- Pourmansour, J. 1997. *NLP und Zahnrangst? - Eine Studie zur Effektivität der NLP-Intervention*. Master's thesis, University of Münster, Münster.
- Roderique-Davies, G. 2009. "Neuro-Linguistic Programming: Cargo Cult Psychology?" *Journal of Applied Research in Higher Education* 1 (2): 57–63. <https://doi.org/10.1108/17581184200900014>.
- Shams, M. 2022. *Psychology in Coaching Practice. A Guide for Professionals*. Maidenhead: Open University Press.
- Sharpley, C. F. 1984. "Predicate Matching in NLP: A Review of Research on the Preferred Representational System." *Journal of Counseling Psychology* 31 (2): 238–248. <https://doi.org/10.1037/0022-0167.31.2.238>.
- Sharpley, C. F. 1987. "Research Findings on Neurolinguistic Programming: Nonsupportive Data or an Untestable Theory?" *Journal of Counseling Psychology* 14 (1): 103–107. <https://doi.org/10.1037/0022-0167.34.1.103>.
- Sturt, J. 2012. "Little Evidence for NLP in Healthcare." Accessed December 12, 2023. <https://www.kcl.ac.uk/archive/news/nmpc/2012/little-evidence-for-nlp-in-healthcare>.
- Sturt, J. 2022. "PTSD Experimental Treatment Trial (PETT): Comparing Two Talking Therapies for the Treatment of Post-Traumatic Stress Disorder in UK Military Veterans." *ISRCTN Registry*. Accessed December 19, 2023. <https://www.isrctn.com/ISRCTN10314773?q=STURT&filters=&sort=&offset=3&totalResults=5&page=1&pageSize=10>.
- Sturt, J., S. Ali, W. Robertson, D. Metcalfe, A. Grove, C. Bourne, and C. Bridle. 2012. "Neurolinguistic Programming: A Systematic Review of the Effects on Health Outcomes." *British Journal of General Practice* 62 (604): e757–e764. <https://doi.org/10.3399/bjgp12X658287>.
- Sturt, J., R. Rogers, C. Armour, N. Greenberg, D. Cameron, L. de Rijk, F. Fiorentino, et al. 2022. *Reconsolidation of Traumatic Memories Protocol Compared to Trauma Focussed Cognitive Behaviour Therapy for Posttraumatic Stress Disorder in UK Military Veterans: A Randomised Controlled Feasibility Trial*. Accessed December 19, 2023. <https://www.researchsquare.com/article/rs-2132714/v1>.
- Tetens, Holm. 2013. *Wissenschaftstheorie: Eine Einführung*. Munich: C. H. Beck.
- Tosey, P., and J. Mathison. 2009. *Neuro-Linguistic Programming: A Critical Appreciation for Managers and Developers*. Basingstoke: Palgrave Macmillan.
- Tylee, D. S., R. Gray, S. J. Glatt, and F. Bourke. 2017. "Evaluation of the Reconsolidation of Traumatic Memories Protocol for the Treatment of PTSD: A Randomized, Wait-List Controlled Trial." *Journal of Military Veteran and Family Health* 3 (1): 21–33. <https://doi.org/10.3138/jmvfh.4120>.
- van Fraassen, Bas. 2008. *The Empirical Stance*. New Haven: Yale University Press.
- Wake, L., R. M. Gray, and F. S. Bourke, eds. 2013. *The Clinical Effectiveness of Neurolinguistic Programming. A Critical Appraisal*. London: Routledge.
- Wiseman, R., C. Watt, L. ten Brinke, S. Porter, S. Couper, and C. Rankin. 2012. "The Eyes Don't Have It: Lie Detection and Neuro-Linguistic Programming." *Public Library of Science One* 7 (7): e40259. <https://doi.org/10.1371/journal.pone.0040259>.
- Witkowski, T. 2010. "Thirty-Five Years of Research on Neuro-Linguistic Programming. NLP Research Data Base. State of the Art or Pseudoscientific Decoration?" *Polish Psychological Bulletin* 41 (2): 58–66. <https://doi.org/10.2478/v10059-010-0008-0>.
- Witt, K. 2008. "Neuro-Linguistic Psychotherapy, (NLPt), Treatment Can Modulate the Reaction in Pollen Allergic Humans and Their State of Health." *International Journal of Psychotherapy* 12 (1): 50–60.
- Zaharia, C., M. Reiner, and P. Schütz. 2015. "Evidence-Based Neurolinguistic Psychotherapy: A Meta-Analysis." *Psychiatria Danubina* 27 (4): 355–363.