Arguing on Unlevelled Playing Fields: Local Management of Reasonableness in a Special Place for Managing Disagreement About Health

Jackson, Sally; Boogaart, Ronny; Garssen, Bart; Jansen, Henrike; Van Leeuwen, Maarten; Pilgram, Roosmaryn; Reuneker, Alex

Boogaart, R., Garssen, B. Jansen, H., Leeuwen, M. van, Pilgram, R. & Reuneker, A. (2024).
Proceedings of the Tenth Conference of the International Society for the Study of Argumentation.
Sic Sat: Amsterdam.

Arguing on Unlevelled Playing Fields: Local Management of Reasonableness in a Special Place for Managing Disagreement About Health

SALLY JACKSON

Department of Communication
University of Illinois at Urbana-Champaign
USA
sallyj@illinois.edu

ABSTRACT: Within the argumentative polylogue framework of Lewiński and Aakhus (2023), the *places* where arguments develop are recognized as potential targets for design. In the UK, an important place for arguing about health is a process for creating clinical care guidelines. A review of the design of this place exposes both its vulnerabilities (e.g., unexamined presumptions) and its resources for overcoming these vulnerabilities (openness to both systemic redesign and to local management by participants).

KEYWORDS: argumentative polylogue, health controversy, clinical guidelines, myalgic encephalomyelitis, chronic fatigue syndrome

1. INTRODUCTION

This paper has to do with the designability of argumentation, including not only its interactional design but also its repertoire of inferential methods. For the former, I draw heavily on the work of Mark Aakhus (especially 1999, 2001, 2002, 2003, 2007, 2013, and 2017) and on the idea that contemporary argumentation practice is already heavily "built up" through technical improvements that call for a comprehensive rethinking of its nature. For the latter, I incorporate my own recent work on innovation in inference methods (Jackson, 2015a, 2015b, 2019; Jackson & Schneider, 2018; Schneider & Jackson, 2018).

A focus on designability almost inevitably challenges legacy thinking about argumentation. Newly designed objects do not generally fit neatly into conceptual systems devised from pre-existing objects that they augment or replace. The "built-up" nature of contemporary communication practice (Aakhus, 2007) is part of the rationale for a wholesale theoretical shift advocated by Lewiński and Aakhus (2023). The polylogue framework they advance reconceptualizes argumentative discourse as complex networks of *positions*, *players* who advance positions, and *places* where players exchange arguments for and against positions. All of these elements have adapted to change in media ecology.

'Place,' though, is my main concern in this paper. Lewiński and Aakhus (2023, p. 81) describe place as "the particular sociolinguistic work, including the physical, technological, and institutional accoutrement, that projects and enables (or limits) various kinds of participation status and the available moves and countermoves in interaction." Even the simplest dyadic exchange involves sociolinguistic work. Argument occurring spontaneously in interpersonal interaction is structured by participant-administered turn-taking and repair-like sequential structures (Jackson & Jacobs, 1980; Jacobs & Jackson,

1982, 1983). From these taken-for-granted elements arguers can spontaneously restructure their own argumentative practices, experimenting with alternative ways of allocating turns at talk. However, the polylogue framework exposes an important fact about all contemporary discourse: No place stands alone as a site for arguing. Places exist within what Lewiński and Aakhus call an "ecology of places." Increasingly, this ecology of places includes some that have been deliberately designed around furtherance of argumentative purposes. And whether spontaneously or deliberately designed, places are open to redesign.

Though I am not at present prepared to explain how the evolving ecology of places in entwined with innovation in reasoning, some important hints have been given by Toulmin (1958) in his controversial theory of field-dependent reasoning. Further hints will emerge from the work reported here.

2. PLACES FOR ARGUING ABOUT HEALTH AND MEDICINE

Health controversies are replete with unmanaged and mismanaged disagreements, but also with interventions taking the form of places designed for making argumentation more productive. Elsewhere (Jackson, 2023) I have discussed various places that have been constructed on purpose to try to resolve disagreements over an illness so controversial that to refer to it by any name is to align oneself with a contentious position. The illness is now mostly known by initials: ME/CFS, from myalgic encephalomyelitis and chronic fatigue syndrome. Places for attempting progress toward resolution have included (among others): a working group on CFS/ME formed by the UK's Chief Medical Officer in 1998 (CFS/ME Working Group, 2002), a panel formed by the US Institute of Medicine in 2013, whose report (IOM, 2015) proposed an entirely new name, "systemic exercise intolerance disease"; and a Dutch advisory group on ME/CFS formed in 2013 (Health Council of the Netherlands, 2018). Each of these could be the subject of design inquiry. However, I have chosen for this analysis a place designed not for any single controversy but for repeated use in managing disagreements of importance to public health. This place is a guideline development process designed and administered by the UK's National Center for Health and Care Excellence (https://nice.org.uk).

NICE was chartered in 1999 with a mission "to improve health and wellbeing by putting science and evidence at the heart of health and care decision making" (https://www.nice.org.uk/about/who-we-are/our-charter). It is an organization whose existence presupposes a certain ecology of places, and a certain repertoire of reasoning resources, that did not exist a century earlier. Marks (1997) traces the rise of medical science through the twentieth century, showing how the invention of randomized clinical trials (RCTs) was intertwined with aspirations for "therapeutic reform."

To situate the analysis within the recent history of health policy debate, NICE was formed just as evidence-based medicine was taking hold, a time when there was a strong sense worldwide that medical practice was lagging behind what medical science could support. By the turn of the century, the ecology of places already included clinical care settings, heterogeneous science places, various kinds of agency processes, and more. In this context, NICE began designing a process for translating scientific findings into clinical guidelines to provide a routine path from science to practice.

Figure 1 is an effort to show what was designed and how it changed the ecology of places. Pre-dating the creation of NICE, public health authority would have been part of a network of player nodes, place nodes, and the connections between player nodes and place nodes, all of which are indicated in Figure 1 with solid borders and lines. There was already plenty of science present in the network (coming from "science places" that are not shown), but there were only indirect and uncertain links between the scientific literature and clinical practice. Elements that were added to this network are shown with dotted borders and lines, the most significant of which are those that create an established path from the scientific literature to clinical care through a new place node connected directly to both.

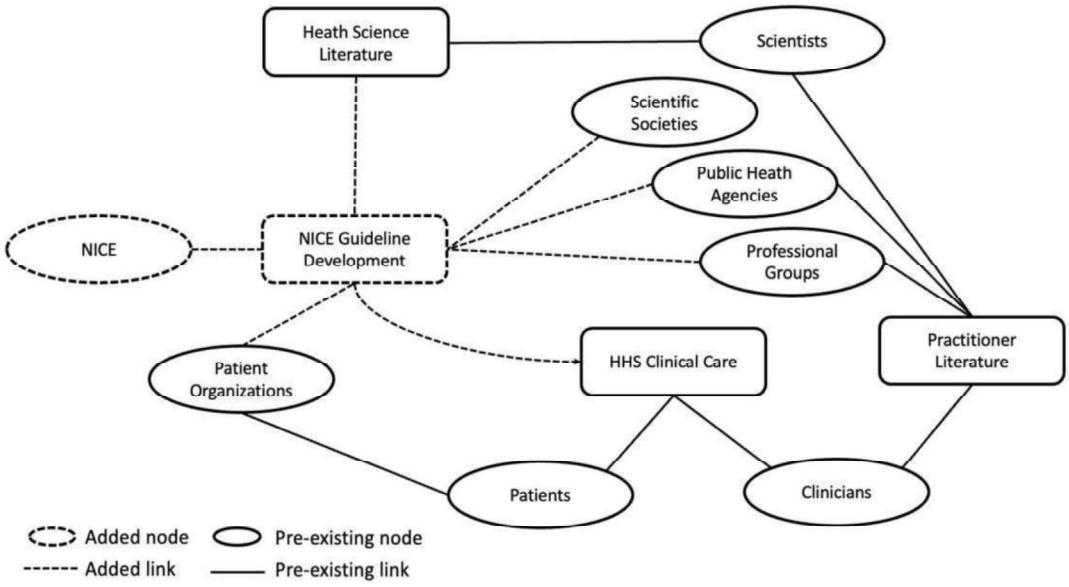

Figure 1. Players (ovals) and places (rounded rectangles) in a generic health polylogue before (solid lines) and after (dotted lines) NICE intervened with a new place.

A basic outline of the new process is shown in Figure 2, with the many tasks involved sorted by who has primary responsibility for each set of tasks. An important feature of the design of this process is the delegation of guideline writing to those most affected by the guideline. NICE assures that the writers of the guideline include representatives from the most important stakeholder groups (those who receive treatment and those responsible for offering treatment). The committee includes experts (professionals engaged in practice) and lay members (generally patient representatives). Organizations representing scientific expertise may register as stakeholders and offer written comments to the committee, but individual scientists are rarely included as voting members of the guideline committee. For patient stakeholders other than the very small number of lay members of any committee, the NICE process created a need for an entity to register on their behalf if no such patient organization had already formed around a "disease constituency" (Epstein, 2016).

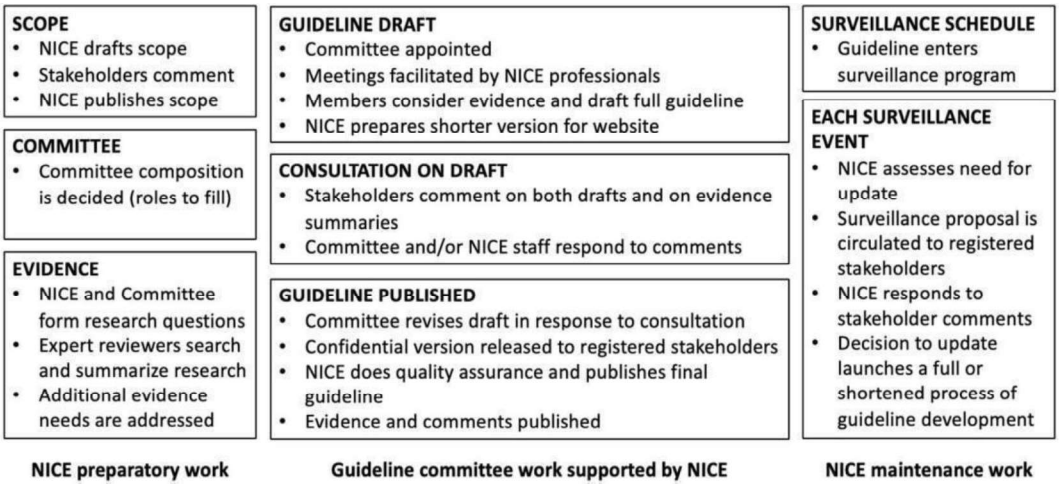

Figure 2. The NICE guideline development process in broad outline.

Although this outline has not changed since NICE's formation, details of how tasks are performed have undergone considerable revision. This can be seen by comparing successive versions of the manual NICE uses to structure the process. Comparing the initial manual (NICE, 2006a) and the most recent revision (NICE, 2023), the most striking difference is the incorporation of a vast number of innovations in reasoning about health that have occurred as a result of dilemmas faced in attempting to turn raw science into something useful at the clinical level. These include literature search methods, research synthesis methods, and evidence grading methods. In 2006, guidance on how to work with research evidence was rather minimal; today it is intricate and technical, with an increasing list of unusual evidence judgment problems to consider and suggestions on how to solve them. For example, as early as 2007 NICE had noticed that its preferred methods for representing evidence quality were too simplistic and had begun experimenting with new methods that were soon incorporated into all guideline development (Thornton et al., 2013; see esp. p. 124). As these new methods have been incorporated over time, the judgments made by participants in the guideline writing process should be expected to have become more nuanced.

Two things have not changed over successive revisions of the manual. First is that the guideline committee has authority to decide what to recommend. NICE personnel play supporting roles (including various kinds of quality assurance) before, during, and after the committee's term of service. The second is a nonprescriptive stance toward disagreement management methods for the guideline committee itself. All appointees must agree to a code of conduct, but the committee is then left to manage disagreements however it deems best.

3. DESIGN LOGIC IN THE NICE GUIDELINE PROCESS

Aakhus (2002) suggested that designs for arguing be evaluated against normative models of argumentation such as the critical discussion model of pragma-dialectics, and Aakhus

and Bzdak (2015) elaborated this suggestion for processes that involve stakeholder engagement. According to Aakhus and Bzdak, a design's logic can be described in terms of its (1) exigency, (2) purpose, (3) orchestration methods, and (4) systemic rationality. For brevity, these will be explained as applied to the NICE process.

Exigency has to do with what prompts the creation of the design. For NICE, the exigency leading to the guideline development process was the gap between expectations for benefits from investment in medical science and actual improvements visible in medical practice. It was generally assumed at the time that the main barrier to evidence-based medicine was that no practitioner could be expected to master the scientific literature for every condition encountered in practice. The exigency in this case and the general form of NICE's response can be seen quite clearly in Figure 1: The lack of a path for content to move from research places to clinical care places is remedied by creation of research-based guidelines that can be consulted as a routine part of clinical practice.

Purpose is different from exigency and also distinct from the goals of any of the identifiable stakeholders. The purpose of the NICE process has always been to produce guidelines for practice, adapting both to the exigency just noted and to general assumptions about democratic participation. Under similar exigencies, it has been common for governments to assemble scientists to come to consensus on how their findings should inform practice, but NICE intended from the start that people with the most stake in the guideline should be responsible for ultimate decisions about whether a body of scientific work made a clear case for any particular approach to a given disease. This did not mean, however, that stakeholders could simply bargain among themselves to reach guidelines acceptable to as many as possible; the result had to be defensible as a considered response not just to stakeholder preferences but to the research evidence (including evidence of cost-effectiveness). A clear design hypothesis here is that ordinary stakeholders can and will take scientific evidence into account if well enough supported. The support needed by stakeholders had to be learned from experience (as is evident from the evolution of the manual).

Orchestration techniques are procedures that structure contributions to the discourse. In NICE's case, there are several distinguishable activities to orchestrate, of which the most important are assembly of a body of scientific evidence, identification of legitimate stakeholders, formation and support of a writing committee, consultation between the writing committee and the rest of the stakeholders, publication of the guideline, and ongoing "surveillance" for signs that the guideline needs updating. Each of these can be orchestrated in ways different from what NICE actually chose to do. For example, the writing committee's deliberations take place under strong confidentiality agreements, while the consultations are entirely public. The first of these decisions reflects a design hypothesis that a closed group engaged in sustained collaborative work will resolve more disagreements than an open-ended set of interested parties who are not obliged to respond to one another. The latter decision (to make consultations public) reflects a design hypothesis that all players' positions should be open to inspection by all others. Assuming NICE does not undercut its own design, this means that no player can influence the content of a guideline by lobbying NICE directly; all must address the guideline committee.

Systemic rationality refers to whatever it is about the process that can be pointed to as a reason to trust its outputs. Two specific design features of importance to systemic

rationality of the guideline process are the highly regimented methods for reviewing and evaluating scientific literature and the strong commitment to inclusiveness. A design hypothesis evident in the former is that the fact base should be assembled by those with expertise but with no stake in the decision. For the latter, the bearing of the fact base on any practical decision requires legitimation by those with the most at stake. The claim made for the credibility of the system's output is that it brings a shared body of fact to a decision made by people directly affected by what happens.

One overarching design hypothesis in the NICE guideline development process is the idea of dividing up a decision-making task and delegating different parts to different performers. As Figure 2 makes clear, the process of writing a guideline is quite complex, involving at least four groups of task performers (not counting scientists who produce primary research reports): NICE professionals with many different technical skills, outside experts to whom tasks can be delegated, an independent committee of representative stakeholders, and a much large set of other stakeholders who contribute written argumentative content to be considered by NICE and its committee. NICE's most important actions are two delegations: one to internal or external experts trusted with impartially reviewing the scientific research literature, and the other to a NICE-appointed committee entrusted with bringing stakeholder views to bear on formation of the guideline. This overarching hypothesis has been tested in many prior designs, not always with good results (see, e.g., Aakhus, 1999). There is no guarantee that it will have the benefits expected—but the systemic rationality of this particular design depends on both delegations performing as expected. NICE does not "delegate" stakeholders to comment but invites them to do so and promises to see that comments are addressed by the performers of delegated tasks.

Another overarching design hypothesis, most visible in the surveillance program, is that decisions about health care are never really final; by design, NICE holds its institutional positions accountable to new evidence, and by extension, to new thinking *about* evidence. Obviously, research done subsequent to the publication of a guideline may call a recommended treatment into question, or it may identify some new treatment that is preferable to the best treatment known at the time the guideline was published. Much less obviously, innovation in reasoning about health may lead to new perspectives on old research. Starting around the turn of the twenty-first century, interest in more rigorous methods for evaluating the strength of research evidence escalated in connection with other innovations in research synthesis methods. Several frameworks for assessing evidence appeared independently, including a system known as GRADE designed specifically for evaluating the relevance and strength of evidence relative to practice guidelines (Guyatt, Sackett, et al., 1995; Guyatt, Oxman, et al., 2008). When NICE adopted the GRADE system, it formalized a presumption established in the original guideline manual (that RCTs should form the core of evidence for treatment effects if possible) while adding explicit mention of conditions under which any actual instance of a presumptively strong form of evidence might have to be "downgraded."

While quite different choices might have been made in the design of this process or in its occasional improvements, nothing here looks manifestly unreasonable. Departures from idealizations of argumentation are evident (especially those involving the establishment of presumptions) but they all have an intelligible purpose. But whether different choices would produce better decisions, or better-defended decisions, is actually

an empirical question, not a theoretical or philosophical one. The main reason for this is that, as with all designed things, we cannot know in advance how the design's action possibilities will be absorbed into actual practice. So unless a design *is* manifestly unreasonable (for example, actively suppressing dissent), its evaluation will inevitably require observation of how people do their own sociolinguistic work within it. Local management by participants themselves often exposes strengths and weaknesses of the design that are impossible to anticipate in the abstract.

4. THE NICE DESIGN UNDER LOCAL MANAGEMENT

NICE has used its guideline process hundreds of times. In each use it must be enacted by a different group of participants who engage in local management of the interactions occurring under NICE's orchestration. No one case can show how often the design performs as expected. However, a well-chosen case *can* show ways in which it might fail and ways it might avert failure. Looking closely at how the NICE guideline development process has worked for the extremely vexing case of ME/CFS exposes certain vulnerabilities, but as will be shown, it also puts on display the impressive capacity of this design for locally managed repair and for systemic improvement over time.

NICE's involvement with the ME/CFS controversy began in 2004 and is ongoing. Long before NICE's intervention, the controversy had taken shape around two core concerns: (1) the nature of the disease and (2) the treatments available to patients. Patients mostly believed that they were suffering from bodily illness requiring medical treatment, and clinicians—especially in the UK—were much more likely to suspect psychosomatic illness. The guideline development process produced an institutional position on treatments but did not take a position on the exact nature of the disease. Omitting scientists, Figure 3 shows roughly what positions were circulating and who their proponents were.

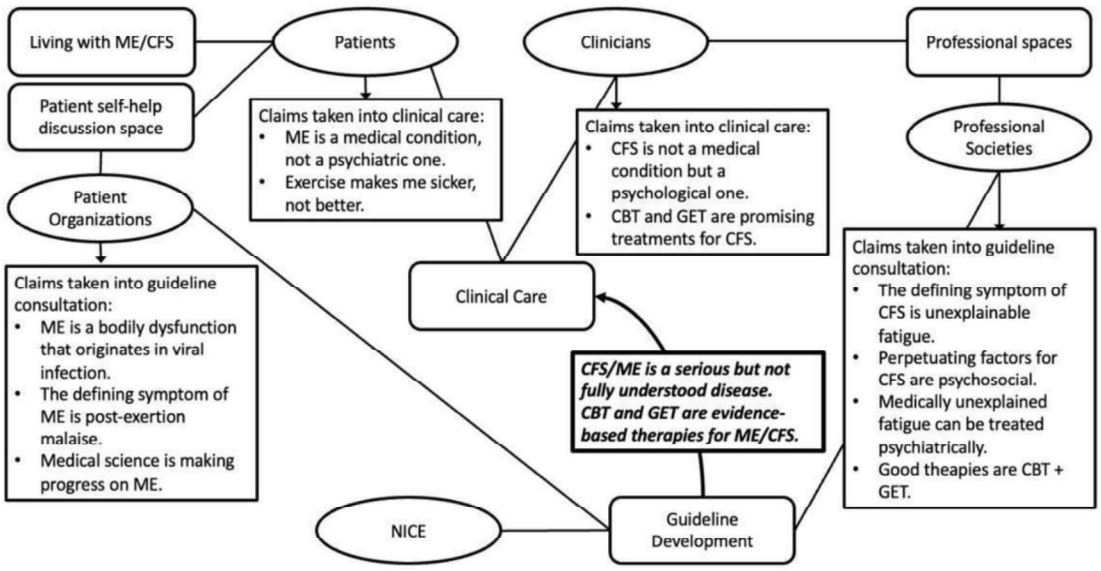

Figure 3. Players, positions, and places involved in or affected by CG53 (2007).

Argumentative texts generated in the development of a guideline are important data for assessing the design. These texts include preliminary exploration of a topic by NICE itself, consultation over its initial ideas about how to charge a guideline committee, preparation of one or more expert reviews of literature, notes on committee meetings, drafts of documents, and extensive written contributions from stakeholders. Questions that can be answered by examining the arguments in these texts have to do with whether, under this design, stakeholders elaborate their reasoning in normatively positive or negative ways. Do they attempt to resolve issues, or simply press their own preferences? Do they pay attention to the reasoning of their opponents, or simply contradict opposing views? Do they engage in the form of argument extension that is the hallmark of good debating, or simply escalate pressure? Do they embrace the systemic rationality of the process, or seek to undermine it in single-minded pursuit of their individual goals? It is important to look not only at what arguments are made but also at how these arguments fare within the various steps of the process. How are constructive and obstructive moves taken up by those able to respond?

Figure 4 shows diagrammed arguments from one thread involving arguments presented during consultation over NICE's first guideline on ME/CFS (the draft guideline, NICE, 2006b; consultation comments, NICE, 2006c; and the final version of the guideline authored by the committee, NICE, 2007a). The issue in this thread is whether the guideline should recommend a form of treatment known as Graded Exercise Therapy (GET). This treatment had been tested in randomized clinical trials (RCTs) that appeared to show moderate benefits for patients being treated for chronic fatigue, but patient organizations had accumulated many reports that GET worsened patients' symptoms. Similar arguments centered on Cognitive Behavior Therapy (CBT), though here patients reported only that it was ineffective, not that it was harmful. The arguments on both sides of the issue center on what to count as evidence, and it is clear that the NICE process allows argument to go quite deeply into evidence assessment. But the thread concludes in a disappointing way, with a choice of which evidence to believe based on placement of the evidence type within a prescribed hierarchy in which RCTs and systematic reviews of RCTs are placed at the top.

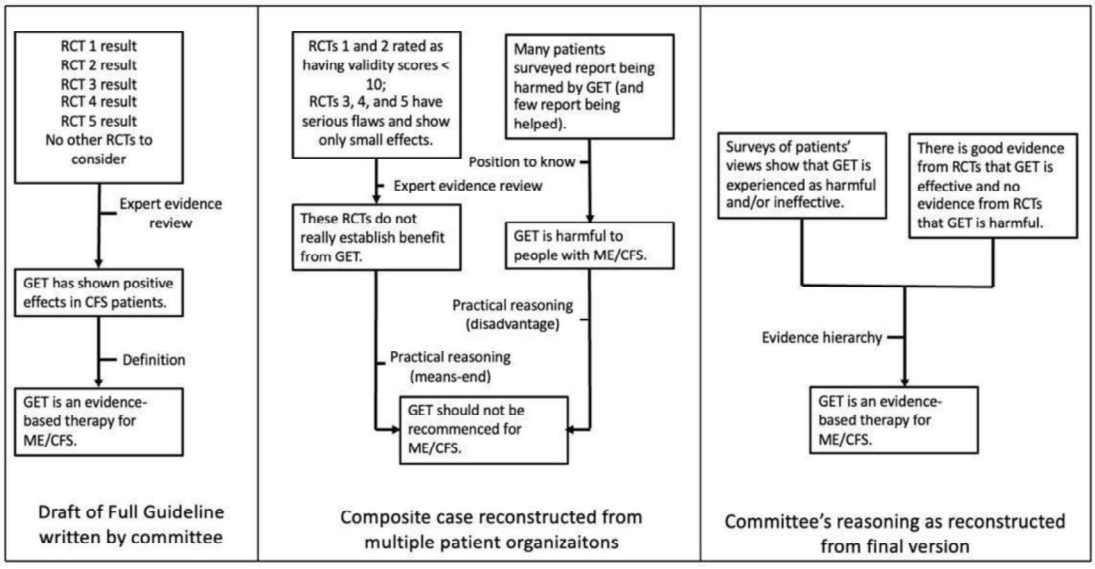

Figure 4. One thread in debate over content of CG53.

There are several design hypotheses deeply embedded here that do not come to the surface in the consultation over this guideline. These have to do with the topic-independent assumption that RCTs are ideal for evaluating effects of medical treatments and the subsidiary idea that a formal evidence search should begin by searching for RCTs, resorting to other kinds of evidence only when RCTs are lacking. The rationale for this is clear, but it has an unintended consequence: It creates an unlevel playing field for competing ideas about treatments, advantaging any treatment that someone would invest in evaluating within an RCT and disadvantaging ideas that have only civilian patients as champions. In the 1990s, as the illness was becoming better known, many ideas about possible treatments were circulating, including some that were discovered by patients and widely shared among themselves. While RCTs were done on many ideas about drugs that could be prescribed or therapies that could be professionally delivered, patients' ideas were largely ignored.

A presumption in favor of treatments backed by scientific evidence is not unreasonable, so long as it is possible to overturn the presumption through reasonable exception. In 2006, the strength of the presumption was at a high point. RCTs and evidence pyramids were still recent inventions for which there was great optimism and great enthusiasm. It was certainly realized by then that some RCTs are poorly designed (as is evident in inventories of flaws threatening validity, such as Campbell & Stanley, 1966). It was not so well understood that some RCTs produce weak evidence through no fault of the experimenter, and that *many* produce evidence that supports no one position (e.g., because effects happen to be very small or too highly variable to be predictable for any patient). Even less well understood in the general public were the processes leading some treatments to be extensively researched and others to be ignored. The guideline development committee embraced top-of-the-pyramid thinking rather uncritically, but they can be faulted for this only if we are willing to fault them for not discovering *right away, on their own*, that arguments relying on RCTs are defeasible for any or all of these reasons.

Clinical Guideline 53 (CG53, NICE, 2007b) identified GET and CBT *and no other treatments* as evidence-based for ME/CFS patients. The practical consequence of this would be that for the next 15 years patients seeking treatment in the NHS would be ushered into these treatments, a deeply distressing outcome for the patients and their advocacy organizations. Had any other treatments (such as pharmaceuticals) been shown through RCTs to be helpful for ME/CFS patients, the small benefits observed for GET and CBT would have been less likely to dominate care. Patient organizations were appalled by this situation and redoubled both their efforts to assemble evidence against these treatments and their efforts to identify promising alternatives as research priorities meriting support. What happened after is very pertinent to the evaluation of the NICE design.

NICE (the organization) is not a place but a player that takes institutional positions of its own, mostly by accepting reports from its committees. In 2007, NICE accepted the committee report, taking on as its own institutional position that GET and CBT were "evidence-based" for treatment of moderate CFS. NICE resisted many efforts by stakeholders to get it to withdraw its endorsement of CG53, but as noted earlier, the process has a built-in repair mechanism consisting of a regular schedule of surveillance reviews. Originally, these reviews were meant to identify new research that might change any component of a guideline. CG53 came due for routine surveillance in 2010, and NICE duly consulted with the stakeholder community before postponing a decision until 2011, pending results from an ongoing RCT known as the PACE trial (which, significantly, also

included a treatment meant to represent a technique favored by patients known as "pacing"). When PACE findings became available and appeared to confirm the benefits of GET and CBT (and to show no benefit from pacing), NICE stood by the CG53 committee's conclusions. Skirmishes around surveillance reviews continued (along with a considerable amount of other activism outside the NICE process). NICE remained implacable—until 2017, when patient advocates at last succeeded in putting together a challenge capable of overturning the presumption *in this case* without undermining the value of RCTs in general. This challenge depended on showing that CG53's original evidence base was weaker than formerly believed and that new research purporting to *confirm* the value of GET and CBT instead raised new concerns about the entire body of evidence.

In a document known as a surveillance proposal, NICE (2017a) argued that there was no reason to update CG53. Reports had continued appearing from the PACE trial, purporting to add evidence in favor of CG53's treatment recommendations. But the validity of the PACE trial was very widely questioned in the research literature as well as in patient communities, especially after the experimenters were forced by court action to release anonymized data that others could re-analyze. Figure 5 shows how NICE attempted to handle the critiques of PACE that had appeared by 2017. NICE reasoned (panel a) that if PACE were discounted entirely, there would be no justification for reconsidering the conclusions drawn by the CG53 committee. This reasoning drew many direct rebuttals in stakeholder consultation (NICE, 2017b), of which two are shown. VIRAS (panel b) pointed out that taking PACE at face value, the evidence it produced was evidence that the objective benefits of GET and CBT were too small to matter at all to patients. BRAME (panel c) pointed out that it had never been reasonable to consider updates only on the appearance of new evidence, a clear step forward in recognizing that problems with treatments in actual use may not appear during trials and may never be documented in research literature.

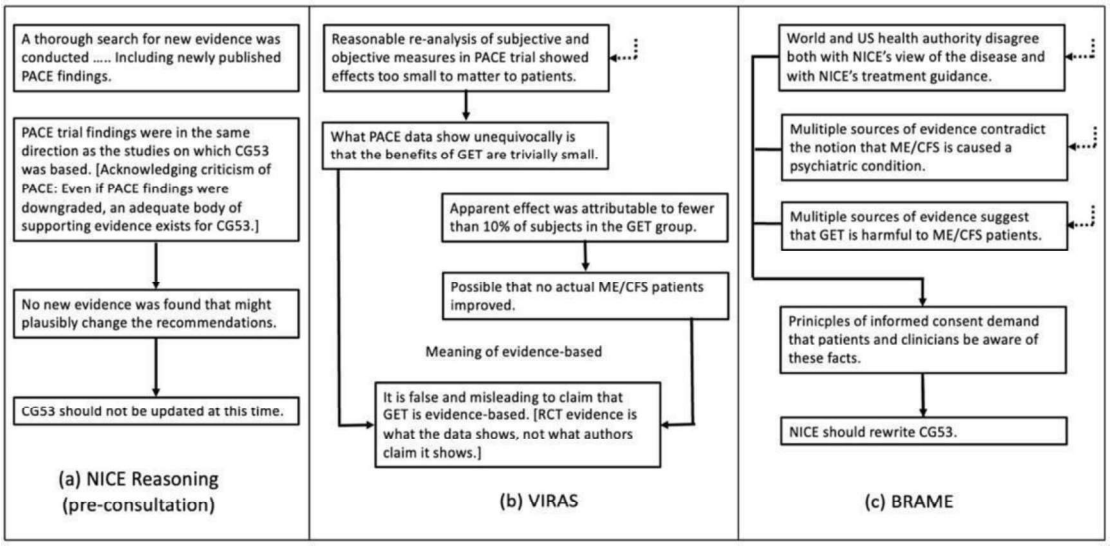

Figure 5. Two rebuttals of a position taken by NICE in surveillance review of CG53.

In its final surveillance report, NICE (2017c) reversed position, based on new evidence brought forward by stakeholders during consultation. Thereafter, a new guideline

development committee was appointed, and a new literature review was conducted from scratch. The original evidence base for everything included in CG53 was re-reviewed using updated methods that were part of the then-current manual. This produced a new wave of protest from a different set of stakeholders (a subset of Royal Colleges representing interested medical specialties). But in the end, NICE Guideline 206 (NG206; NICE, 2021) cautioned against referring ME/CFS patients to GET and recommended CBT only in a very limited role as compared with its centrality to treatment under CG53.

NICE's involvement is still not over. Subsequent to publication of NG206, a powerful and prestigious group of stakeholders have urged NICE to set NG206 aside, publishing harsh critiques in medical journals (Flottorp, et al., 2022; White, et al., 2023). So far, NICE has stood by its committee and defended its process (Hall, 2023). NICE's confidence in its guideline committees is anchored in the consistency of its process from topic to topic and in the belief that a heterogeneous committee arguing among themselves will have considered the arguments like those now being made in the press. How the present dispute will play out is not at present predictable; it depends largely on how the critiques of NG206 are answered.

5. CONCLUSION

Strong presumptions like those favoring RCT evidence over all other evidence create an unlevel playing field in NICE's guideline process, especially under conditions in which some stakeholders have more access to the performance of RCTs than others. But whether an unlevel playing field resulting from an established presumption is a threat to the quality of argumentation depends on what possibility there is for the presumption to be overturned. And overturning presumptions is necessarily a matter for local management; it is a kind of repair mechanism to be used infrequently, in circumstances that cannot be predicted.

Recall that the presumption is here orchestrated through the performance of the literature review, which becomes a starting point for the work of the guideline committee. This bit of orchestration is easily tweaked if experience shows that it creates more problems than it solves. But if experience shows that problems can be handled through local management, it is better to safeguard local management than to try to achieve infallibility. This is because delegating questions to experts is generally a good design hypothesis, but attempting to put them beyond criticism is a terrible one.

In the case considered here, stakeholders on the patients' side were eventually able to overturn the presumption enjoyed by this specific body of RCT evidence, after which the defense of GET and CBT began to crumble. Perhaps the most encouraging thing to notice in this case is that patients' persistent critiques of CG53 were not attacks on NICE's design, but a limited effort at repair within one instantiation of that design. Patients never attacked reliance on RCTs in the abstract. Patient organizations never suggested that NICE should rethink its general preference for RCTs or restructure its process to de-emphasize the reliance on science. Though ME/CFS patients have long been stereotyped as militant in their attacks on science (Blease & Geraghty, 2018), what actually happened here was that patient organizations quickly saw that procedures that are reasonable in the abstract had broken down *in this particular instance.* And, importantly, the process left room for local management to repair the breakdown.

The NICE process is not foolproof, but as a place designed for managing disagreement about health, it withstands scrutiny quite well. Much credit goes to NICE's institutional commitment to making full use of innovations in reasoning about health, but so too is credit due to how the process leaves many avenues open for local management of unforeseeable dilemmas in reasoning from medical science to health care practice.

ACKNOWLEDGEMENTS: Thanks are due to Scott Jacobs and Mark Aakhus, both for comments on this paper and for many years of collegial support in thinking about the role of design in argumentation.

REFERENCES

Aakhus, M. (1999). Science court: A case study in designing discourse to manage policy controversy. *Knowledge, Technology & Policy, 12*(2), 20–37. https://doi.org/10.1007/s12130-999- 1020-6

Aakhus, M. (2001). Technocratic and design stances toward communication expertise: how GDSS facilitators understand their work, *Journal of Applied Communication Research, 29*(4), 341-371. https://doi.org/10.1080/00909880128113

Aakhus, M. (2002). Modeling reconstruction in groupware technology. In F. H. van Eemeren (Ed.), *Advances in pragma-dialectics* (121-136). Newport News: Vale Press.

Aakhus, M. (2003). Neither naïve nor critical reconstruction: Dispute mediators, impasse, and the design of argumentation. *Argumentation*, *17*, 265-290. https://doi.org/10.1023/A:1025112227381

Aakhus, M. (2007). Communication as design. *Communication Monographs*, *74*(1), 112-117. https://doi.org/10.1080/03637750701196383

Aakhus, M. (2013). Deliberation digitized: Designing disagreement space through communication-information services. *Journal of Argumentation in Context*, *2*(1), 101-126. https://doi.org/10.1075/jaic.2.1.05aak

Aakhus, M., & Bzdak, M. (2015). Stakeholder engagement as communication design practice. *Journal of Public Affairs*, *15*(2), 188-200. https://doi.org/10.1002/pa.1569

Aakhus, M. (2017). The communicative work of organizations in shaping argumentative realities. *Philosophy & Technology*, *30*(2), 191-208. https://doi.org/10.1007/s13347-016-0224-4

Blease, C., & Geraghty, K. (2018). Are ME/CFS patient organizations "militant"? Patient protest in a medical controversy. *Bioethical Inquiry, 15*(3), 393–401. https://doi.org/10.1007/s11673-018-9866-5

Campbell, D. T., & Stanley, J. C. (1966). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.

CFS/ME Working Group. (2002). *A report of the CFS/ME working group: A report to the Chief Medical Officer of an independent working group.* UK Department of Health.

Epstein, S. (2016). The politics of health mobilization in the United States: The promise and pitfalls of "disease constituencies". *Social Science & Medicine*, *165*, 246-254. https://doi.org/10.1016/j.socscimed.2016.01.048

Flottorp, S. A., Brurberg, K. G., Fink, P., Knoop, H., & Wyller, V. B. (2022). New NICE guideline on chronic fatigue syndrome: More ideology than science? *The Lancet*, *399*(10325), 611-613. https://doi.org/10.1016/S0140-6736(22)00183-0

Guyatt, G. H., Sackett, D. L., Sinclair, J. C., Hayward, R., Cook, D. J., Cook, R. J., ... & Wilson, M. (1995). Users' guides to the medical literature: IX. A method for grading health care recommendations. *JAMA*, *274*(22), 1800-1804. https://doi.org/10.1001/jama.1995.03530220066035

Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ, 336,* 924-926. https://doi.org/10.1136/bmj.39489.470347.AD

Hall, R. (2023, July 11). ME/CFS guidance that discourages exercise is flawed, say researchers. *The Guardian.* https://www.theguardian.com/society/2023/jul/11/chronic-fatigue-guidance-discouraging- exercise-is-flawed-say-researchers

Health Council of the Netherlands. (2018). *ME/CFS*. The Hague: Health Council of the Netherlands, publication no. 2018/07

Institute of Medicine. (2015). *Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness*. Washington, DC: The National Academies Press. https://doi.org/10.17226/19012

Jackson, S. (2015a). Deference, distrust, and delegation: Three design hypotheses. In F. H. van Eemeren & B. Garssen (Eds.), *Reflections on theoretical issues in argumentation theory* (ch. 17). Cham, Switzerland: Springer.

Jackson, S. (2015b). Design thinking in argumentation theory and practice. *Argumentation, 29*, 243-263. https://doi.org/10.1007/s10503-015-9353-7

Jackson, S. (2019). Warrant-establishing arguments: Defending a new kind of inference. In B. Garssen, D. Godden, G. R. Mitchell & J. H. M. Wagemans (Eds.), *Proceedings of the Ninth Conference of the International Society for the Study of Argumentation,* pp. 537-547. Amsterdam: SICSAT.

Jackson, S. (2023). Online health communities in controversy over ME/CFS and Long Covid. *European Journal of Health Communication*, *4*(2), 49-72. https://doi.org/10.47368/ejhc.2023.203

Jackson, S., & Jacobs, S. (1980). Structure of conversational argument: Pragmatic bases for the enthymeme. *Quarterly Journal of Speech, 66*, 251-265. https://doi.org/10.1080/00335638009383524

Jackson, S., & Schneider, J. (2018). Cochrane Review as a "warranting device" for reasoning about health. *Argumentation, 32*(2), 241-272. https://doi.org/10.1007/s10503-017-9440-z

Jacobs, S., & Jackson, S. (1982). Conversational argument: A discourse analytic approach. In J. R. Cox & C. A. Willard (Eds.), *Advances in argumentation theory and research* (pp. 205-237). Carbondale and Edwardsville, IL: So. Ill. U. Press.

Jacobs, S., & Jackson, S. (1983). Speech act structure in conversation: Rational aspects of pragmatic coherence. In R. T. Craig & K. Tracy (Eds.), *Conversational coherence: Form, structure, and strategy* (pp. 47-66). Newbury Park, CA: Sage.

Lewiński, M., & Aakhus, M. (2023). *Argumentation in complex communication: Managing disagreement in a polylogue*. Cambridge, England: Cambridge U. Press.

Marks, H. M. (1997). *The progress of experiment: science and therapeutic reform in the United States, 1900-1990*. Cambridge: Cambridge U. Press.

National Institute for Health and Care Excellence (NICE). (2006a). The guideline manual. Retrieved from https://www.bsuh.nhs.uk/wp-content/uploads/sites/5/2016/09/The-guidelines-manual-NICE.pdf

NICE. (2006b). *Chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy): diagnosis and management of CFS/ME in adults and children.* (NICE guideline draft for consultation.) https://web.archive.org/web/20170920112048/https://www.nice.org.uk/guidance/cg53/documents/cfsme-guideline-consultation-nice-guideline2

NICE. (2006c). *Chronic fatigue syndrome/myalgic encephalomyelitis consultation table.* https://web.archive.org/web/20170920112210/https://www.nice.org.uk/guidance/cg53/documents/nice-version-table-of-comments2

NICE. (2007a). *Chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy): diagnosis and management of chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy) in adults and children*. https://web.archive.org/web/20210410215921/https://www.nice.org.uk/guidance/cg53/evidence/full-guideline-pdf-196524109

NICE. (2007b). *Chronic fatigue syndrome/myalgic encephalomyelitis or encephalopathy: diagnosis and management.* https://web.archive.org/web/20210913115350/https://www.nice.org.uk/guidance/cg53/resources/chronic-fatigue-syndromemyalgic-encephalomyelitis-or-encephalopathy-diagnosis-and-management-pdf-975505810885

NICE. (2017a) *Surveillance proposal consultation document: chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy): NICE guideline CG53 – 10-year surveillance review.* https://web.archive.org/web/20210206014037/https://www.nice.org.uk/guidance/cg53/evidence/appendix-a-preconsultation-summary-of-evidence-from-surveillance-pdf-4602203534

NICE. (2017b). *Surveillance report 2017 ... NICE guideline CG53: Appendix C Stakeholder Comments Table.* https://web.archive.org/web/20210410215933/https://www.nice.org.uk/guidance/cg53/evidence/appendix-c-stakeholder-consultation-comments-table-pdf-4602203536

NICE. (2017c). *Surveillance report 2017 – chronic fatigue syndrome/myalgic encephalomyelitis (or encephalopathy): diagnosis and management (2007) NICE guideline CG53*. https://web.archive.org/web/20200614130721/https://www.nice.org.uk/guidance/cg53/resources/surveillance-report-2017-chronic-fatigue-syndromemyalgic-encephalomyelitis-or-encephalopathy-diagnosis-and-management-2007-nice-guideline-cg53-pdf-5964455783941

NICE. (2020). *COVID-19 rapid guideline: managing the long-term effects of COVID-19*. https://www.nice.org.uk/guidance/NG188

NICE. (2021). *Myalgic encephalomyelitis (or encephalopathy)/chronic fatigue syndrome: diagnosis and management.* https://www.nice.org.uk/guidance/ng206

NICE. (2023). *Developing NICE guidelines: The manual.* https://www.nice.org.uk/process/pmg20

Schneider, J., & Jackson, S. (2018). Modeling the invention of a new inference rule: The case of 'Randomized Clinical Trial' as an argument scheme for medical science. *Argument & Computation*, *9*(2), 77-89. https://doi.org/10.3233/AAC-180036

Thornton, J., Alderson, P., Tan, T., Turner, C., Latchem, S., Shaw, E., …, & Chamberlain, K. (2013). Introducing GRADE across the NICE clinical guideline program. *Journal of Clinical Epidemiology*, *66*(2), 124–131. https://doi.org/10.1016/j.jclinepi.2011.12.007

Toulmin, S. E. (1958). *The uses of argument*. Cambridge: Cambridge U. Press.

White, P, Abbey, S., Angus, B., Ball, H. A., Buchwald, D. S., Burness, C., … & Zeman, A. (2023). Anomalies in the review process and interpretation of the evidence in the NICE guideline for chronic fatigue syndrome and myalgic encephalomyelitis. *Journal of Neurology, Neurosurgery & Psychiatry.* https://doi.org/10.1136/jnnp-2022-330463