2022

Artificial Intelligence and the Situational Rationality of Diagnosis: Human Problem-Solving and the Artifacts of Health and Medicine

Michael W. Raphael
CUNY Graduate Center

Received: 29 July 2021 | Revised: 5 October 2022 | Accepted: 7 October 2022

DOI: 10.1111/soc4.13047

REVIEW ARTICLE

WILEY

Artificial intelligence and the situational rationality of diagnosis: Human problem-solving and the artifacts of health and medicine

Michael W. Raphael

City University of New York Graduate Center, New York, New York, USA

Correspondence
Michael W. Raphael, City University of New York Graduate Center, 365 Fifth Avenue, New York, NY 10016, USA.
Email: mraphael@gradcenter.cuny.edu

Abstract

What is the problem-solving capacity of artificial intelligence (AI) for health and medicine? This paper draws out the cognitive sociological context of diagnostic problem-solving for medical sociology regarding the limits of automation for decision-based medical tasks. Specifically, it presents a practical way of evaluating the artificiality of symptoms and signs in medical encounters, with an emphasis on the visualization of the problem-solving process in doctor-patient relationships. In doing so, the paper details the logical differences underlying diagnostic task performance between man and machine problem-solving: its principle of rationality, the priorities of its means of adaptation to abstraction, and the effects of seeking optimization in the problem-solving process. Using these parameters as a heuristic for evaluating the capacity of AI to address issues of diagnostic error through design, the paper presents a conceptual review of the discipline of AI in medicine. Studies relying on procedural rationality describe models that treat diagnosis as a "natural artifact" by employing symbolic methods designed to simulate human problem-solving. Research adhering to probabilistic rationality describes models that treat diagnosis as a "natural artifact" of an ecological image by utilizing sub-symbolic methods designed to simulate neural networks. Research guided by situational rationality describes models that require treating diagnosis as a "socio-cognitive artifact," the artificiality of which is organized in discourses of patient-centered decision-making. The paper concludes with a commentary on the ethical application of AI in health and medicine, given the logical differences underlying diagnostic task performance.

KEYWORDS
artificial intelligence, artificial intelligibility, cognitive sociology, human problem-solving, practical ritualistic activity, quasi-decomposability of complexity, situational rationality

> Medicine is not a science. Instead, it is a rational, science-using, inter-level, interpretive activity undertaken for the care of a sick person. As an interpretive activity turned toward an endless succession of individuals, it takes the patient as its text and seeks to understand his or her malady in the light of current biological, epidemiological, and psychological knowledge.
>
> Montgomery Hunter (1991, p. 25)

1 | INTRODUCTION

Medical encounters are the crux of the diagnostic problem-solving process. Sociology has long understood this process by critically examining how medical discourse shapes the relation of mind and body through medical knowledge, the medical profession, and the deployment of medical expertise (Foucault, 1973; Freidson, 1970; Patel et al., 2012; Prior, 2003; Rose, 2007; Turner & Samson, 1995). However, with the study of artificial intelligence (AI) in health and medicine, these conceptual frameworks are insufficient to capture the ethical issues necessarily surrounding medical authority and diagnostic error. In the context of understanding the limits of AI as a function of human discourse, the missing heuristic I propose emphasizes artificiality and its effects on the organization of the problem-solving process. The point is to provide sociologists with a practical guide to the technicalities of the literature on engineering AI in health and medicine, thereby enabling them to sort, select, and use the relevant literature to guide their own research. The definition of this heuristic will draw from the perspective of cognitive sociology (Raphael., 2017, 2021) to clarify *what the AI actually* ***does*** *and* ***does not*** *do in the course of problem-solving*. This is in contrast with the different senses in which AI is understood (e.g., Liu, 2021).

This review analyzes artificiality by evaluating diagnostic task performance in medical encounters according to three heuristic criteria: the principle of rationality, the prioritization of design or discourse in its means of adaptation, and the effect of its capacity for optimization in human and machine problem-solving. The principle of rationality posits a model of intelligence as a means of adaptation to a problem and arriving at a solution. These means reveal how priorities of design or discourse estimate information from medical encounters. The priority of design treats health and medicine as "natural artifacts." In that case, the capacity of human intervention is conditional upon and constrained by natural law, and the heuristic assumption is that "essence precedes existence." The priority of discourse indicates that health and medicine are "socio-cognitive artifacts." In this case, the scope of human intervention in nature is secondary to the capacity of human participation to interfere with designable solutions. Here, the heuristic assumption is that "existence precedes essence." By distinguishing between "natural artifacts" and "socio-cognitive artifacts," we are acknowledging that a deconstructed truth does not necessarily provide a solution to biomedical problems (Clarke, 2010; Kerr, 2004), that is, until we understand whether or not the optimization capacity of human and machine involves the imaginative potential of situational rationality. In that case, the lack of certainty may be at least as much a resource as an obstacle, and its status as a resource is what is crucial in evaluating clinical judgment.

These criterial heuristics will help us grasp the degree to which clinical judgment focuses the problem-solving process in a way that acknowledges three basic facts related to the doctor-patient relationship. First, *the case* does not merely pose a *technical* problem, be it of organizational, occupational, professional, ideological, or profitable concern. Second, *the subject of the case* is *embodied* as a patient who has to live with the consequences of a diagnosis. Third, *a patient is a participant in their own case*, a participant who has to reconcile their own narrative (with, e.g., attempts to offer alternative accounts of potential causes) with necessities of treatment, its efficacy as a solution, and the contingencies involved in continuing treatment past the point of ostensible solution. These facts point to the *cognitive sociological context* of how human problem-solving contributes to medical encounters and locate the prospects of AI relative to that context. The remainder of this article presents the major concepts that support this claim. It begins by locating the idea of "artificiality" in its cognitive sociological context in order to show how the different priorities of design and discourse configure the problem-solving process and its application to diagnosis in medical encounters. Next, the AI literature in health and medicine will be reviewed according to three principles of rationality that allow us to distinguish the problem-solving capacity of the human from that of a machine. Research informed by procedural rationality reveals how the priority of design guides the development of knowledge-based and case-based expert systems for supporting clinical decision-making. Research adhering to probabilistic rationality reveals how the priority of design guides the development of machine learning algorithms for modeling clinical decision-making. Research based on situational rationality reveals how the priority of discourse guides the contribution of human-computer interaction to diagnosis within the doctor-patient relationship. Finally, a broader notion of artificial intelligibility will be described as presenting an ethical issue for applying AI in health and medicine.

In selecting the work to be reviewed, I began with the journal *Artificial Intelligence in Medicine* to identify relevant review articles. Then, I used Google Scholar to identify cross-references and subsequent studies. In this respect, the paper offers a conceptual overview of the problem-solving necessary for medical encounters to evaluate the extent to which AI *can contribute to medical encounters*. For a comprehensive survey of technical implementations of AI, see a review of 185 methods (Pandey & Mishra, 2009); for the scope of more recent applications within the medical domain, see two recent surveys of intelligence-based medicine (Chang, 2020; Topol, 2019b). For a historical overview of the debates within the field of AI, see a review of the controversy between symbolic and sub-symbolic information processing (Cardon et al., 2018; Mitchell, 2019).

2 | THE ARTIFICIALITY OF HEALTH AND MEDICINE AND MEDICAL ENCOUNTERS

Sociological studies of science and technology describe artificiality as "the results of work by scientists and engineers" in integrating "human and nonhuman actors in analyses of the construction of knowledge and things" (Cf. Krippendorff, 2007; Sismondo, 2008). However, for medical sociology and its relationship to AI and problem-solving, we require a more archaic meaning and explanation: a state of affairs that provides "evidence of human artifice," as Herbert Simon described it in *The Sciences of the Artificial* ([1968]1996). Simon distinguishes artificial things from natural things. Artificial things (a) are synthesized by human beings, (b) imitate appearances in natural things while lacking the reality of natural things, (c) specify functions, goals, and adaptation that characterize them, and (d) their design is discussed in descriptive and imperative terms. It is in regard to these characteristics that Simon characterizes "artificiality" as connoting "perceptual similarity but essential difference, resemblance from without rather than within." (Simon, [1968]1996, p. 13) The philosophical challenge is that "those things we call artifacts are not apart from nature," that "a forest may be a phenomenon of nature; a farm certainly is not." In this sense, health, disease, illness, and sickness are irrefutably artificial in light of how human organization, as the distribution of conditions and constraints that shape the social environment and its resources, intervenes with natural life expectancies. The challenge for diagnosis is to sort out the configuration of artificiality by "the visualization of symptoms and signs as forming a clinical picture of some pathological process" (Houghton et al., 2010). This section will describe how problem-solving processes differentially prioritize design and discourse and their impact on this process of visualization.

Studies of problem-solving reveal two basic models of how natural artifacts prioritize design in explaining how the "adaptation of means to environments is brought about." Design is undertaken within an information processing system to arrive at a solution that aligns human ends and natural means.[1] This alignment is achieved either through simplification or through complication in the problem-solving process. The design of simplification utilizes procedural rationality to hierarchically, iteratively, and recursively identify nontrivial redundancies that structure a maze; the structure of this maze enables the removal of ambiguity and the selection of criteria for halting search. That is, it defines a procedure for verifying that "the end of the maze" has been reached (Kahneman & Tversky, 1979; Simon, 1981, 1995a, 1995b, 1995c). The design of complication utilizes probabilistic rationality to iteratively and recursively simulate mathematically tractable possibilities that provide a "shape to the mold." This shape is provided as an ecological image in which a branching tree of achievable situations is defined either by a well-structured domain (e.g., rules of a game), or a classification of examples and non-examples, that is, to refine hypotheses about the fit of multi-level representations to external data (Gibson, 1977; Turner, 2018). Accordingly, adaptation by the problem-solver is a response to the complexity of the system or to the complexity of the environment: a response in which the degree of selectivity is a property of the design for which the scope of an intervention is primarily constrained by natural laws. In this respect, human problem-solving favors procedural rationality due to the constraints of bounded rationality, the fact that attention is a scarce resource in which the size of memory and processing capacity are inextricably linked (Simon, [1947]1997, p. 129). Machine problem-solving, contrastingly, is only constrained by the scalability of resources in which the efficacy of iterative and recursive processing is a function of software and hardware (Nagarajan & Stevens, 2008).

Studies emphasizing the priority of discourse reveal the constitution of socio-cognitive artifacts through "processes of reification" in arriving at a solution in which human participation interferes with our own human interventions with nature (Raphael, 2017). By "discourse," we refer to "doing and speaking" in the sense of "collective enunciation," that "to participate in discourse is to be caught up in a certain momentum that is independent of individually specific motives or intentions."[2] Processes of reification have aspects that are functionally ineradicable and aspects that are invariably in motion. The functionally ineradicable aspects refer to the fact that the arrival at a solution requires a momentary representation of the problem—to be treated as a thing—in which the truncation of the representation of our experience is unavoidable. The aspect that is not unavoidable is the degree to which the semblance of an artifact presents a "predicament of intelligibility," thereby presenting an invitation to participate in the problem-solving process through "practical ritualistic activity" (Goffman, 1967, 1974; Raphael, 2013).[3] The functionally ineradicable aspect relies on a principle of reification to guide the selectivity of its processes; the invariably in motion aspects constitute the complexity of the situation: "inferential functions" that describe the cognitive relation between the "degree of relatedness" and the "degree of decomposability" in discourse (an "inferential function" is what is accomplished by the impression of a set of expressions). In other words, processes of reification seek to explain the situational rationality of whatever is "retrieved and held up for inspection," instead of resorting to a natural belief or a natural artifact.[4] Thus, the priority of discourse reveals the immanence of information itself within the problem-solving process. This implies that reference, representation, and computation in the problem-solving process are not topically given to perception; the relatedness and decomposability of its "quasi-essence" is thoroughly situational. This is consequential for problem-solving since a proper grasp of the imaginative aspect of the representation in the situation indicates the degree to which participation allows for engrossment, and the degree to which engrossment realizes the balance of inferential functions appropriate to the object, its transformation, and the motion of its rhetorical effect. In sum, the complexity of discourse highlights how the distinctions between "situation," "occasion," and "context" refer to a non-recursive ambiguity that accounts for the social aspects of explaining information processing—without treating the social as the accidental or elective presence of other individuals or as an aggregation of individual cognitions (e.g., Dazeley et al., 2021). Crucially, then, socio-cognitive artifacts explain how participants may succeed in utilizing the same principle of reification while they may nevertheless fail to participate appropriately in practical ritualistic activity.

For medical encounters, the two different priorities in arriving at solutions determine whether or not *the visualization process of the clinical picture is dependent on the doctor-patient relationship*. If the case is a natural artifact, then a diagnosis is merely a technical problem. The accreditation of selectivity, the source of complexity, and the authority of the solution are external to the problem-solving process in that the design of diagnostic possibilities is determined in advance by standardized single disease management guidelines that provide step-by-step instructions about the most intimate details of medical care (Collins, 2017; Elstein & Schwarz, 2002; Timmermans & Angell, 2001). If the case is a socio-cognitive artifact, then a diagnosis is not merely a technical problem. The accreditation of selectivity, the source of complexity, and the authority of the solution are derived from the situational rationality of the doctor-patient relationship. Crucially, since this artificiality does not primarily prioritize human intervention in nature, it must be noted here that the concern is with how human participation generates symptoms and signs that obscure technical problems. For example, Mishler (1984) observed that physicians and patients often pursue distinctly, and sometimes conflicting, agendas in a medical visit, leading physicians recurrently to suppress the patients' concerns, even though they can be important sources of evidence of further medical problems. Similarly, Waitzkin (1991) reports that denying the expression of personal troubles reinforces the patient's accommodation to the social contexts in which illness arises. This clinical picture becomes particularly important when mental symptoms are the subjects of scrutiny. As Goffman (1969) explains, mental symptoms are typically disruptions in social relationships and organizations. These disruptions are experienced through socio-cognitive artifacts of place in which alienation, rebellion, insolence, untrustworthiness, hostility, apathy, and intrusiveness are expressed as indications of pathological identity. Processes of reification thus reveal the depths to which discourse reaches into our situational sense of sanity; without the confirming or disconfirming the reality of social status relative to the expectations and obligations found in public places, workplaces, and the family, whether analog or digital, we would have no frame of reference to draw from as the basis for participation.

Thus, the diagnostic problem-solving involved in medical encounters is oriented toward building a clinical picture based on two incommensurate tasks: assessing the intervention required for the body and assessing the extent to which human participation interferes with the focus on the body. For the first task, symptoms and signs are natural artifacts of a biomedical framework: the case and its medical history is independent of the patient in which highly detailed information about the patient's past and present medical conditions, past diagnoses, disease progressions, treatments, lab test results, and the like are reliably retrievable such that any technical problems are algorithmically discernible as properties in need of treatment through some procedural intervention (e.g., pharmaceutical, surgical, chronic monitoring). For the second task, the existence of symptoms and signs precedes their essence as socio-cognitive artifacts: the constitution of the case is dependent on the doctor-patient relationship in which a documented medical history serves as an indication for memory and not a substitute for the narrative of the patient's embodied experience. In this sense, symptoms and signs have an ambiguous character in which the discourse among the doctor and the patient have to participate appropriately in practical ritualistic activity such that the deferred meaning of doing and speaking is reconciled with what has been done and said, especially in the accreditation of selectivity. In other words, clinical questions are not formulated by treating the patient as a demographic profile; the information the patient seeks to provide—and the information the doctor seeks—mutually acknowledges the experiential consequentiality of the patient's condition and the difficulties the patient faces in coming to terms with the information the doctor provides. That is, an acknowledgment that continuing treatment past the point of solution is not merely a commitment to procedural rationality. It requires explaining how procedural rationality is reconcilable with the situational rationality that the patient continually experiences, especially in the case of chronic conditions where outpatient treatments face many contingencies in being properly administered by less-skilled practitioners (e.g., family, home-health aides). Because of this, the task cannot be adequately designed in advance, nor can its performance be subjected to selective trial-and-error or full simulation. The discourse of this visualization process is thoroughly a non-recursive conversation in which discerning treatment requires handling the complexities of discourse and the motion of its rhetorical effects.

3 | THE SITUATED RATIONALITY OF DIAGNOSIS: TASK PERFORMANCE AND THE PROSPECTS FOR ARTIFICIAL INTELLIGENCE

The previous section highlighted how diagnostic problem-solving involves a visualization composed of two incommensurate tasks: one prioritizing design and one prioritizing discourse. The first task takes the body as its focus: the burden of the problem-solving process on the physician's own bounded rationality can be reduced considerably through design while improving their problem-solving performance in visualizing symptoms and signs. The second task focuses on the situation: the burden of the problem-solving process must be shared with the patient so that the visualization of symptoms and signs relates the scope of reference to its degree of specialization in specifying its realm of meaning. The following sub-sections examine how the performance of each of these tasks is described in the literature regarding three principles of rationality: procedural, probabilistic, and situational. Elaborating on these principles will help us distinguish the problem-solving capacity of a human from that of a machine to understand (and evaluate) the relationship between AI and medical ethics.

3.1 | Procedural rationality and the symbolic visualization of bodily diagnosis

The principle of procedural rationality is that the structure of the maze and its simplification determines whether or not the design of a natural artifact is optimizable. If the size and scope of the maze are consistent with a well-defined logic of representation, then an algorithm can be designed to approach and then reach the end of the maze. An algorithm is a temporally ordered set of actions taken by design to anticipate success, that is, to arrive at an optimal solution, for example, the maximization of utility. This assumes that each discrete step can be evaluated in advance with enough foresight to validate their contribution to achieving an optimal solution. If the logic of representation is ill-defined, then the hierarchically guided elimination of ambiguity will progress until the size and scope of the maze are fully determined. If the size and scope of the problem representation exceed computational capacity, then a heuristic is designed. A heuristic is a set of actions taken, prescribed as "rules of thumb," toward a satisfactory solution by enabling a degree of selectivity otherwise inadequate to the task. Rather than systemically searching through the maze, rules of thumb are utilized to derive a criterion for halting search when a satisfactory solution is found. This criterion is iteratively and recursively derived by decomposing the maze through a succession of means-ends analyses (trial and error) so that solving smaller problems will lead to solving the larger problem.

The essential assumption of these goal-seeking instructions is that the difference between these computational forms lies in the degree to which the solvability of the problem is given by the "symbolic" character of a domain or concept. This implies that computation merely involves "storing symbols, and inputting, outputting, organizing and reorganizing such symbols and symbol structures, comparing them for identity or difference, and acting conditionally on the outcomes of the tests of identity" (Newell & Simon, 1956). The simplifying priority of design for the use of algorithms implies that achieving an ideal solution is possible with enough engineering foresight (Chandel & Sood, 2014). The simplifying priority of design for the use of heuristics implies that achieving foresight regarding all possible alternatives is too costly, and is not merely a matter of engineering. It thereby requires us to "satisfice" instead by aiming "at the good when the best is incalculable" (Simon, 1956, 1979). In other words, the efficacy of algorithms depends on how the maze is systematically searched, whereas the efficacy of heuristics expresses the integration of information and skill in order to forgo an extensive systematic search.

The significance of this symbolic character lies in its prospect of encompassing an entire domain of goal-seeking instructions, referred to as an "expert system." The programming of expert systems identifies the problems appropriate to the domain. It gradually develops a body of strategies as more is learned about the domain in which larger sets of inference rules are necessary to maintain the logical coherence of the symbolic information processing system. In this respect, since the design of algorithms anticipates a successful systematic search, procedural rationality tends to fail when facing large domains. It was only through the addition of heuristic evaluation functions that enabled expert systems to develop enough to construct and resolve logical inconsistencies. The medical encounter equivalent is the development of clinical support systems that provide computer-assisted diagnoses (Miller et al., 1982). These systems help doctors cope with the overwhelming amount of data and knowledge produced by medical technology (Cristiani et al., 2019; Huang et al., 1993). These systems tend to symbolize natural artifacts in two ways: knowledge-based and case-based.

Knowledge-based expert systems separate parts from the whole by storing information about the domain separately from information about classification that brings a whole together. For diagnosis, the system stores individual disease profiles that list findings that can occur in patients with each illness. This enables a computer program to utilize heuristic principles to select a hierarchical classification principle to produce an exhaustive differential diagnosis for each finding. Extensive studies have led to the development of computer programs to act as quick medical references, like a simple electronic textbook customized by the user to generate hypotheses (Lucas, 2001; Miller, 2009). From an intelligence perspective, these systems offer nothing like a visualization of the patient within the discourse of the medical encounter. Its rules merely allow doctors to improve their own decision-making by mediating the difference between acknowledging the *validity* of the list of profiles (possible cases) and converting data into situationally rationalized information, data that is generated to a significant extent through discourse.

Case-based expert systems construct wholes by identifying similarities and differences by recalling and reusing specific knowledge obtained from records of past experience. This enables the system to handle unexpected cases by imposing otherwise missing input values based on formal inference from previously indexed wholes (profiles). These wholes are accessed by a user through a series of question prompts that updates the lists of similar relevant stored cases and suitable questions through "dialogue inferences" (Ekerholt et al., 2014, pp. 51–55). For diagnosis, this adds an additional interactive component to a customized electronic textbook by highlighting matching features between the target case and the problem description (Hillig & Müller, 2020). From an intelligence perspective, while these systems provide an alternative interface for establishing relevance in the search process, there is still no visualization of the patient, and this is necessary if the exceptional case is to be recognized medically in its possible (and probable) exceptionality.

These two different implementations of procedural rationality noticeably stress the doctor's need to cope with bounded rationality, external limitations, given the conservative estimate that about 12 million significant misdiagnoses occur per year in the United States (Singh et al., 2014). These diagnostic errors often arise from how humans deal with natural artifacts (e.g., failing to order the right test, misinterpreting a test that was performed, failing to consider a proper differential diagnosis, or missing an abnormal finding). At the same time, while clinical support systems can build in redundancies to reduce these sorts of errors, these systems are nevertheless limited in their capacity to simulate the aspects of human problem-solving that contribute to successful task performance, namely how heuristics computationally operate heuristically in a situation, and not as a formalized algorithmic evaluation function. In other words, the greatest strength of expert systems—to make rule-based inferences toward a goal—is also its greatest weakness since the crux of rule-following in problem-solving is *knowing when rules must be broken*, and the visualization of that possibility requires intelligibility, and *not mere intelligence*. For diagnosis, this means that the visualization of the body must not merely be an algorithmic model of the patient's case, but a dynamic image of the body as a natural artifact.

3.2 | Probabilistic rationality and the sub-symbolic visualization of bodily diagnosis

The principle of probabilistic rationality is that the complication of an ecological image enables the optimization of a natural artifact's design. This implies that the alignment of human ends with natural means requires fitting the case to the environment rather than to a domain-defined maze. Search methods instead optimize the probabilities in the data, probabilities that cannot be programmed by a symbolic logic of representation. Algorithms instead model networks of neurons where the strength and weakness of connections between neurons depend on the weight given to the inputs calculated in its sum. Crucially, these networks allow for the parallel processing of information in which neurons that are wired together, fire together. This computational mode complicates the problem-solving process by "sub-symbolically" encoding the problem in a mathematically tractable manner. In this sense, the solvability of a problem is a function of the training of the network and its methods of machine learning. A basic artificial neural network begins its training with randomized weights and connections, which are gradually organized into layers. Such a layer is modeled as a column of units—with each unit representing a neuron. Layers are divided between hidden layers and an output layer. Hidden layers contain non-output units that connect the inputs with the output layer. Optimization thus describes how a model learns the activation values of the neurons as a natural artifact of "the parameters in the objective function from the given data" (Sun et al., 2020).

Machine learning optimizes the probabilities in the model by training on examples to determine the weights and thresholds that would produce correct answers. This self-organizing training occurs in one of three ways: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning trains the network by beginning with a data structure defined by a programmer who defines a set of desired outcomes for a range of inputs. Through continual feedback, hypotheses about the relevant features are tested to fine-tune its classification of examples and non-examples. Unsupervised learning trains the network by beginning with a data structure that has no desired outcomes. Instead, co-occurring features are found to produce patterns or clusters in the data; it mines for patterns to create its own classification of examples and non-examples. In short, these two methods utilize feedback to create a path with no conception of the maze; it merely identifies a well-refined prediction. Reinforcement learning, in contrast, utilizes feedback to refine decision-making through reward or punishment, represented by a non-binary evaluation function as a positive or negative score. The evaluation function does not need to remember previous states; it only needs to learn values that are good predictions for rewards. Reinforcement learning thus creates a way to navigate the maze with no conception of the maze or its paths (Silver et al., 2021). The efficacy of these training methods comes from their integration into multilayered neural networks where the multiple hidden layers allow for deep learning, a method of encoding the results as multi-level representations (Koteluk et al., 2021).

The significance of this sub-symbolic character lies in its capacity to forgo the weaknesses of expert systems. Whereas expert systems allow for patterns to be recognized through crafted inference rules, machine learning develops its own evaluation function, enabling the development of natural language processing (speech recognition) and computer vision (image recognition) without the prescribed limitations imposed by the serial processing of rules. The medical encounter equivalent to this is the use of medical imaging and electronic health records (EHR) as natural artifacts. The use of medical imaging visualizes the body in a way that is imperceptible to the naked eye. In this respect, four datasets are generated: magnetic resonance images, computerized tomography scans, X-ray images, and ultrasound images. Each of these images provides different diagnostic information. Accordingly, a classifier needs to be trained for each purpose, typically by differentiating between anatomical template images and histopathology images within each dataset—to encode a multi-level representation that recognizes how the arrangement of pixels in one set of images indicates a visualized natural artifact of the body in a healthy condition and a body in a diseased condition. Additionally, at a minimum, this diagnostic modeling needs to account for scanner variability, scan acquisition settings, subject demography, and heterogeneity in disease characteristics across subjects—all of which affect the design of the natural artifact (Barragán-Montero et al., 2021). The use of EHR systems visualizes the body as data. In this respect, these data serve two functions that guide their basic design: diagnostic support and organizational communication. The diagnostic support function requires storing and maintaining patient demographics, medication lists, allergies, diagnoses, test results, clinical assessments, and treatments as well-defined data structures that include longitudinal data and hierarchically organized information. The organizational communication function supports the diagnostic and treatment process by streamlining the management and coordination of orders, results, and case reporting. When these natural artifacts are built into a dataset for machine learning, the scale numbers in hundreds of billions of data points. The primary challenge in building these datasets is that the natural artifacts of the organizational communication function are often structured in non-standardized machine-readable formats across institutional lines, although attempts are being made through the development of the openEHR clinical modeling program (Fernandes et al., 2020; Wulffet al., 2018).

From an intelligence perspective, while the visualization of the patient's symptoms and signs are static in medical imaging and dynamic in EHRs, the application of probabilistic rationality through machine learning diagnostic algorithms remains largely untested as part of a clinician's regular course of work (Rajkomar et al., 2019). For medical imaging analysis, when the task is to retrospectively identify abnormalities across a variety of image types, these algorithms perform well. For the analysis of EHRs, when the task is to retrospectively predict overall patient outcomes, the treatment of sepsis, and cancer therapy, these algorithms performed well (Lin et al., 2021). However, for the goal of precision medicine, to fit data analytics to a single patient rather than a population, many machine learning studies continue to be limited to a single data modality, meaning that such algorithms are still at the early development stage (MacEachern & Forkert, 2021). In other words, sub-symbolic methods of diagnosis are substantive areas of research that remain under-utilized in prospective real-world settings. They are under-utilized because dataset size is not everything. While datasets are slowly getting larger (including more incomplete and noisy data), predictive performance is not improving at the same rate. In overcoming the rules of procedural rationality, with their limitations acting as a source of error, the socio-cognitive context of probabilistic rationality and its weaknesses are slowly being revealed: systemic human biases or errors in producing data are being reinforced by the natural artifact's design (Kudina & Boer, 2021). This socio-cognitive context implies that the ability of these machine learning algorithms to attach meaning to detected patterns remains limited to its scope of probabilistic rationality. For example, the diagnoses from which the models are built may be provisional or incorrect, in which conditions that do not manifest recordable symptoms are outside the scope of the data. Accordingly, these models should be restricted from autonomous diagnosis-making and relegated to helping doctors cope with the limits of their own bounded rationality.

3.3 | Situational rationality and the visualization of diagnosis in human-computer interaction

The principle of situational rationality is that the demands of discourse—without and within practical ritualistic activity—generate the configuration of artificiality and complexity in the problem-solving process as socio-cognitive artifacts of how human participation responds to predicaments of intelligibility. This implies that the means of adaptation for human interventions with nature cannot achieve optimization by merely prioritizing design over discourse since human participation necessarily involves socially constituted processes of reification. These processes of reification situate conditions and constraints in which the information processing capacity of participants in the problem-solving process is a function of the distribution of resources and the constitution of bounded rationality. In other words, the socio-cognitive context of the problem-solving process—the *immanence* of information (its degrees of selectivity, decomposability, and relatedness) and the authority of the solution—interferes with the scaling of information processing systems to the level of human organization precisely because the priority of design cannot solve the demands of discourse in advance. Accordingly, principles of reification are situationally selected to momentarily address the deferral of meaning (uncertainty), ambiguity, and the imaginative potential of continuing past the point of immediate solution. Thus, search methods are not merely a function of a computation in which iterative recursion simplifies or complicates a logic of representation—either through procedural or probabilistic rationality. Instead, for human problem-solving, search methods have to make intelligible the symbolic serial and sub-symbolic parallel processing of discourse by balancing and reconciling situationally appropriate principles of reification. Accordingly, the correspondence between reference, representation, and computation cannot be given in advance; its ratio can only be estimated as a momentary belief about the size and scope of the problem and its solvability.

Momentary beliefs describe how socio-cognitive artifacts acquire their "quasi-essence" through discourse and the motions of its rhetorical effects. In this sense, the existence of domains and concepts are dependent on participation in practical ritualistic activity: the manner in which they are "retrieved and held up for inspection" affects how they are socio-cognitively perceived, constituted in memory, and made meaningful for estimating the demands of a solution and its implementation. The literature reveals that this situational rationality is consequential for design in two respects. First, design must communicate the limits of machine problem-solving, that is, how its solutions are natural artifacts. Second, design must communicate that its lack of capacity for socio-cognitive artifacts implies that automation itself will not address how communication is itself a significant source of error in decision-making. These two challenges of discourse for design show the effect of socio-cognitive artifacts in medical encounters through clinical support systems and the use of patient decision aids in shared decision-making.

Studies of diagnostic human-computer interaction show that the efficacy of machine problem-solving for clinical support systems is limited by how socio-cognitive artifacts determine the context, quality, and prospective value of medical information. The *contextual* character of medical information is a socio-cognitive artifact of medical discourse that demanded its generation: doctors are aware of the constantly evolving temporal movement of a medical case, and, to some extent, they allow the case to build itself. However, in order to build a model, design eliminates the contextual character of medical information in EHRs by decontextualizing the information into data, which is then modeled in its own context, a context that is wholly dependent on a strictly formal construction that is unlikely to be compatible with that revealed in the reciprocity of discourse. The design of this natural artifact cannot, on its own, make the data speak for themselves. As Berg and Goorman (1999) explain, “the further information has to be able to circulate, the more work is required to disentangle the information from the context of its production,” meaning that a socio-cognitive artifact will result from the motion of its rhetorical effects when it is placed in discourse to arrive at a diagnosis. The *quality* of medical information is a socio-cognitive artifact of how poor clinical evidence is deeply embedded within published medical research and the treatment guidelines informed by them. Ioannidis et al. (2017) report that this “medical information mess” poses an obstacle to the effectiveness of clinical support systems for two reasons. First, few clinical studies, even in the major general medical journals, are able to avoid exaggerating their results or overestimating treatment effects because of a reliance on methodological designs that have limited generalizability. Second, among those practitioners who report feeling confident to evaluate the medical literature, there is often “a lack of the basic skills required for determining a study's reliability and applicability.” For these reasons, the design of natural artifacts, a machine learning model based on clinical studies, further contributes to misinformation, even if it relies randomized controlled trials (Ioannidis, 2016). The *prospective value* of medical information is a socio-cognitive artifact of its *explainability* and its *intelligibility*. The explainability of a model for medical diagnosis is contestable in terms of the data used, its bias, and its accuracy (Ploug & Holm, 2020; Zhang et al., 2021). The intelligibility of a model relies on how its visual design communicates ambiguity within clinical discourse. This poses unique challenges for formalized design since the goal of such visualizations are to become “so highly situated, so fitting, so natural” that they are “unremarkable,” even in the context of purely professional clinical decision making (Yang et al., 2019, 2020). Yet, such design principles are at odds with what doctors require for “effective human–AI collaboration” as information about the “basic, global properties of the model” (i.e., its overall design objective, its known strengths and limitations, and its subjective point-of-view) requires placing the model within a discourse (Cai et al., 2019). As Marathe and Toyama (2020) report, this quantification and visualization of physiological indicators are insufficient for expert problem-solving in which the narrow focus of formalized systems on “the individual patient and their body is not borne out in real-world hospital settings.” Even the diagnosis of epileptic seizures is a deeply situated process in which design decisions may exclude significant portions of data that may become salient later on, as the interactive situation changes. Hence, the prospective value of medical information is limited when treated as a natural artifact without acknowledging situational rationality.

Studies of patient decision aids in shared decision-making reveal that communication remains a significant obstacle in medical encounters when clinicians underestimate patients' desires for information beyond mere data (Scalia et al., 2019). Instead, for most medical encounters involving primary care, despite intensive efforts at implementation, current evidence strongly suggests that doctors “simply jettison shared decision-making and patient-centered care” since intermediate-stakes decisions are “overwhelmingly numerous and lead to real-time limits.” When this communication is attempted, doctors are likely to provide only 1–2 min for such a conversation (Caverly & Hayward, 2020).

However, as Elwyn (2021) reports, this discourse involves "the cognitive-emotional work of becoming aware of options, of processing details about the harms and benefits of options, and coming to terms about facets of uncertainty and existential issues." When contrasted with the time therapists require to assist their patients (often 30-min weekly sessions), it is clear that many doctors prefer to address patients as merely natural artifacts. In this respect, recent attempts to integrate EHRs and patient decision aids overemphasize their own utility; by assuming that providing more "information" improves decision-making, doctors fail to assess the extent to which human participation interferes with emphasis on the body and thereby reinforces the probabilistic rationality of the modeled EHRs (Coylewright et al., 2020).

From an intelligence perspective, the situational rationality of these studies reveals that task performance for the visualization of diagnosis within medical encounters is a function of misaligning the priorities of design and discourse. Design, whether procedural or probabilistic, inevitably fails when its means of adaptation ignores the complexity and movement of discourse. While expert systems and machine learning algorithms may successfully contribute to visualizing a patient's case as a natural artifact, both applications of rationality fail to achieve an understanding of medicinal practice "beyond the conception of patients as mere bundles of diseases" (Bjerring & Busch, 2020). Human problem-solving succeeds insofar as we acknowledge that natural artifacts, though necessary, are insufficient for problem-solving. As a function of our bounded rationality, we estimate the socio-cognitive artifacts of our situational rationality as we adapt to the failures of design to account for the necessary deferral of meaning in discourse (e.g., as in implementing vaccination programs and public health guidelines in the context of misinformation). Machine problem-solving cannot design its way out of this predicament since the priority of our goals is a matter of intelligibility rather than formally-oriented intelligence. For diagnosis, this requires understanding how the meanings of natural artifacts is deferred in the problem-solving process. This point becomes difficult to ignore when we summarize what the AI actually does and does not do in the course of diagnostic problem-solving.

4 | THE ARTIFICIAL INTELLIGIBILITY OF THE CASE AND MEDICAL ETHICS

Technological innovation often chases its own ends unless it avoids the course of creating a design. Accordingly, critics of science and technology are routinely required to raise ethical concerns about the alignment of means and ends as a corrective to the ambitions and imperatives of engineering. While an ethical outline of epistemic, normative, and legal concerns is well-documented for AI in medicine (Morley et al., 2020; Peek et al., 2015; Schönberger, 2019), the focus on inconclusive, inscrutable, or misguided evidence (with its potential for unfair discriminatory outcomes and liabilities), as it is traditionally discussed, overlooks the cognitive sociological context of the problem-solving process. This paper has sought to introduce a heuristic for revealing this context: to show how the artificiality and complexity of the problem-solving process are demonstrable according to principles of rationality. To refocus the ethical debate, these principles reveal differences in models of intelligence that delineate between human and machine problem-solving, namely the degree to which an operative optimization assumption alters the priorities of design and discourse, that is, the handling of natural and socio-cognitive artifacts. I have argued that this optimization assumption poses the central ethical issue in the clinical problem-solving process for the concerns of medical ethics. By demonstrating how diagnostic problem-solving involves two incommensurate tasks, the literature reviewed in this paper reveals the sizable impact of this optimization assumption on the relationship between task performance and principles of rationality.

For procedural rationality, this optimization assumption shapes the "symbolic" character of expert systems in which the visualization of diagnosis focuses doctor-patient relationships on simplifying natural artifacts. In this respect, to maximize the utility of a knowledge-based or case-based system, information has to be simplified in order to adhere to organizationally-determined procedures and guidelines. This encoding enables search methods to structure the diagnosis as a maze in which computation is the only obstacle in the way of arriving at a solution through an evaluation function crafted by hierarchical iterative recursion.

For probabilistic rationality, this optimization assumption shapes the "sub-symbolic" character of machine learning in which the visualization of diagnosis focuses doctor-patient relationships on complicating natural artifacts. In this respect, to maximize the learning performance of the model by fitting it to its environment, the structure of information has to be complicated so that it is mathematically tractable. This encoding enables search methods to train an evaluation function as a multi-level representation without domain-crafted rules. The sub-symbolic character of this multi-level representation (with its parallel processing) implies that the computation of an evaluation function is a natural artifact of classifying examples and non-examples in a supervised, unsupervised, or a reinforced iterative recursion of a refined prediction.

For situational rationality, acknowledging the demands of discourse (with its predicaments of intelligibility) requires rejecting the optimization assumption as the achievement of an ideal solution since the artificiality and complexity of discourse cannot be *merely* encoded as natural artifacts. Instead, the visualization of diagnosis must focus on doctor-patient relationships in resolving the predicament of intelligibility by assessing the sort of intervention required for the body and assessing the extent to which human participation interferes with the body as purely natural. Accordingly, in order to determine the meaning of natural artifacts (i.e., how the context, quality, and prospective value of medical information are placed within discourse), the encoding of socio-cognitive artifacts must be estimated in terms of how the immanence of information to the decision process and the authority of the solution constitute momentary beliefs in the problem-solving process about the size and scope of the domain. This discourse allows for the visualization of symptoms and signs in order to achieve intelligibility through participation in practical ritualistic activity. Such search methods balance the internal and external conditions and constraints on the problem-solving process with the constitution of bounded rationality (and its dynamic demands of serial-parallel processing of non-recursive ambiguity).

Now, if we accept three basic facts about the doctor-patient relationship, that a patient's case does not merely pose a technical problem, that a patient is an embodied consciousness, and that a patient is a participant in establishing the intelligibility of their own case, then the medical ethical concern regarding the optimization assumption is clear for the prospects of AI and organizational design. For AI, we must remember that means and ends really do matter; that AI (autonomous or not) is of limited value without "artificial intelligibility." An information processing system achieves artificial intelligibility when it is capable of situational rationality and its requisite participation in practical ritualistic activity. This means that its task performance in the problem-solving process easily handles both natural and socio-cognitive artifacts by sorting through what demands iterative recursion and what demands a highly selective momentary belief about non-recursive ambiguity and its rhetorical effects.

As things stand, without the artificial intelligibility of the case, AI will effectuate diagnostic errors for at least three reasons. First, the *real world is not the ideal world* in which the best adaptation for *a moment* is not an ideal adaptation *in the long run*. Second, without accounting for the priority of discourse, *a system cannot account for how communication becomes formalized to match the input*, leading to a troubling "computer knows best" attitude (McDougall, 2019). Third, without accounting for the fact that human problem-solving is itself unable to optimize (that medical diagnosis necessarily involves different degrees of uncertainty), *unremarkable designs of output are likely to add more uncertainty than clarity* to the problem-solving process (Doraiswamy et al., 2020). Accordingly, artificial intelligibility acknowledges the interference of the constitution of bounded rationality with the ideal optimization of human problem-solving.

For organizational design, this cognitive sociological context reveals that the demands of adequate diagnostic problem-solving are not being met (Heritage & Maynard, 2006; Kuper et al., 2017; Topol, 2019a). Task performance continues to treat the patient's case as a technical problem in which the delivery of care and the acquisition of informed consent are often formulaic means designed to maximize organizational resources across all possible cases (Grote & Berens, 2020). In other words, the ethical issue surrounding technological innovation is that it provisions *general solutions to particular problems*: for health and medicine, this problem is generally understood to be a source of diagnostic error. The literature reviewed in this paper makes it clear that designing technology itself can only improve (and optimize) our procedural and probabilistic rationality. The ethical consequence is that we are easily distracted from addressing sources of error in situational rationality, most notably, the more difficult challenge of being relentless in facing our own limitations and inadequacies. In the course of patient-centered diagnosis and care, these inadequacies are a minefield, one we must trek by engaging in a temporally open-ended conversation about the risks to patients' lives and the character of their existence.

ACKNOWLEDGMENTS

This paper wishes to acknowledge Jeremy J. Gross, Todd Isenstadt, Michael E. Brown, and Saskia C. Hooiveld for their assistance in the preparation of this manuscript. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The author declare that there are no conflicts of interest for this work.

ORCID

Michael W. Raphael https://orcid.org/0000-0002-7744-6778

ENDNOTES

[1] It is important for readers to understand that this is a description of adaptation according to a principle of least effort, and not a principle of most effort, such as the complexities involved in scientific practices (e.g., Mol & Law, 2004; Moser et al., 2006). A principle of least effort tends to result in dichotomous polarizations whereas a principle of most effort seeks to balance the need for simplification with the need for complication.

[2] The conception of discourse as "collective enunciation" is too complex to further elaborate here beyond highlighting its thoroughly social character: "In other words, the momentousness and value of an ostensibly distinct gesture or utterance depends on its participation as a moment of a course of activity and as an instance of collective enunciation or, more broadly, of sociality as the basic fact." (Brown, 2014, p. 434)

[3] "Practical ritualistic activity" refers to the doing that balances the "substance" and "ceremony" of phasing inferential functions. In this respect, "substance" describes the aspect of activity that indicates participation in its doing has a primary importance in its own right that is independent of the situation. "Ceremony" describes the aspect of activity that indicates participation in its doing—in its own right—has a secondary importance; its primary importance indicates that participation in its doing meets the participatory demands of the situation (Goffman, 1967). These aspects are important because they highlight how the size and scope of what participants can visualize in memory has a significant bearing on the sense of what is going on, as in the playing of a game of chess or within the course of learning in a classroom. The predicament of the chess player (where the point of focus can vary from problem representations present on the chessboard to the audience) and the predicament of the student (where the point of focus often ranges from books to presentation materials and other designed mazes to the audience) describe how there is still the matter of balancing the "substance" and "ceremony" of what is going on, as in the way Goffman describes interaction rituals. The medical encounter involves the same kind of predicament with far greater consequences: a conversation between a doctor and a patient—each of which have an asymmetrical relationship in their problem-solving capacity to participate in addressing the patient's case (Cf. Foucault, 1973).

[4] Situational rationality, from the perspective of cognitive sociology, seeks to understand how the transcendental aspect of situations affects the course of the problem-solving process—where ambiguity is non-reducible as a function of the ongoing course of activity, that is, "the essential sociality of humans" (Brown, 2014, p. 176). This perspective shares similarities with the interests of science and technology scholars, but there are subtle differences in the meanings and evaluations made of the constitution of performance, materiality, reflexivity, and creativity. For example, Marres et al. (2018, pp. 31–34) argue situational rationality is "a project to recover the specificity of social forms in the face of ANT's indifference." For cognitive sociology, situational rationality is not a project; rather, it is an illumination of the communicative aspects that are irreducibly social—aspects that are not merely a matter of social research describing the particularities of practices in need of change.

REFERENCES

Barragán-Montero, A., Javaid, U., Valdes, G., Nguyen, D., Desbordes, P., Macq, B., Willems, S., Vandewinckele, L., Holmstrom, M., Lofman, F., Michiels, S., Souris, K., Sterpin, E., & Lee, J. A. (2021). Artificial intelligence and machine learning for medical imaging: A technology review. *Physica Medica*, *83*, 242–256. https://doi.org/10.1016/j.ejmp.2021.04.016

Berg, M., & Goorman, E. (1999). The contextual nature of medical information. *International Journal of Medical Informatics*, *56*(1–3), 51–60. https://doi.org/10.1016/s1386-5056(99)00041-6

Bjerring, J. C., & Busch, J. (2020). Artificial intelligence and patient-centered decision-making. *Philosophy & Technology*, *34*(2), 349–371. https://doi.org/10.1007/s13347-019-00391-6

Brown, M. E. (2014). *The concept of the social in uniting the humanities and social sciences*. Temple University Press.

Cai, C. J., Winter, S., Steiner, D., Wilcox, L., & Terry, M. (2019). "Hello AI": Uncovering the onboarding needs of medical practitioners for human-AI collaborative decision-making. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–24. https://doi.org/10.1145/3359206

Cardon, D., Cointet, J.-P., & Mazières, A. (2018). Neurons spike back. The invention of inductive machines and the artificial intelligence controversy. *Réseaux*, *211*(5), 173–220. https://doi.org/10.3917/res.211.0173. Retrieved from https://www.cairn-int.info/article-E_RES_211_0173--.htm

Caverly, T. J., & Hayward, R. A. (2020). Dealing with the lack of time for detailed shared decision-making in primary care: Everyday shared decision-making. *Journal of General Internal Medicine*, *35*(10), 3045–3049. https://doi.org/10.1007/s11606-020-06043-2

Chandel, A., & Sood, M. (2014). Searching and optimization techniques in artificial intelligence: A comparative study & complexity analysis. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, *3*(3), 866–871. https://doi.org/10.32628/CSEIT2062167

Chang, A. C. (2020). *Intelligence-based medicine: Data science, artificial intelligence, and human cognition in clinical medicine and healthcare*. Elsevier.

Clarke, A. E. (2010). *Biomedicalization: Technoscience, health, and illness in the U.S*. Duke University Press.

Collins, R. D. (2017). *Algorithmic diagnosis of symptoms and signs: A cost-effective approach* (4th ed.). Wolters Kluwer Heath.

Coylewright, M., Keevil, J. G., Xu, K., Dodge, S. E., Frosch, D., & Field, M. E. (2020). Pragmatic study of clinician use of a personalized patient decision aid integrated into the electronic health record: An 8-year experience. *Telemedicine and e-Health*, *26*(5), 597–602. https://doi.org/10.1089/tmj.2019.0112

Cristiani, M., Olivieri, F., & Tomazzoli, C. (2019). Diagnostics as a reasoning process: From logic structure to software design. *Journal of Computing and Information Technology*, *27*, 43–57. https://doi.org/10.20532/cit.2019.1004411

Dazeley, R., Vamplew, P., Foale, C., Young, C., Aryal, S., & Cruz, F. (2021). Levels of explainable artificial intelligence for human-aligned conversational explanations. *Artificial Intelligence*, *299*, 103525. https://doi.org/10.1016/j.artint.2021.103525

Doraiswamy, P. M., Blease, C., & Bodner, K. (2020). Artificial intelligence and the future of psychiatry: Insights from a global physician survey. *Artificial Intelligence in Medicine*, *102*, 101753. https://doi.org/10.1016/j.artmed.2019.101753

Ekerholt, M., Follesø, S. L., & Heimark, Ø. (2014). *A system for conversational case-based reasoning in multiple disease medical diagnosis*. (Master Thesis). Norwegian University of Science and Technology.

Elstein, A. S., & Schwarz, A. (2002). Evidence base of clinical diagnosis: Clinical problem solving and diagnostic decision making: Selective review of the cognitive literature. *BMJ*, *324*(7339), 729–732. https://doi.org/10.1136/bmj.324.7339.729

Elwyn, G. (2021). Shared decision making: What is the work? *Patient Education and Counseling*, *104*(7), 1591–1595. https://doi.org/10.1016/j.pec.2020.11.032

Fernandes, M., Vieira, S. M., Leite, F., Palos, C., Finkelstein, S., & Sousa, J. M. (2020). Clinical decision support systems for triage in the emergency department using intelligent systems: A review. *Artificial Intelligence in Medicine*, *102*, 101762. https://doi.org/10.1016/j.artmed.2019.101762

Foucault, M. (1973). *Birth of the clinic: An archaeology of medical perception*. Pantheon.

Freidson, E. (1970). *Profession of medicine: A study of the sociology of applied knowledge*. University of Chicago Press.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology (p. ix, 492p.)*. Lawrence Erlbaum Associates.

Goffman, E. M. (1967). *Interaction ritual: Essays on face-to-face behavior*. Doubleday.

Goffman, E. M. (1969). The insanity of place. *Psychiatry: Journal of Interpersonal Relations*, *32*(4), 357–387. https://doi.org/10.1080/00332747.1969.11023600

Goffman, E. M. (1974). *Frame analysis: An essay on the organization of experience*. Harper & Row.

Grote, T., & Berens, P. (2020). On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics*, *46*(3), 205–211. https://doi.org/10.1136/medethics-2019-105586

Heritage, J., & Maynard, D. W. (2006). Problems and prospects in the study of physician-patient interaction: 30 years of research. *Annual Review of Sociology*, *32*(20062006), 351–374. https://doi.org/10.2307/29737743

Hillig, S., & Müller, R. (2020). How do conversational case-based reasoning systems interact with their users: A literature review. *Behaviour & Information Technology*, *40*(14), 1–20. https://doi.org/10.1080/0144929x.2020.1767207

Houghton, A. R., Gray, D., & Chamberlain, E. N. (2010). *Chamberlain's symptoms and signs in clinical medicine: An introduction to medical diagnosis* (13th ed.). Hodder Arnold.

Huang, J., Fox, J., Gordon, C., & Jackson-Smale, A. (1993). Symbolic decision support in medical care. *Artificial Intelligence in Medicine*, *5*(5), 415–430. https://doi.org/10.1016/0933-3657(93)90034-z

Ioannidis, J. P. A. (2016). Why most clinical research is not useful. *PLoS Medicine*, *13*(6), e1002049. https://doi.org/10.1371/journal.pmed.1002049

Ioannidis, J. P. A., Stuart, M. E., Brownlee, S., & Strite, S. A. (2017). How to survive the medical misinformation mess. *European Journal of Clinical Investigation*, *47*(11), 795–802. https://doi.org/10.1111/eci.12834

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263. https://doi.org/10.2307/1914185

Kerr, A. (2004). *Genetics and society: A sociology of disease*. Routledge.

Koteluk, O., Wartecki, A., Mazurek, S., Kolodziejczak, I., & Mackiewicz, A. (2021). How do machines learn? Artificial intelligence as a new era in medicine. *Journal of Personalized Medicine*, *11*(1), 32. https://doi.org/10.3390/jpm11010032

Krippendorff, K. (2007). An exploration of artificiality. *Artifact*, *1*(1), 17–22. https://doi.org/10.1080/17493460600610848

Kudina, O., & Boer, B. (2021). Co-designing diagnosis: Towards a responsible integration of machine learning decision-support systems in medical diagnostics. *Journal of Evaluation in Clinical Practice*, *27*(3), 529–536. https://doi.org/10.1111/jep.13535

Kuper, A., Veinot, P., Leavitt, J., Levitt, S., Li, A., Goguen, J., Schreiber, M., Richardson, L., & Whitehead, C. R. (2017). Epistemology, culture, justice and power: Non-bioscientific knowledge for medical training. *Medical Education*, *51*(2), 158–173. https://doi.org/10.1111/medu.13115

Lin, A. L., Chen, W. C., & Hong, J. C. (2021). Electronic health record data mining for artificial intelligence healthcare. In *Artificial intelligence in medicine* (pp. 133–150).

Liu, Z. (2021). Sociological perspectives on artificial intelligence: A typological reading. *Sociology Compass*, *15*(3). https://doi.org/10.1111/soc4.12851

Lucas, P. (2001). Symbolic diagnosis and its formalisation. *The Knowledge Engineering Review*, *12*(2), 109–146. https://doi.org/10.1017/s0269888997002026

MacEachern, S. J., & Forkert, N. D. (2021). Machine learning for precision medicine. *Genome*, *64*(4), 416–425. https://doi.org/10.1139/gen-2020-0131

Marathe, M., & Toyama, K. (2020). The situated, relational, and evolving nature of epilepsy diagnosis. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW3), 1–18. https://doi.org/10.1145/3432916

Marres, N., Guggenheim, M., & Wilkie, A. (Eds.). (2018). *Inventing the social*. Mattering Press.

McDougall, R. J. (2019). Computer knows best? The need for value-flexibility in medical AI. *Journal of Medical Ethics*, *45*(3), 156–160. https://doi.org/10.1136/medethics-2018-105118

Miller, R. A. (2009). Computer-assisted diagnostic decision support: History, challenges, and possible paths forward. *Advances in Health Sciences Education*, *14*(S1), 89–106. https://doi.org/10.1007/s10459-009-9186-y

Miller, R. A., Pople, H. E., & Myers, J. D. (1982). Internist-I, an experimental computer-based diagnostic consultant for general internal medicine. *New England Journal of Medicine*, *307*(8), 468–476. https://doi.org/10.1056/nejm198208193070803

Mishler, E. G. (1984). *The discourse of medicine: Dialectics of medical interviews*. Ablex Pub. Corp.

Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. Farrar, Straus and Giroux.

Mol, A., & Law, J. (2004). Embodied action, enacted bodies: The example of hypoglycaemia. *Body & Society*, *10*(2–3), 43–62. https://doi.org/10.1177/1357034x04042932

Montgomery Hunter, K. (1991). *Doctors' stories: The narrative structure of medical knowledge*. Princeton University Press.

Morley, J., Machado, C. C., Burr, C., Cowls, J., Joshi, I., Taddeo, M., & Floridi, L. (2020). The ethics of AI in health care: A mapping review. *Social Science & Medicine*, *260*, 113172. https://doi.org/10.1016/j.socscimed.2020.113172

Moser, I., Jacucci, E., & Law, J. (2006). Fluids or flows? Information and qualculation in medical practice. *Information Technology & People*, *19*(1), 55–73. https://doi.org/10.1108/09593840610649961

Nagarajan, N., & Stevens, C. F. (2008). How does the speed of thought compare for brains and digital computers? *Current Biology*, *18*(17), R756–R758. https://doi.org/10.1016/j.cub.2008.06.043

Newell, A., & Simon, H. A. (1956). The logic theory machine. *IRE Transactions on Information Theory (IT)*, *2*(3), 61–79. https://doi.org/10.1109/TIT.1956.1056797

Pandey, B., & Mishra, R. B. (2009). Knowledge and intelligent computing system in medicine. *Computers in Biology and Medicine*, *39*(3), 215–230. https://doi.org/10.1016/j.compbiomed.2008.12.008

Patel, V. L., Arocha, J. F., & Zhang, J. (2012). Medical reasoning and thinking. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning*. Oxford University Press.

Peek, N., Combi, C., Marin, R., & Bellazzi, R. (2015). Thirty years of artificial intelligence in medicine (AIME) conferences: A review of research themes. *Artificial Intelligence in Medicine*, *65*(1), 61–73. https://doi.org/10.1016/j.artmed.2015.07.003

Ploug, T., & Holm, S. (2020). The four dimensions of contestable AI diagnostics – A patient-centric approach to explainable AI. *Artificial Intelligence in Medicine*, *107*, 101901. https://doi.org/10.1016/j.artmed.2020.101901

Prior, L. (2003). Belief, knowledge and expertise: The emergence of the lay expert in medical sociology. *Sociology of Health & Illness*, *25*(3), 41–57. https://doi.org/10.1111/1467-9566.00339

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, *380*(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259

Raphael, M. W. (2013). Learning informal interaction: The problem of self-help literature for expression games. Retrieved from http://hdl.handle.net/2047/d20003073

Raphael, M. W. (2017). *Cognitive sociology*. Oxford Bibliographies Online in Sociology. Oxford University Press. https://doi.org/10.1093/obo/9780199756384-0187

Raphael, M. W. (2021). The situational rationality of task performance: Artifacts of belief in human problem-solving and artificial intelligence. *SocArXiv*. https://doi.org/10.31235/osf.io/yt9vz

Rose, N. S. (2007). *Politics of life itself: Biomedicine, power, and subjectivity in the twenty-first century*. Princeton University Press.

Scalia, P., Durand, M.-A., Berkowitz, J. L., Ramesh, N. P., Faber, M. J., Kremer, J. A. M., & Elwyn, G. (2019). The impact and utility of encounter patient decision aids: Systematic review, meta-analysis and narrative synthesis. *Patient Education and Counseling*, *102*(5), 817–841. https://doi.org/10.1016/j.pec.2018.12.020

Schönberger, D. (2019). Artificial intelligence in healthcare: A critical analysis of the legal and ethical implications. *International Journal of Law and Information Technology*, *27*(2), 171–203. https://doi.org/10.1093/ijlit/eaz004

Silver, D., Singh, S., Precup, D., & Sutton, R. S. (2021). Reward is enough. *Artificial Intelligence*, *299*, 103535. https://doi.org/10.1016/j.artint.2021.103535

Simon, H. A. ([1947]1997). *Administrative behavior* (4th ed.). The Free Press.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138. https://doi.org/10.1037/h0042769

Simon, H. A. ([1968]1996). *The sciences of the artificial* (3rd ed.). The MIT Press.

Simon, H. A. (1979). System principles. In *Models of thought* (pp. 1–6). Yale University Press.

Simon, H. A. (1981). Studying human intelligence by creating artificial intelligence. *American Scientist*, *69*, 300–309.

Simon, H. A. (1995a). Artificial intelligence: An empirical science. *Artificial Intelligence*, *77*(1), 95–127. https://doi.org/10.1016/0004-3702(95)00039-H

Simon, H. A. (1995b). Explaining the ineffable: AI on the topics of intuition, insight and inspiration. In *Proceedings of the fourteenth international joint conference on artificial intelligence* (pp. 939–948). https://doi.org/10.5555/1625855.1625977

Simon, H. A. (1995c). Near decomposability and complexity: How a mind resides in a brain. In H. J. Morowitz & J. L. Singer (Eds.), *The mind, the brain, and complex adaptive systems, Santa Fe Institute studies in the sciences of complexity, proceedings Vol. XXII* (pp. 25–44). Addison-Wesley.

Singh, H., Meyer, A. N. D., & Thomas, E. J. (2014). The frequency of diagnostic errors in outpatient care: Estimations from three large observational studies involving US adult populations. *BMJ Quality & Safety*, *23*(9), 727–731. https://doi.org/10.1136/bmjqs-2013-002627

Sismondo, S. (2008). Science and technology studies and an engaged program. In *The handbook of science and technology studies* (Vol. *1*, pp. 13–31).

Sun, S., Cao, Z., Zhu, H., & Zhao, J. (2020). A survey of optimization methods from a machine learning perspective. *IEEE Transactions on Cybernetics*, *50*(8), 3668–3681. https://doi.org/10.1109/tcyb.2019.2950779

Timmermans, S., & Angell, A. (2001). Evidence-based medicine, clinical uncertainty, and learning to doctor. *Journal of Health and Social Behavior*, *42*(4), 342–359. https://doi.org/10.2307/3090183

Topol, E. J. (2019a). *Deep medicine: How artificial intelligence can make healthcare human again* (1st ed.). Basic Books.

Topol, E. J. (2019b). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Turner, B. S., & Samson, C. (1995). *Medical power and social knowledge*. Sage Publications.

Turner, S. P. (2018). *Cognitive science and the social: A primer*. Routledge.

Waitzkin, H. (1991). *The politics of medical encounters: How patients and doctors deal with social problems*. Yale University Press.

Wulff, A., Haarbrandt, B., Tute, E., Marschollek, M., Beerbaum, P., & Jack, T. (2018). An interoperable clinical decision-support system for early detection of SIRS in pediatric intensive care using openEHR. *Artificial Intelligence in Medicine*, *89*, 10–23. https://doi.org/10.1016/j.artmed.2018.04.012

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. In *Paper presented at the proceedings of the 2020 CHI conference on human factors in computing systems*.

Yang, Q., Steinfeld, A., & Zimmerman, J. (2019). Unremarkable AI: Fitting intelligent decision support into critical, clinical decision-making processes. In *Paper presented at the proceedings of the 2019 CHI conference on human factors in computing systems, Glasgow, Scotland, UK*.

Zhang, Z., Genc, Y., Wang, D., Ahsen, M. E., & Fan, X. (2021). Effect of AI explanations on human perceptions of patient-facing AI-powered healthcare systems. *Journal of Medical Systems*, *45*(6), 64. https://doi.org/10.1007/s10916-021-01743-6

AUTHOR BIOGRAPHY

Michael W. Raphael is in the Sociology Department at the Graduate Center of the City University of New York specializing in cognitive sociology, with an emphasis on the politics of situational rationality and its relationship to human and machine problem-solving. By focusing on these politics, he explores domains such as photography, law, medicine, education, and artificial intelligence to examine how communication in the development of competence and expertise impacts theories of justice that guide both academic and public discourse. He teaches courses in sociology, criminology, and anthropology.

How to cite this article: Raphael, M. W. (2022). Artificial intelligence and the situational rationality of diagnosis: Human problem-solving and the artifacts of health and medicine. *Sociology Compass*, *16*(11), e13047. https://doi.org/10.1111/soc4.13047