

Mastery Without a Master: Clinical Deskillling and the Post-Hegelian Paradox of Artificial Intelligence

andrea mangiameli

mangiaferla@gmail.com

Humanitas Gavazzeni

Research Article

Keywords: artificial intelligence, clinical judgment, deskillling, Hegel, professional agency, meaningful human control

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10792015/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Automation can produce a form of dependent mastery in which delegated activity erodes the know-how needed for later takeover. Contemporary artificial intelligence extends this problem into activities through which clinical judgment is formed, including case framing, synthesis, generation of alternatives, and reason-giving. Repeated use can affect two outcomes at once: the immediate decision and the future condition of the decision-maker. Hegel's account of formative activity, together with contemporary work on skill cultivation and AI-induced deskilling, supports a diachronic possibility: recurrent epistemic delegation can change the clinician who delegates. Erosion of an already acquired role-level clinical capability constitutes abdicated mastery; AI-mediated training raises a related developmental risk when formative activity is reduced before mastery is established. Medicine gives this trajectory distinctive ethical significance because professional authority and patient-facing responsibility may persist as substantive clinical agency narrows. Meaningful human control requires adequate human ability. In AI-mediated medicine, that ability is partly shaped by the history of delegation and by whether practice preserves the capacity to contest, reframe, suspend, or bypass mediated judgment. A system can remain under formal human control while progressively changing the human capacity on which that control depends.

1. INTRODUCTION: THE MASTER'S ABDICATION

Automation can relieve human work and still alter what the human remains able to do once the work has been delegated. This possibility was already visible in industrial automation, where routine performance could be transferred to machines while the practical know-how needed for takeover became less frequently exercised (Bainbridge 1983; Coeckelbergh 2015). Formal mastery can survive a decline in the capability through which mastery is exercised.

Contemporary clinical AI enlarges the field in which this history can unfold. Cognitive automation has long been part of medicine through expert systems, risk models, automated interpretation, and clinical decision support. General-purpose language models add breadth, accessibility, and continuity. A single interface can organize a case, generate a differential diagnosis, compare therapeutic options, synthesize evidence, draft explanations, and present reasons favoring one course. These functions already enter clinical reasoning in experimental settings (Goh et al. 2024; Hager et al. 2024). Deliberative work has become readily available for offloading.

That location changes the philosophical stakes. Clinical reasoning develops through repeated acts of framing, comparison, uncertainty calibration, explanation, and revision. Each act produces an answer and also exercises the person who must later judge a similar case. Repeated use can affect two outcomes at once: the immediate decision and the future condition of the decision-maker. The second accumulates slowly and may remain invisible during high assisted performance. The temporal structure raises a question: what happens when the work delegated to AI is also part of the work through which the clinician remains capable of judging?

Natural-language interaction makes that temporal problem practically important because the same system can accompany several stages of a case. Its use can appear as a sequence of small, sensible substitutions: the first problem list, the first differential, the first synthesis of evidence, and the first explanation of why one option appears preferable. A generated problem list can shape the differential, the differential can organize the evidence search, and the resulting synthesis can structure the explanation offered to the patient. No single handoff need appear decisive. The cumulative effect within one episode is that an early generated representation can become the starting point for later reasoning while the clinician remains actively engaged at every stage.

Across cases, repeated use can move continuously from clerical assistance into substantive framing without a new tool, workflow, or explicit decision to delegate judgment. The cumulative pattern determines how much unassisted reasoning remains in routine practice, how often clinicians originate a frame before evaluating one already supplied, and which cognitive activities continue to receive feedback and responsibility. Capability change is difficult to observe at the moment it begins because each consultation can remain useful, reasonable, and compatible with good immediate care. Months or years of individually sensible use can nevertheless redistribute the practice through which professional capability is maintained.

Professional roles change more slowly than the practices that sustain their capabilities. Credentialing, responsibility, documentation, and patient expectations can continue to identify the physician as the person who judges after daily work has become heavily mediated. The visible architecture of responsibility may remain stable as the cognitive architecture supporting it changes. Abdicated mastery names erosion of an already acquired role-level clinical capability across that history of delegation. A related educational risk arises when the same redistribution of formative activity occurs before mastery has securely formed.

Medicine gives this trajectory ethical force because physicians and clinical teams continue to explain, endorse, contest, and answer for decisions. Meaningful human control already presupposes adequate human ability. Formal authority can persist while the practice that sustains that ability changes. A system can remain under formal human control while progressively changing the human capacity on which that control depends. That possibility rests on a formative premise: professional activity can shape the person who performs it.

2. FROM WORK TO JUDGMENT: WHY DELEGATION CHANGES THE DELEGATOR

Hegel's account of lordship and bondage provides a relational grammar for the formative problem. In the *Phenomenology of Spirit*, the master's satisfaction is mediated through the servant's activity and through the object shaped by that activity (Hegel 2018, §§ 189–192). The relation already places activity between the master and the world he seeks to command.

Hegel's account of the servant's labor adds the temporal dimension. Work participates in the formation of the acting subject by shaping the servant's relation to the object and to himself (Hegel 2018, §§ 195–196). Contemporary accounts of skilled agency and medical education operate in different registers and preserve a structurally related insight: agential control, habituated exercise, and cognitive struggle contribute to the development and maintenance of expertise (Ferdman 2026; Keren et al. 2026). Capability has a history of practice. The formative relation between activity and subject is philosophical, while the maintenance of professional capability depends concretely on repeated exercise within environments that afford practice.

Delegation acquires a formative dimension once capability is understood in this way. Moving an activity from clinician to system redistributes opportunities for exercise, and the developmental significance varies across activities. Medicine can relinquish transcription, arithmetic, repetitive measurement, and other historical tasks without compromising clinical agency; such displacement can release attention for more demanding work. Case formulation, generation of alternatives, calibration of uncertainty, recognition of omitted information, and justification of a patient-specific recommendation are more closely tied to the exercise through which judgment is renewed. Later judgment depends on having practiced these activities under uncertainty, feedback, and responsibility. Their value lies both in what they produce and in what repeated performance allows the clinician to remain able to do. Recurrent offloading of these activities changes the pattern of practice available to later judgment without implying that every historical task deserves preservation.

The formative character of activity also changes what mastery can mean when access to practical activity is increasingly mediated. The internal problem of mastery reaches into its means-ends structure, where mastery itself becomes an end pursued through a being treated as a means (Cunniff 2026). Locating the contradiction in this means-ends structure avoids making unacknowledged dependence the core of Hegel's own critique. An instrumental relation can then be examined for the way it reorganizes the practical activity through which the user pursues and specifies ends.

Clinical AI makes that question concrete because tools enter an already organized practice in which attention, time, evidence, and patient priorities compete for salience. A clinician may begin from a general end such as helping this patient and use a system to summarize the case, define salient considerations, generate alternatives, compare evidence, or articulate candidate reasons. Medicine contains several legitimate ends at once - survival, symptom relief, function, treatment burden, uncertainty, and patient commitments - and clinical judgment organizes their weight in a particular case. AI mediation can shape the practical field within which those ends acquire case-specific form.

Technological mediation can also become a source of dependence when routine work is transferred away from the person later expected to resume it. Automation has long exposed operators to abnormal conditions after reducing the routine engagement through which takeover skill was maintained, and technological mastery can become vulnerable when distance from work erodes the practical know-how needed to reclaim it (Bainbridge 1983; Coeckelbergh 2015). General-purpose AI broadens the set of

activities exposed to that temporal mechanism by entering diagnostic and decision processes that belong to professional deliberation (Goh et al. 2024; Hager et al. 2024).

Everyday accessibility makes repeated offloading easy to normalize. A narrow decision aid usually appears at a recognizable point in a workflow; a general-purpose system can move across reading, evidence synthesis, differential diagnosis, multidisciplinary preparation, documentation, and patient communication within one episode. The transition from clerical support to substantive reasoning can occur within the same conversation, making the distribution of cognitive work less visible. An early generated representation can become the starting point for several later activities, leaving subsequent reasoning active and sophisticated while increasingly focused on refining a supplied frame. Repeated across cases, this reduces opportunities to originate alternatives independently and allows the same framing to travel through team discussion and patient communication. The pattern of practice allocated across time determines which forms of cognitive work continue to receive independent exercise.

The analogy remains bounded by machine subjectivity. Hegel's relation involves self-conscious agents and recognition, whereas current AI remains below the threshold required for a recognitive artificial subject (Plevrakis 2024). The post-Hegelian relation remains a user-tool relation whose movement occurs on the human side as altered patterns of practice reshape role-level clinical capability. Deliberative activity is both formative and readily delegable; repeated offloading can reshape the clinician who still occupies the formal role of judge.

3. WHEN DELEGATION CHANGES THE DELEGATOR

A capability that depends on exercise becomes vulnerable to the organization of exercise. Skilled action develops through habituation in environments that continue to offer opportunities for independent performance, making workflow part of the conditions through which capability is maintained (Ferdman 2026). When a system repeatedly supplies a formulation, interpretation, or recommendation before the clinician performs the corresponding work, fewer occasions remain for independent practice.

Capability change follows two temporal forms. Previously acquired competence can degrade when practice diminishes, and opportunities during formation can become too sparse for competence to stabilize (Natali et al. 2025). Established clinicians and trainees occupy different points in this capability history: one can lose fluency in an activity previously performed independently; the other can progress without enough unassisted experience for the same capacity to stabilize.

During training, early and repeated reliance on AI may allow correct outputs to be reached before the cognitive foundations for independent reasoning are securely established, a prospective risk termed never-skilling (Ke et al. 2026). Direct longitudinal evidence in medical training remains limited. Cognitive struggle itself contributes to the formation of clinical expertise (Keren et al. 2026). Foreclosed mastery names the corresponding situation in which an AI-mediated educational environment provides too little formative activity for a role-relevant capability to stabilize.

The educational pathway can interact with erosion across generations. Clinical expertise is transmitted through demonstration, correction, graded responsibility, exposure to tacit reasoning, and progressively independent performance. When experienced clinicians routinely begin from generated summaries or option sets, trainees encounter fewer examples of how a vague presentation becomes a structured clinical problem, even if supervisors remain knowledgeable. Tacit cues can become less visible when early formulation is externalized, and fewer occasions remain for supervisors to show how uncertainty is narrowed before an answer is available. One generation may abdicate a capacity that the next generation is never given the conditions to acquire. Curricula, supervision, time allocation, and local norms become part of the same history because they determine which forms of reasoning remain observable, practiced, corrected, and eventually entrusted to the learner.

Reliance can become self-reinforcing. Cognitive automation can create a loop in which reliance reduces mindful engagement and competence maintenance; declining competence increases the relative utility of automation and encourages further reliance (Rinta-Kahila et al. 2023). When the delegated activity contributes to professional judgment, the loop reaches the capacities that govern future reliance itself. Reduced independent framing makes generated framing more attractive, and greater reliance leaves fewer occasions for independent formulation.

Organizational incentives can stabilize that loop. Independent formulation often takes longer than reviewing a pre-structured output, and clinical environments reward speed, throughput, documentation completeness, and standardization. Metrics that capture assisted performance and leave retained capability unmeasured make immediate advantages visible and long-term human effects difficult to detect. Capacity maintenance can depend on protected opportunities for practice, feedback, and responsibility even when individual clinicians value expertise.

Conversational AI can reinforce the same pattern at the level of individual workflow. Consulting a system can shift part of the cognitive sequence from originating a frame to evaluating one already supplied. Those activities may place different demands on clinical reasoning, and a workflow dominated by evaluation leaves fewer occasions for generative formulation from the outset. Evaluation remains cognitively demanding, yet a workflow dominated by supplied frames changes the distribution of practice available for originating them.

Successful assisted care provides abundant information about the performance of the combined human-AI arrangement and relatively little about what the clinician could do if that arrangement became unreliable. Confidence can remain well calibrated to routine assisted performance while rare situations requiring independent reconstruction remain largely untested. Retained capability is a distinct outcome. Lower unassisted adenoma detection after implementation of computer-aided polyp detection has been observed in a multicenter study, with the observational design supporting association rather than causation (Budzyń et al. 2025). Evidence across other medical domains remains small and heterogeneous (Heudel et al. 2026). Assisted performance and capability retained after exposure are

empirically separable questions, and routine performance metrics can miss the human resource most needed when the ordinary mediated pathway fails.

Supervisory competence is especially difficult to infer from average assisted accuracy. A clinician can perform below a system on a narrow benchmark and still retain the role-level capability to recognize that the benchmark is irrelevant, that the case lies outside the system's competence, that a recommendation conflicts with information absent from the input, or that a different problem should have been posed. These capacities are not redundant copies of the delegated task. They allow the problem to be reopened under discordance, novelty, incomplete representation, or system absence. Their preservation distinguishes useful dependence from loss of substantive clinical agency.

As the target moves from procedural performance to higher-order judgment, the evidentiary boundary becomes sharper. Language models can already enter diagnostic and clinical decision processes, but the long-term effects of that participation on human capability remain unresolved (Goh et al. 2024; Hager et al. 2024). Current evidence combines observed changes in unassisted performance in some automated domains with a plausible mechanism linking reduced exercise to capability erosion and a technological environment that makes deliberative offloading widely accessible. Direct longitudinal evidence for erosion of higher-order clinical reasoning after sustained use of general-purpose generative AI remains sparse. The strongest longitudinal test is still ahead. Medicine's longstanding dependence on distributed expertise also makes the preservation target role-relative.

Modern medicine is epistemically distributed, and substantive clinical agency remains compatible with that dependence. Complex knowledge practices rely on testimony and expertise that no individual can fully reconstruct (Hardwig 1985). Clinical practice depends on laboratories, imaging specialists, pathologists, multidisciplinary teams, guidelines, and evidence distributed across institutions. Role-level clinical capability lies in governing these contributions: understanding their relevance, recognizing grounds for doubt, seeking another source, and deciding how a contribution enters the overall judgment. Recurrent dependence becomes ethically significant when it erodes the capacities required for that governance. The governing capacity becomes most visible when sources disagree, when a contribution falls outside its domain, or when the patient introduces information that changes the significance of otherwise reliable evidence.

Role-level clinical capability can erode even when the professional role and object of judgment remain comparatively stable; this subject-side mechanism complements accounts in which deskilling follows transformation of the object of clinical judgment itself (Germenis 2026). The distinction matters because the clinician may continue to face recognizably similar professional demands while the practical resources available for meeting them have changed. A clinician who delegates a task today may face a similar task years later with a different field of practical possibilities. If reasoning can be delegated and capacities depend on their exercise, the performance of judgment and the preservation of the capability to judge become two different problems.

4. THE MASTER'S ABDICATION

Abdication describes progressive erosion of a previously acquired role-level clinical capability through a history of recurrent delegation, until formal authority increasingly exceeds the substantive capability through which that authority is exercised. It develops by degrees and can remain domain-specific. A physician may retain deep disease knowledge and lose fluency in generating an unprompted differential, or remain capable of criticizing an AI recommendation and become less able to construct a viable alternative plan. The threshold is set by the professional role and by situations in which independent judgment is expected, especially disagreement, novelty, system failure, and patient-specific complexity.

The language of 'abdicating medical decision making' has already been applied to decision substitution associated with automation complacency (Saadeh et al. 2025). Abdicated mastery concerns a different temporal relation: recurrent delegation progressively alters acquired role-level clinical capability while formal decisional authority may remain with the clinician. Decision substitution can contribute to this trajectory, but the two concepts are not coextensive.

Role-relativity keeps the concept from requiring universal technical self-sufficiency. The capacities expected of a trainee, generalist, subspecialist, and multidisciplinary team differ, as do the demands of routine and high-stakes decisions. The standard is the capability needed to occupy a given professional role when the ordinary mediated pathway fails to fit the case. Abdication can be partial: a clinician can remain highly capable across most of a domain and still lose a specific capacity that becomes decisive under exceptional conditions.

Ordinary delegation remains compatible with clinical agency when role-level governance is preserved. A physician can rely on a laboratory, specialist, guideline, calculator, or AI system and retain enough capability to understand the contribution's relevance, recognize discordance, seek another epistemic route, suspend reliance, and integrate the contribution into a judgment she can defend (Hardwig 1985). Abdication appears when repeated offloading weakens the acquired capacities required for that governance across time.

Team organization can mitigate or amplify individual abdication by determining whether complementary human capacities remain accessible when a judgment must be reopened. Distributed expertise can preserve substantive agency even when no single clinician can reproduce every specialist contribution. The risk increases when a team is organized around the same generated synthesis and no member retains a credible route for reframing the case, identifying discordance, or constructing an alternative course. Team structure becomes part of the environment in which individual role-level capability is exercised and sustained.

The corresponding educational trajectory remains conceptually separate. Foreclosed mastery arises before role-level capability has stabilized, when the environment provides too little formative practice for its adequate development. Experienced clinicians and trainees can inhabit the same workflow with

different temporal vulnerabilities, and their interaction gives the problem an intergenerational dimension (Ke et al. 2026; Keren et al. 2026).

Institutional design shapes these trajectories through electronic records, procurement decisions, default interfaces, staffing models, workload pressures, educational policies, and performance metrics. Each choice can be locally reasonable, yet their accumulation determines how cognitive work is distributed across clinicians, teams, and systems. Assisted performance can preserve visible success while the practices sustaining independent capability become less frequent. Because the capabilities at risk are tested mainly by unusual cases, disagreement, or system failure, routine outcomes may provide little reason to question confidence, credentialing, or formal scope. Professional self-assessment can lag behind actual change, and the first visible sign may appear only after the margin for independent action has narrowed. Abdication can coexist with sustained clinical success during the interval in which role-level capability and formal authority begin to diverge.

Abdication names the resulting relation between persisting authority and diminished capability; its causal history may be distributed across procurement, education, supervision, staffing, workflow, credentialing, and quality measurement and may unfold without a single intentional act of surrender. A clinician may reasonably use a high-performing system because time is scarce, because the institution expects its use, or because assistance releases attention for other patients. Institutional choices simultaneously determine which activities remain available for independent practice and which forms of judgment are still demonstrated to trainees. Ordinary professional rationality at each step can coexist with a pattern of delegation that progressively narrows independent practice. The resulting capability loss need not be attributable to negligence by any single clinician.

The post-Hegelian inversion lies entirely on the human side of the relation. Current AI remains an instrument and lacks the self-consciousness required for Hegelian mastery. The servant need not become conscious if the master becomes capable of abdicating. Formal authority can remain fixed as the capacities supporting that authority are exercised less often and progressively weakened. Clinical action then depends more heavily on mediation than the formally authoritative human has fewer resources to contest or replace. Moral subjecthood remains on the human side, and abdication directs attention to the human capability that professional authority still presupposes.

Medicine gives this gap a distinctive ethical shape because the visible locus of judgment often remains human throughout the process. Physicians and clinical teams continue to explain recommendations, respond to patients, integrate competing reasons, and answer for decisions. The patient encounters a clinician whose authority carries expectations of interpretation and judgment even when substantial upstream reasoning has been mediated. Continuing clinical authority requires a substantive human act of judgment. Its content determines which capacities abdication places at risk.

5. THE RESPONSIBLE CLINICIAN

Clinical authority remains epistemically and relationally demanding even within distributed responsibility. Responsibility for AI-mediated care can extend across clinicians, teams, institutions, developers, manufacturers, and regulators. At the clinical interface, physicians still commonly explain recommendations, relate them to a patient's circumstances, decide whether to act, and answer professionally for the resulting course. Oversight can become performative when dependence makes independent scrutiny increasingly difficult (Natali et al. 2025). The human checkpoint can remain visible after the epistemic work expected from it has weakened.

That epistemic work begins with the status of the recommendation itself. Statistical success can support a recommendation, and clinical authority also draws on understanding, context, and the capacity to relate evidence to a particular patient. Medical AI enters an epistemic practice in which authority depends on forms of understanding and human expertise as well as output quality (Kerasidou and Kerasidou 2026; Ferrario et al. 2024). An AI output must still enter a practice of appraisal and justification before it can function as a clinical judgment.

Independent confirmation gives that task an operational form. A clinician using AI as a co-pilot should retain enough professional knowledge and experience to assess its conclusion against an independent clinical standpoint (Chan 2023). Independent confirmation is itself an exercised capacity, and its availability in a difficult future case depends partly on the practice through which it has been maintained during easier ones. Confirmation alone still leaves open whether the clinician can make the recommendation her own.

Endorsement deepens the requirement. Clinical judgment requires the clinician to stand behind a recommendation as a reason for action and to retain the capacities to revise, refuse, and defend it (Matsui 2026). Recurrent AI-mediated deliberation can alter the capacity needed for that endorsement. A safeguard framed as 'the physician must be able to judge' depends on a developmental and practice history capable of preserving the physician who can satisfy it.

The distinction between defending and relaying makes this trajectory visible at the bedside. A clinician who can defend a judgment can articulate enough of the clinical reasons she endorses to respond to challenge, recognize missing considerations, revise the plan, and explain why the recommendation fits this patient. A clinician who can only relay an AI-mediated rationale may still communicate fluently and document appropriately. The difference becomes clinically significant when a patient introduces a reason the system never represented, a case falls outside the expected pattern, or disagreement requires a fresh account of what should be done.

Reason-giving at the bedside makes the distinction between defending and relaying more than an epistemic ideal. Patients can ask why a recommendation fits their situation, what alternatives were considered, how uncertainty was weighed, and what would change the plan. New preferences, treatment burdens, fears, family circumstances, competing goals, or information omitted from the initial representation can alter the practical problem after a recommendation has already been generated. A defensible judgment gives the clinician resources to answer these questions, revise the frame, and

explain why the revised course fits this patient. A fluent generated rationale can preserve the language of explanation even after independent access to the reasons needed for such revision has narrowed. Responsible explanation remains responsive to reasons that emerge during the encounter.

Patient particularity makes this capability especially important. Clinical decisions combine survival, symptom burden, functional goals, treatment burden, uncertainty, family context, and values that can reorganize the significance of population-level evidence. Meaningful human control in medical AI includes responsiveness to patients' epistemic contributions (Pozzi and Santoni de Sio 2026). Reopening the system's frame allows reasons introduced during the encounter to alter the clinical problem rather than merely be appended to an existing recommendation.

Such cases require more than error detection. A missing history, an unrepresented patient value, a new competing risk, or a shift in goals can change what the evidence means for this case. Clinical judgment includes the ability to reformulate the problem and generate a new option set when the existing frame has become inadequate. Recurrent dependence on AI-generated framing can narrow the occasions on which this generative work is practiced. The same capacity that allows a clinician to contest a recommendation also allows her to see that a different problem should have been posed.

Clinical phronesis gives medical content to this form of judgment by integrating experience-grounded expertise, attentiveness to the patient, embodied perception, professional character, and reasoning within complex situations (Boudreau et al. 2024). It concerns knowing what a situation calls for when multiple forms of evidence, perception, experience, and patient particularity must be integrated. Role-level clinical capability depends on an integrated set of practiced capacities: independent problem framing, recognition of omitted information, generation and appraisal of alternatives, detection of discordance, access to alternative epistemic resources, and integration of patient reasons. A clinician may retain one capacity and lose fluency in another, yet professional judgment depends on their organization into a conclusion that can be revised and defended. Substantial erosion of this integrated capability would weaken the clinician's ability to occupy a role defined by judgment, accountability, and responsiveness to the particular patient even when many replaceable task skills have been delegated successfully.

Explainability gains its value from these standing capacities. A clinician may receive a transparent account of how an AI recommendation was produced and still lack the practiced ability to evaluate whether the framing, alternatives, and weighting of reasons fit the present patient. Explanation supports oversight when it enters a capable practice of judgment; its protective value declines as the recipient has fewer opportunities to exercise that practice independently. Formal participation can remain intact as the capability to constitute, contest, and defend the decision narrows.

These role-level capabilities give clinical authority its substantive content. They allow physicians and teams to govern distributed epistemic dependence, respond to patient reasons, reopen a decision when the system's representation becomes inadequate, and construct an alternative whose rationale they can defend. The ability to depart from a mediated path depends on more than formal permission; it depends

on a practiced capability to recognize why departure is warranted and what should replace the supplied course. Meaningful human control is substantive only where the people assigned to exercise it retain the ability required by the role.

6. WHEN HUMAN CONTROL DEPENDS ON ITS OWN HISTORY

Meaningful human control links authority to human capability by requiring system behavior to track relevant human reasons and by assigning responsibility to agents with actual capacity to exercise it (Santoni de Sio and van den Hoven 2018; Cavalcante Siebert et al. 2023). In medicine, the requirement includes the clinician's capacity to decide with decision support, without it, and against it when circumstances require (Kara 2025). Human control rests on an ability that must remain present in the people assigned to exercise it.

That ability has a practice history. Habituation, opportunities for exercise, and the distribution of cognitive work shape the role-level clinical capability available at later moments. A clinician may satisfy the ability condition early in adoption and reach a later stage with fewer independent resources because recurrent reliance has changed what she routinely practices. The ability required for meaningful human control is partly shaped by the history of delegation in AI-mediated medicine. The capability available at a later decision is partly the residue of which cognitive activities were repeatedly performed, supported, or bypassed in earlier ones.

Interfaces, permissions, and responsibility assignments can remain stable while the human resources underlying them evolve. A system may still display an override button, require a clinician's signature, record the same accountable role, and present the same formal route for departure after years of use. These features establish architectural opportunities for control; they do not show that the human participant still possesses the role-level clinical capability needed to recognize grounds for departure and construct a viable alternative. A workflow can look formally unchanged across years of use even as the capability needed to make its permissions effective changes. Enduring substantive control depends on whether previous practice has preserved enough role-level capability for the clinician or team to recognize when mediation no longer fits and to act competently outside it.

History-dependence changes the meaning of reversibility. A workflow can permit override, suspension, or exit and still leave the human participant progressively less able to recognize when those actions are needed or to execute them competently. Reversibility has an architectural layer and a capability layer: the arrangement must allow a different course, and the human participant must retain sufficient role-level capability to use that possibility. Formal permission to depart from an AI recommendation can survive after practical reversibility has weakened.

Subject-preserving accounts already place non-delegability, contestability, and reversibility among the conditions of substantive human judgment and recognize institutional responsibilities for forming agents capable of judgment (Torrecilla-Pinero 2026). Repeated delegation adds a temporal vulnerability

to these conditions because it can alter the later capability required to exercise them. Reversibility itself can decay as the capability required to exercise it erodes.

Treating capability as a maintained condition extends responsibility for control into training sequences, workflow design, and the distribution of unassisted practice before a difficult case ever occurs. The sociotechnical arrangement must continue to produce clinicians and teams capable of exercising the human authority formally assigned to them when reliance becomes unsafe or clinically inadequate.

Team organization can preserve practical reversibility when at least one credible route remains for challenging the generated frame and constructing another. Complementary expertise can compensate for individual limits when the team preserves genuinely distinct epistemic routes. A team organized around a single AI-mediated synthesis can lose that resource even when many humans remain formally involved, because multiple participants may inherit the same initial representation. Health-specific meaningful human control includes conditions for enabling, exercising, and evaluating control; capability maintenance belongs within those enabling conditions and becomes visible when control is assessed after prolonged use (Hille et al. 2026).

Selective capability maintenance can include protected unassisted reasoning during training, structured opportunities for independent disagreement or alternative formulation, and periodic assessment under system absence or discordance. These practices need not preserve every technical skill that technology can safely replace. Role-level clinical capability requires enough integrated capacity to reopen a case frame, incorporate patient reasons, identify grounds for departure, and construct a defensible alternative when mediation no longer fits the situation.

Successful deployment should be interpreted across the same temporal horizon. High assisted performance shows what the combined human-AI system can achieve under ordinary conditions. Retained capability shows whether substantive human control remains available when the ordinary conditions of reliance no longer hold. The two outcomes can diverge as present assistance narrows the independent practice that future oversight will require. Ethical evaluation must follow both the decision produced today and the human capability being produced for tomorrow.

7. CONCLUSION

AI-mediated medicine redistributes cognitive work, and cognitive work participates in the formation and maintenance of clinical capability. Repeated delegation can change the clinician who delegates even when assisted performance remains high and formal professional authority remains intact. For clinicians who already possess the relevant capabilities, that trajectory can become abdicated mastery as role-level capability narrows across a history of offloading. The same redistribution can also alter the formative environment in which future clinicians learn to judge.

Medicine preserves a human locus of judgment across these trajectories. Physicians and clinical teams continue to explain, endorse, contest, and answer for decisions, and patients continue to introduce

reasons that require interpretation within a particular life. Meaningful human control depends on preserving the role-level clinical capability that gives those responsibilities substance. Technical authority to override an AI system has practical force only where clinicians or teams retain enough capability to recognize when departure is warranted and to construct a defensible alternative.

A system can remain under formal human control while progressively changing the human capacity on which that control depends. That possibility is the central post-Hegelian paradox of AI-mediated medicine. The future quality of clinical judgment will be shaped partly by how today's practice distributes the cognitive work through which clinicians learn, maintain, and exercise the capacity to judge.

Declarations

Author Contribution

A.M. wrote the main manuscript text

References

1. Bainbridge, L. 1983. Ironies of automation. *Automatica* 19(6):775–779. [https://doi.org/10.1016/0005-1098\(83\)90046-8](https://doi.org/10.1016/0005-1098(83)90046-8)
2. Boudreau, D., H. Wykretowicz, E. A. Kinsella, A. Fuks, and M. Saraga. 2024. Discovering clinical phronesis. *Medicine Health Care and Philosophy* 27:165–179. <https://doi.org/10.1007/s11019-024-10198-8>
3. Budzyń, K., M. Romańczyk, D. Kitala, P. Kołodziej, M. Bugajski, H. O. Adami, J. Blom, M. Buszkiewicz, N. Halvorsen, C. Hassan, T. Romańczyk, Ø. Holme, K. Jarus, S. Fielding, M. Kunar, M. Pellisé, N. Pilonis, M. F. Kamiński, M. Kalager, M. Bretthauer, and Y. Mori. 2025. Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study. *The Lancet Gastroenterology & Hepatology* 10(10):896–903. [https://doi.org/10.1016/S2468-1253\(25\)00133-5](https://doi.org/10.1016/S2468-1253(25)00133-5)
4. Cavalcante Siebert, L., M. L. Lupetti, E. Aizenberg, N. Beckers, A. Zgonnikov, H. Veluwenkamp, D. Abbink, E. Giaccardi, G.-J. Houben, C. M. van den Jonker, D. Forster, and R. L. Lagendijk. 2023. Meaningful human control: actionable properties for AI system development. *AI and Ethics* 3:241–255. <https://doi.org/10.1007/s43681-022-00167-3>
5. Chan, B. 2023. Black-box assisted medical decisions: AI power vs. ethical physician care. *Medicine Health Care and Philosophy* 26:285–292. <https://doi.org/10.1007/s11019-023-10153-z>
6. Coeckelbergh, M. 2015. The tragedy of the master: automation, vulnerability, and distance. *Ethics and Information Technology* 17(3):219–229. <https://doi.org/10.1007/s10676-015-9377-6>
7. Cunliff, S. 2026. The master's problem: revisiting Hegel's critique of social domination. *European Journal of Philosophy* 1–20. Advance online publication. <https://doi.org/10.1111/ejop.70089>
8. Ferdman, A. 2026. AI deskilling is a structural problem. *AI & Society* 41:3001–3013. <https://doi.org/10.1007/s00146-025-02686-z>

9. Ferrario, A., A. Facchini, and A. Termine. 2024. Experts or authorities? The strange case of the presumed epistemic superiority of artificial intelligence systems. *Minds and Machines* 34. <https://doi.org/10.1007/s11023-024-09681-1>. Article 30.
10. Germenis, A. E. 2026. Ontological deskilling: clinical expertise and the changing disease object in precision medicine. *Medicine, Health Care and Philosophy*. Advance online publication. <https://doi.org/10.1007/s11019-026-10377-9>
11. Goh, E., R. Gallo, J. Hom, E. Strong, Y. Weng, H. Kerman, J. A. Cool, Z. Kanjee, A. S. Parsons, N. Ahuja, E. Horvitz, D. Yang, A. Milstein, A. P. J. Olson, A. Rodman, and J. H. Chen. 2024. Large language model influence on diagnostic reasoning: a randomized clinical trial. *JAMA Network Open* 7(10):e2440969. <https://doi.org/10.1001/jamanetworkopen.2024.40969>
12. Hager, P., F. Jungmann, R. Holland, K. Bhagat, I. Hubrecht, M. Knauer, J. Vielhauer, M. Makowski, R. Braren, G. Kaissis, and D. Rueckert. 2024. Evaluation and mitigation of the limitations of large language models in clinical decision-making. *Nature Medicine* 30:2613–2622. <https://doi.org/10.1038/s41591-024-03097-1>
13. Hardwig, J. 1985. Epistemic dependence. *The Journal of Philosophy* 82(7):335–349. <https://doi.org/10.2307/2026523>
14. Hegel, G. W. F. 2018. *The Phenomenology of Spirit*, ed. T. Pinkard. Trans.). Cambridge University Press. (Original work published 1807).
15. Heudel, P. E., H. Crochet, Q. Filori, T. Bachelot, and J. Y. Blay. 2026. Artificial intelligence in medicine: a scoping review of the risk of deskilling and loss of expertise among physicians. *ESMO Real World Data and Digital Oncology* 12:100693. <https://doi.org/10.1016/j.esmorw.2026.100693>
16. Hille, E. M., P. Hummel, and M. Braun. 2026. Meaningful human control over AI for health? A review. *Journal of Medical Ethics* 52(e1):e50–e58. <https://doi.org/10.1136/jme-2023-109095>
17. Kara, M. A. 2025. Clouds on the horizon: clinical decision support systems, the control problem, and physician-patient dialogue. *Medicine Health Care and Philosophy* 28:125–137. <https://doi.org/10.1007/s11019-024-10241-8>
18. Ke, Y., L. Jin, J. C. L. Ong, A. J. Thirunavukarasu, J. Car, C. Y. Cheung, Y. C. Tham, D. S. W. Ting, M. E. H. Ong, S. Compton, A. Narayan, P. A. Keane, T. Y. Wong, D. W. Bates, P. Tan, and N. Liu. 2026. AI-induced never-skilling in medical education. *Nature Medicine* 32:1997–2006. <https://doi.org/10.1038/s41591-026-04438-y>
19. Keren, R., B. R. Desai, and D. C. West. 2026. Promoting clinical expertise in the age of AI: no struggle, no mastery. *Journal Of The American Medical Association* 335(21):1837–1838. <https://doi.org/10.1001/jama.2026.6097>
20. Kerasidou, A., and C. X. Kerasidou. 2026. Epistemic authority and medical AI: epistemological differences and challenges in medical practice. *Medicine Health Care and Philosophy* 29:89–95. <https://doi.org/10.1007/s11019-025-10306-2>
21. Matsui, M. 2026. Is there a judgment here? Why medicine cannot endure decision-making without a judging subject in the age of AI. *Medicine, Health Care and Philosophy*. Advance online publication.

<https://doi.org/10.1007/s11019-026-10346-2>

22. Natali, C., L. Marconi, L. D. Dias Duran, and F. Cabitza. 2025. AI-induced deskilling in medicine: a mixed-method review and research agenda for healthcare and beyond. *Artificial Intelligence Review* 58. <https://doi.org/10.1007/s10462-025-11352-1>. Article 356.
23. Plevrakis, E. 2024. Can AI be a subject like us? A Hegelian speculative-philosophical approach. *Discover Computing* 27. <https://doi.org/10.1007/s10791-024-09472-3>. Article 46.
24. Pozzi, G., and F. Santoni de Sio. 2026. Epistemic justice as a condition for meaningful human control over medical AI. *Minds and Machines* 36. <https://doi.org/10.1007/s11023-026-09762-3>. Article 10.
25. Rinta-Kahila, T., E. Penttinen, A. Salovaara, W. Soliman, and J. Ruissalo. 2023. The vicious circles of skill erosion: a case study of cognitive automation. *Journal of the Association for Information Systems* 24(5):1378–1412. <https://doi.org/10.17705/1jais.00829>
26. Saadeh, M. I., J. Janhonen, E. Beer, C. Castelyn, and D. N. Hoffman. 2025. Automation complacency: risks of abdicating medical decision making. *AI and Ethics* 5:5783–5793. <https://doi.org/10.1007/s43681-025-00825-2>
27. de Santoni, F., and J. van den Hoven. 2018. Meaningful human control over autonomous systems: a philosophical account. *Frontiers in Robotics and AI* 5. Article 15. <https://doi.org/10.3389/frobt.2018.00015>
28. Torrecilla-Pinero, J. A. 2026. Judgment cannot be delegated: a subject-preserving framework for AI governance. *Philosophy & Technology* 39. <https://doi.org/10.1007/s13347-026-01140-2>. Article 135.