

2024-12-01

The cultural search: AI, sustainability, and the human touch

J. Floyd. 2024. "The Cultural Search: AI, Sustainability, and the Human Touch" Journal of Artificial Intelligence for Sustainable Development, Volume 1, Issue 1. <https://doi.org/10.69828/4d4kkb>
<https://hdl.handle.net/2144/50261>

"Downloaded from OpenBU. Boston University's institutional repository."



POSITION PAPER

The Cultural Search: AI, Sustainability, and the Human Touch

Juliet Floyd^{1,*}

¹Boston University, USA

*jfloyd@bu.edu

Abstract

The increasing growth in diversity and complexity of forms of life, accelerated by new media and the development of AI, raises fundamental questions about how best to sustain educational systems that serve democratic forms of life. Alan Turing predicted the need for a “cultural search” “involving the human community as a whole”, something not reducible to the evolution of algorithmic or biological processes alone. On his view education involves the development over the human life course of both “discipline” and “initiative” and the inheritance of “culture”. It is these, including human “phraseology”, that he felt would most fatefully shape, sustain, transform, and translate intergenerational, intersocial, and intercultural values and symbolic forms over time. On Turing’s view, “intelligence” is an “emotional” concept rooted in “common sense”, consisting in the appreciation of the manifold differences among different types of search. It is primarily constituted neither by the ability to mimic human activities, nor to optimize rational choices or achieve consensus about the outcomes of only one kind of search. This paper argues that Turing’s idea of “the cultural search” 1) illuminates the global, specifically human educational challenges to democratic forms of life in a world of AI, including the limits of algorithms to master ethical and social problems of the public and 2) draws out the importance of what Rawls, analyzing the philosophical foundations of democratic liberalism, called *the sense of justice* and *the problem of democratic stability*. Both are centrally involved in the possibility of sustaining democratic forms of life over time. Given the variety of diverse “comprehensive doctrines” by which people live, as well as the differences among particular situations of democratic forms of government around the world, education must foster the development in citizens of a “sense of justice” sufficient to sustain an “overlapping consensus” on the positive good of principles of tolerance and difference over time. The result is 3) that AI design for democratic education should constantly attend to the centrality and the plasticity, connectivity and resilience of human-to-human interactions in the presence of machines to meet the challenge of an evolving matrix of democratic values, emotions, aspirations, and human capacities that alone make it possible for democratic societies to be sustained and developed.

Keywords: Cultural search, Democratic forms of life, Sense of justice, Limits of algorithms, Human-to-human interactions

The Turing Test as Education

The Turing Test or “Imitation Game”, as Turing called it (Turing; 1950), was designed to draw us into, explore, and refine our sense of whether machines can “think” or not. The Test retains its fascination in popular culture three quarters of a century later, despite the many ways in which computers have substantially developed since that time. The Test is by now a *meme*, a socially significant symbolic myth, revealing the spirit of our age and forming a matrix of feeling shaping the forms of our self-presentations, our sense of mentality and life, as well as a particular horizon of questions we take to be centrally relevant (compare (Langer; 1957), chapter 1 on questions). At the moment there is much discussion of “AGI”, i.e., “artificial general intelligence” – a concept not very well defined, but used as a kind of marker or aspiration for the design of networked computer devices that can “outperform” humans across a wide range of cognitive

(perhaps economically valuable?) tasks. The public and experts are confused about the nature of this marker and the software industry's use of it: there are many fears about machines "replacing" humans with "superintelligence" and actually taking over the world, although in many respects our fears remain emotional: ghostly and unclear. Thus the point of revisiting the Turing Test, which lies at the origin of the development of AI.

As the first to design the stored program computer concept, analyze satisfactorily the notion of algorithm, and work with actual computing machines to solve numerical and cryptological problems (Hodges; 2012), Turing found himself periodically "surprised" by how powerful the machines he worked with proved to be in solving the problems which he devised for them (Turing; 1950), pp 450. Foreseeing the future powerful development of artificially "intelligent" machinery (AI), in his (Turing; 1950) he proposed to investigate the question "Can Machines Think?" by replacing the straightforward demand for definitions of the concepts of "machine" and "think" with a game, now known as the "Turing Test".

In the first version of his game in his (Turing; 1950), pp 433, a man A and a woman B sit in a room walled and screened off from an Interrogator C, who may be man or woman. C poses typewritten questions to A and B, ideally by using a teleprinter. C's aim is to determine which of A and B is a woman, and which a man. A attempts to fool C with his answers, whereas B attempts to help C determine the right answer with hers. Turing conjectures that "probably" the best strategy for (the cooperative woman) B is to always tell the truth, to help C judge accurately. We see here a very particular construal of gender built into the first step of the Turing Test's presentation (see (Sterrett; 2000; Floyd; 2023)).

Turing next shifts this gender-guessing game into another, in which A is replaced with a computer. C then must face a computer and B. The test of the "intelligence" of the machine, now in A's place, is whether it can fool C with its answers into thinking that it is a human, and not a machine, a significant percentage of the time.

Turing suggested that

... in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10⁹, to make then play the imitation game so well that an average interrogator will not have more than 70 percent. chance of making the right identification after five minutes of questioning (Turing; 1950), pp 492.

There is much debate today about whether or not the Turing Test has been passed, and what the character of the response-dependent concept of "intelligence" really comes to. From Searle's *Chinese Room* (Searle; 1980, 1984) to IBM Watson's victory on *Jeopardy!* (in 2011), and the movies *Bladerunner* (1982, 2017) and *Her* (2013), the very idea of a human fooled by, or drawn into an emotional matrix with, a machine has captured the cultural imagination. Now we have deep fakes produced easily and well by generative AI, and the first suicide of a fourteen-year-old who appears to have fallen in love with a ChatBot has been reported (Rose; 2024). Concerted efforts to use large language models (LLMs) to produce poems and creative works that will strike us as works of genuine art, literature and creativity are going on increasingly. In general this obsession with AI and human-machine interface form a kind of "spirit" of our time, which has been called by Katz and Aakhus, who conducted some of the earliest significant sociological studies of the impact of mobile technology on our lives (Katz and Aakhus; 2004), the *Apparatgeist*, the age of the machine.

At its root, Turing's Test is an educational test. It is, furthermore, built to be democratic. Viewed in this way it is less an objective way to measure the relative "intelligence" of machines and human beings in a contest of task-completion, and more an experiment in human response, perception, feeling, and phraseology: the making of sense among human beings (Floyd; 2019, 2023). Turing emphasized that "intelligence" is an "emotional" concept, evolving over time and entangled with the many ways in which we embed our words in forms of life, including those in which computers figure (Turing; 1948; Proudfoot; 2014). Of course, forms of life are not mechanisms: they have biological, algorithmic, ethological, linguistic, cultural, historical, symbolic, and political aspects: all the panoply of organism (Wittgenstein; 2009) called "forms of life" (Floyd; 2020). And Turing emphasizes – as is all too often forgotten – human-to-human interaction in the presence of machines. His Test investigates an emotional matrix of human significance, not merely mechanical or behavioral processes.

Confronting the limits of algorithmic approaches, Turing stressed the importance of what he called "the cultural search" for the future of "democratic community": our ways of devising surveyable representations and concepts to help human beings communicate, navigate, and meaningfully integrate their lives into the coming algorithmic, biological and scientific avalanche of information (Turing; 1948), pp 516. And he predicted that in the future the cultural search would take place at a global level, amongst "the human community as a whole" (Turing; 1948), pp 516. Today, with the world-wide penetration of mobile technology and the use of social media – the technological effects being rapidly reinforced by AI and machine learning – we see his prophecy coming true.

Turing explicitly worried about the human tendency to uniformize and impose one mode of language in handling computers: this is an all-too-human tendency. As he wrote, "no democratic mathematical community would stand for" the idea of a uniform mathematical notation, "nor would it be desirable" (Turing; 1944), pp 245. He worried about the choice of data sets and results being skewed toward English:

The works and customs of mankind do not seem to be very suitable material to which to apply scientific induction. A very large part of space-time must be investigated if reliable results are to be obtained. Otherwise we may (as most English children do) decide that everybody speaks English, and that it is silly to learn French (Turing; 1950), pp 448.

Turing rightly worried repeatedly about our all-too-human biases, and especially at the societal level. He raised the prospect of religious and intellectual authorities imposing arbitrary orthodoxies on our behavior. He himself – not without his prejudices – expressed the erroneous view that Muslims deny women souls (Turing; 1950), pp 443. On the topic of race, he stressed the frailty of our all-too-human reliance on social luxuries ("strawberries and cream") to grease the wheels of social acknowledgement in a world where there is little genuine "friendliness" between humans and machines or between humans of differing races (Turing; 1950), pp 447. The technological plan, he pointed out, is to treat machines as our "slaves" (Turing; 1947), pp 496, gradually absorbing human rote forms of labor into the mechanisms, e.g., "girls' read[ing] values and punch[ing] them into cards", thereby treating human activities as part of the machines' "limbs" (Turing; 1947), pp 495. Women, who formed 75% of the workforce at Bletchley Park during World War II, were cast into the supplementary role of "wrens" (Wikipedia; 2025), or birds; Turing pictures their "slave" labor being replaced by the mechanisms. There

has been increasing interest in the role of “hidden figures” in the history of computing and cryptography (see the film of this title and the TV series *The Bletchley Circle*), and there remain widely expressed concerns that the emotional matrix of our human-machine interactions is and has always been shaped through gender lenses around the globe (Adam; 1998; Trauth and Quesenberry; 2023), and also racialized (Chude-Sokei; 2022). The marginalized have frequently been classified, with their forms of labor, as themselves rote, mechanical, unthinking beings, despite the fact that their labor was often intellectually exhausting and terribly demanding (Daston; 2018).

Turing predicted that AI would disrupt the social power of the intelligentsia, specifically mathematicians, in light of the fact that the Universal Turing machine is capable of incorporating all particular routines into the work of one:

The Masters [the mathematicians] are liable to get replaced because as soon as any technique becomes at all stereotyped it becomes possible to devise a system of instruction tables which will enable the electronic computer to do it for itself. It may happen however that the masters will refuse to do this. They may be unwilling to let their jobs be stolen from them in this way. In that case they would surround the whole of their work with mystery and make excuses, couched in well-chosen gibberish, whenever any dangerous suggestions were made. I think that a reaction of this kind is a very real danger (Turing; 1947), pp 495.

This difficulty remains one for pure mathematicians today. Will their conceptual and proof-work be utterly replaced by machines? Turing thought not: he felt that the development of everyday “types” and phraseology, human modes of surveyability and phrasing, were essential, not only to the progress of knowledge, but to the development of “intelligent” machinery: a point he derived from attending Wittgenstein’s philosophical lectures on the foundations of mathematics (Floyd; 2013; Cooper and van Leeuwen; 2013). We will help make the machines’s rule-following activities “intelligent” and significant through our modes of everyday conversation and phraseology.

Nowadays, however, we see strong and influential tendencies among some AI engineers to replicate and eradicate all human work, which would be fatal, if Turing was right, for the interaction between human insight and science. The crucial importance to our future with AI of democratic conversation, including conversation about forms of justice and care in a world in which AI is increasingly imbricated, must be emphasized as part of the meaning-making process, which is essentially an educational one. This alternative approach – a search – involves ethics being built into each step of design and implementation, constantly revisited in light of human beings’ emotional responses to AI in an open conversation about “emotional” responses. This will require a far more open effort at the accountable implementation of AI throughout the globe, and a far more serious focus on education, searching, dialogue and listening than the framing of short-term marketing jargon and the hidden mechanisms of surveillance and software development have so far delivered. On what could such a cultural search be based? What are the prospects for its reality? Is it too ideal, or does it have a grounding in certain “emotional” human realities?

Rawls on the Sense of Justice

To answer this question of reality let us turn, in the sphere of democratic philosophy, to John Rawls, the theorist par excellence of democratic liberalism rooted in the idea of a “social contract” among citizens (Rawls; 1963, 1999b,a, 2005). Working within “ideal theory”, Rawls’s idea of “justice as fairness” was intended to help members of a democratic society reconcile themselves to principles governing the “basic structure” of their social and political forms of organization – “basic” meaning that we exclude from consideration all detailed and specific mechanisms and methods of governance and control, as well as the actual historical and psychological positions of particular members of society and their instrumental estimation of empirical goods and ends. The focus of the “basic structure” is on a fundamental idealized setup of frameworks and schematic rules orienting *principled* discussion of – and argumentation over – the outcomes generated by specific democratic mechanisms and institutions. This is a structure that should, if a modern industrial democratic society is to abide over time, clearly convey how citizens are to develop, mutually recognize, and communicate a “sense of justice” among themselves in a democratic community. The philosophical aim, conceptually speaking, is to transform the human and moral idea of a *person* into that of a *democratic citizen* by way of the generation of fundamental principles. Such a concept of a citizen is intended to express a constitutive node in our understanding of democracy, functioning as a normative ideal. It does not simply denote a particular legalized national status. Rawls’s idea is to transform a *modus vivendi* – a temporary peace in which each party that cannot dominate the other waits until they have power to do so – into the positive good of a just democratic regime where tolerance is respected and fostered.

Rawls’s theory is quite limited and abstract, though its power and potential for realization comes precisely from this. His project is a matter of distilling principles to test the intellectual coherence of democratic ideals in the way a mathematician distills axioms or basic principles for a concept. Here we may remember that Turing, analogously, distilled principles for the ancient mathematical concepts of “algorithm” or “computation” with his notion of a Turing Machine (Turing; 1936/37). Note that despite the abstractness of these idealized models, concrete realities are presupposed in any such exercise of reconstruction: neither Rawls’s reconstruction of a just democratic society nor Turing’s of the notion of an algorithm is either *a priori* conceptual or societal “engineering”. Nor are these accounts simply descriptive. Rawls speaks of the method of achieving “reflective equilibrium”: we test our untutored intuitions about justice as fairness against his principles, revising the principles if the intuitions indicate the need. Just as Turing’s analysis of “computable” turned on a language-game-like picture of a human in the real world reckoning according to a fixed rule with paper and pencil – thereby assuming that a small snapshot has a sufficient degree of reality to underpin the theory (Floyd; 2017) – Rawls’s presentation and acknowledgment of a democratic society governed by a conception of justice as fairness presupposes that a robust modicum of a “sense of justice” – an emotional and philosophical sensibility, emotional matrix, set of commitments and structure – already exists in the reflecting subjects who discuss his theory.

This methodological assumption is something critics of Rawls have questioned, particularly with respect to the issues of feminism and race. In the United States the legacy of slavery has greatly and continuously affected, not only the setup of institutions and laws from the beginning, but the emotional and cultural sensibilities of citizens with regard to gender and race (Abbey; 2013; Mills; 1997). It is true that in presuming that a sense of justice is sufficiently diffused and shared by citizens throughout a democratic community, Rawls’s philosophical construction of the viability of a particular basic structure of democratic institutions is idealized, partly historical, and rationalized – risking falsification through presumed universalism and neutrality¹. Yet the point of emphasizing the importance of citizens’ sharing a “sense of

¹ One may look at Rawls’s lectures on the history of philosophy (Rawls; 2007), however, to see how he integrates his theory of justice with the history of modern

justice” is also psychological, developmental and emotional, showing an aspiration on Rawls’s part to secure the actual reality, or at least feasibility of democracy as a stable, multi-generational form of government capable of enduring over time (Rawls; 1999a,b; Nurock; 2007, 2008).

Rawls famously fashioned an analytical device, “the original position”, to make vivid the derivation of principles of justice for his theory. In the thought experiment, subjects are asked to go behind a “veil of ignorance”, blanketing their particular religious, cultural and socio-economic situations. They are then to imagine how to devise principles governing a basic structure that would be just for all, no matter what their situations. The idea is that they are to act as representatives tasked with setting up the original basic structure – the basic principles of justice – that are to govern their democratic society in the future in the long term, binding themselves to it (Rawls; 1999a,b). Rawls’s critics have questioned the psychological ability of individuals in particular traditional societies to engage in such “veiled”, disinterested reflection, pointing out that the very idea of a social contract presupposes the full standing of all participants – something historically the United States certainly did not grant. Yet one might also argue that the representatives would have to be aware of, and moved by, the historical realities of social structures deemed unjust – from religious intolerance to oppression, slavery and systematic disenfranchisement of groups – if they are to act as “representatives” at all, for this is a capacity presupposed in the very concept of representative democracy.

The important point for our purposes is that Rawls never construed this procedure as an algorithm to be simply computed. Despite the game-theoretic flavor of the derivation of principles in the “original position”, the analytical thought experiment is not intended to eliminate all aspects of bias or feeling. Rather than aiming at utterly neutral abstract principles for an arbitrary economic agent maximizing utility, the procedure is intended to elicit and generate a structure of principles from the ideal tradition of democratic aspirations handed down through – though not exclusively through – the best in the tradition of Locke, Rousseau, and Kant as received in the history of the United States: an historical reality, however mixed. Our sensibilities and untutored responses to situations of justice and injustice are to be factored into the theory as it is formulated, our epideictic responses in situ taken seriously. Suggesting a comparison with challenges faced by Chomsky’s universal grammar in linguistics (Rawls; 1999a), pp 41, Rawls notes that in the quest to formulate principles to describe the sense of grammaticality in natural languages, Chomsky highlighted the fact that the theoretical constructions “far outrun the ad hoc precepts of our explicit grammatical knowledge”: such an investigation requires, not merely a systematizing theory, but one that is sensitive to particular, situated speakers’ responses, for there will be no fully adequate algorithm the linguist can import into the field.

Rawls’s later work on international law, *The Law of Peoples* (Rawls; 2001) expressed doubt about how universally applicable his democratic theory might be, leaving to specific political cultures their task of elaborating their own principles and theories of justice. It is not clear how far Rawls’s view applies, for example, to a constitutional monarchy such as Great Britain’s or other mixed forms of governance, nor how it would fare in non-European, traditionally based political cultures. Pursuing the analogy with Chomsky’s “universal” grammar, we need to distinguish between “justice as such” and particular, culturally embodied forms of justice – just as we should distinguish language capacity as such from its culturally specific, idiomatic forms.

What is important for us, living in a global world of AI, is that Rawls’s approach to developing a philosophy of justice turns, not on what is strictly speaking “rational” (in the sense of economic rules of self-interest), but of what Rawls calls the *reasonable* (Rawls; 1999a,b). The “reasonable” is something for which there is no algorithm, but only the prospect of ongoing discussion and exchange, touchpoints to be gained in the offering by public citizens and legislators of rulings instantiating principles of justice one might “reasonably” expect others to share and be reciprocally bound by. This will include reflection on education: the raising of children in moral frameworks that allow them to take on the capacity to structure their behavior, sensibilities and ongoing lives in ways compatible with a public, shared sense of democratic justice. “The sense of justice” involves a complex psychological and philosophical matrix of capacities for shame, reflection, self-obligation, and obligation to community (Rawls; 1963), providing the moral capacity for life in a modern democratic society with its plurality of rapidly-evolving forms of life.

There is a long American philosophical tradition emphasizing as foundational for democratic life and stability the cultivation of other elements of education: individuality, aesthetic sensibility, and the overcoming of societal conformity through self-reliance, the resilient and tolerant aspiration to moral perfection, and willingness to dissent (Emerson; 1841; Thoreau; 2015; Dewey; 1916; Cavell; 1988, 1989). A corollary in the present day of streaming media draws the importance of popular culture in to the process of education (Laugier; 2024). Perhaps Rawls underestimated such insistence on the perpetual challenges to democratic institutions of devising systems of education that will inspire in citizens not only knowledge and moral capacity, but individual creativity, empathy, generosity and resilience in the face of a variety of conflicting forms of human life. We see today how generative AI has reframed the question of the value and fertility of human creativity – in the arts, moral development, and life – and infer that this places a large question mark over our sense of the future of democratic education. Before, democratic thinkers conceived originality and non-conformity as something to be developed in order to achieve human stability in democratic structures of justice over time. With the advent of mobile technology and the convenience of sharing on the world wide web – as well as the continuing flood of ever-more-convincing AI generated images, texts and sounds – there are threats to the economic valuation and support of human creativity and development, and threats to our sense of its human originality, as well as newly original forms of disinformation, distraction, and nudging. Even more, an idea that AI algorithms may ideally solve all ethical and political questions, transcending the human, has entered the engineering world (Floyd and Katz; 2023).

There are general fears that human beings will be taken more and more out of the loop of technological creation as further and further activities are automated – oftentimes based on “moral” reasoning such as the Trolley problem or the imperative to minimize pollution in the choice of transportation routes in geolocation programs. In the limit, imagining “AGI” involves imagining that no activities of importance – artistic expression, care, love, companionships, political discourse, discernment and commitment – will remain centrally human ones, as machines will be regarded as “better than human” at what they do.

And yet, as we examine the Turing Test and Rawls’s democratic theory more precisely, at each stage of development the sense of justice and of “intelligence” is based upon the development of the associated human emotions and judgments. Without them no sense of justice can be developed, nor any meaningful sense of “intelligence”. Sociologists are already documenting the losses involved when labor with

machines severs connective ties amongst human beings to the point where sensitivity, emotional care, and judgment are undercut (Pugh; 2024).

Such worries return us once again to the Turing Test, which in a delimited way asks us to explore what humans are to say and do when machines may be substituted for humans as interlocutors and not found wanting so far as communication of *information* is concerned. The heart of the matter here is the distinction between information transfer/efficiency and human conversation, meaningfulness, reasoning, feeling, valuation, co-evolution, and self-discovery. The latter take place evolving over time, they are not one-shot problems to be solved in a certain time frame or according to a finite set of principled algorithms or maximizing routines. In the case of Rawls's ideal theory, we are considering how best to orient democratic human conversation that can aid in the development of a sense of justice in individuals living over time in the context of a sufficiently "well-ordered society" (no one is starving, there is no existential threat from war or disease, most are capable of developing and living by way of a sense of justice, and so on) that itself evolves over time. And Rawls's assumption that a democratic society is well-ordered is analogous to Turing's assumption in the context of his "Imitation Game" that the players do not cheat, have some fairly extensive life experience over time with human face-to-face conversation, and are not already filled with resentment toward one another, or toward computing machines, but are open to discussion.

In both the scenario of the Original Position and that of the Turing Test, the possibility of developing what (Rawls; 2005) called an "overlapping consensus" concerning the design of specific systems is not to be construed in terms of a deduction from a list of fixed principles or algorithms, a consensus on outcomes or beliefs, pure analyses of concepts, or authoritarian structures of governance. Rather, the whole "game" of reconstructing principles is pursued in a spirit of democratic reasonableness, including educated gauging of history. In the context of the Turing Test, we imagine interlocutors with a certain degree of knowledge and sagacity. In Rawls's setup justice is construed as *fairness*, for which we are assumed to have an emotional sense, i.e., a capacity to sense equity in terms of status, solidarity, and representative voice, not simply economic position.

An "overlapping consensus" – as opposed to a general consensus as to what is a right way to live – emerges in situations in which we know ahead of time that parties are more than likely to bring to the table of discussion very different "comprehensive doctrines" about their places in the world and the nature of reality, while nevertheless they share a commitment to framing overlapping patterns of reasoning with respect to particular cases (Rawls; 2005). Thus, for example, the religious Catholic thinker, while deeply opposed to abortion or homosexuality, may be able to endorse as best a judicial line of reasoning that allows others to pursue an alternative form of life in which abortion in certain cases is acceptable, within limits, in such a way that room is left for serious discussion of and respect for, the traditional thinker's sensibility and point of view. A trans rights advocate may be able to affirm rulings that allow for some limitations on gender-affirming care, but allow for parental and individual choice and freedom of movement and protection of civil rights for all. If such endorsements are impossible, then democracy in Rawls's sense is impossible: all that remains is something merely "political" in the sense of the "political" as merely and only power, something immune to the concepts of fairness and reasonableness. Analogously, if two parties to Turing's Test bring to the table radically different conceptions of "intelligence" then the game cannot tell us anything at all. To make the game work, sufficient "overlapping consensus" on what is reasonably human linguistic behaviour is presupposed. In Turing's game a sense of embodiment, gender dynamics, creativity and interest are all assumed to form part and parcel of the human players' capacities.

In democratic societies any idealized situation of tolerance and fairness, given the complexities of these unwindings of human capacity, will have to be dynamic over time. It is likely, given biotechnological inventions that are likely to increase with the advent of AI, that an increasing variety of differing forms of life will become clearer and more vivid, perhaps evolving with unprecedented speed. In this case there is no hope of reaching agreement in one generation and finishing the conversation once and for all with a Last Word. Even if consensus were forced or gained in one generation, it could not dictate terms of justice and reasonableness to the next without resorting to state oppression (Rawls calls this "the fact of oppression" in (Rawls; 2005)).

What this shows is the foundational importance – both for Rawls's account of democratic justice and for Turing's Test – of the successful cultivation of resilient and shareable forms of morality, sensibility and feeling on the part of the human beings whose societies are reconstructed and measured by way of the philosophical models. Elemental senses of outrage and fear (e.g., concerning child pornography, sex robots, systematic exclusion and oppression of minorities, catastrophe), as well as ethical capacities and accomplishment ("small" acts of kindness, cruelty and heroism, the importance of care and devotion to friendships, the complexities of moral understanding) – will become an even more fundamentally important human task, and in fact one that is necessarily "democratic" in having to include human beings throughout the globe (as well, perhaps, as AI's).

Recall that Turing emphasized that "intelligence" is an "emotional" concept: it is response-dependent, not merely analytical. For this reason he predicted that, as AI comes to take over the labor of "machine table" work, members of the intelligensia would begin being propagandized and surrounded with "gibberish" and "mystery" (Turing; 1947), pp 495, quoted above. He had no expectation that democratic forms of status-marking would emerge naturally from the offloading of labor to algorithms. What he did predict is that the generation of "types" and genres, the higher-level concepts and vernacular expressions to help humans take in the flux and flow of forces around them, would become of central importance in any democratic scientific culture. Most indicatively, he dubbed the dynamic process of terminological and ontological negotiation a matter of "the cultural search" – a search, not an outcome – a quest in the Medieval sense.

Turing had established by proof the general limits of algorithmic decidability (Turing; 1936/37). This led him to insist that the development of "common sense" among members of society – something partly "emotional" – would remain an uneradicated task no matter how sophisticated and wide-ranging the artificial, algorithmic machinery might become. And in the importance of philosophy of education for democracy, he may be said to have followed Bertrand Russell, who wrote,

The thing, above all, that a teacher should endeavor to produce in his pupils, if democracy is to survive, is the kind of tolerance that springs from an endeavor to understand those who are different from ourselves. It is perhaps a natural human impulse to view with horror and disgust all manners and customs different from those to which we are used. Ants and savages put strangers to death. And those who have never traveled either physically or mentally find it difficult to tolerate the queer ways and outlandish beliefs of other nations and other times, other sects and other political parties. This kind of ignorant intolerance is the antithesis of a civilized outlook, and is one of the gravest dangers to which our overcrowded world is exposed. (Russell; 1940), pp 121.

Turing's "Imitation Game" is asking his readers to educate themselves, to display openness to the issue of whether it might make sense to say, in a very specific context, that a machine is "thinking" ("searching", "acting", etc.) – a matter for ongoing discussion and

investigation. He by no means thinks he has established that it actually does make sense. Sense and significance and meaning are not established by decree or operational procedure, but swim in the cultural context of the test – as Turing himself emphasized by setting up a control experiment for the Turing Test in which the job of the player is to distinguish a “man” from a “woman”.

What Turing emphasized, rightly, was that our future with AI will require ongoing human conversation about this very issue: the development of a sense of humanity. The purpose of the Turing Test – to repeat – was not to operationalize or treat as self-evident the notion of “intelligence”, much less to develop AGI (which is in any case provably limited by the undecidability results Turing himself established). Rather, Turing was setting up a pulse-taking conversation-inducer among human beings about their feelings and conceptualizations of intelligence and significance in the presence of machinery. This is an “experiment” in human-to-human conversation and connection in the presence of machines, rather than simply a test of machines with a view from Nowhere.

For, after the Turing Test is run, the screen comes down, and the human who has been asked to discriminate human from machine will be asked to go forward in conversation and life with the humans who have been discriminated for, or against, as human or as machine, as man or as woman. If a person has been too many times misclassified, she may not be willing to continue conversation with the Interrogator. It may not be reasonable to do so. This then transforms the situation, not into a scientific question, but a very human one: how to proceed with reasonable, democratic conversation – in shared forms of life – over time, where consensus is not to be expected.

Acknowledgments

My gratitude is due to Vanessa Nurock, who provided me with excellent comments on a near final draft of this paper. My thoughts have also been shaped by conversations with members of the audience at the UNESCO conference she organized at the Université de Côte d'Azur, France in September 2023.

References

- Abbey, R. (2013). *Feminist Interpretations of John Rawls*, PA: Penn State University Press.
- Adam, A. (1998). *Artificial Knowing: Gender and the Thinking Machine*, Oxford: Taylor Francis Group.
- Cavell, S. (1988). *Conditions Handsome and Unhandsome: The Constitution of Emersonian Perfectionism. The Carus Lectures*, University Of Chicago Press.
- Cavell, S. (1989). *This New Yet Unapproachable America: Lectures After Emerson After Wittgenstein*, Living Batch Press.
- Chude-Sokei, L. (2022). Race and the Singularity, In *Humanities and Artificial Intelligence*, Freddy Paul Grunert, Max Craglia, Emilia Gómez and Jutta Thielen-del Pozo, eds. *Noema Media and Publishing/European Union Commission Report* pp. 54–55.
- Cooper, B. S. and van Leeuwen, J. (2013). *Alan Turing: His Work and Impact*, Boston: Elsevier Science.
- Daston, L. (2018). Calculation and the division of labor 1750–1950, *Bulletin of the German Historical Institute, Washington, DC*, ed. Richard F. Wetzell pp. 9–30.
- Dewey, J. (1916). *Democracy and Education*, New York: Macmillan.
- Emerson, R. W. (1841). Self-reliance, In *The Portable Emerson*, edited by Malcolm Cowley, New York: Viking Press (Original edition, *Essays, First Series*, Boston: James Munroe and Company).
- Floyd, J. (2013). Turing, Wittgenstein and types: Philosophical aspects of Turing's 'the reform of mathematical notation' (1944–5), In *Cooper and van Leeuwen*, eds pp. 250–253.
- Floyd, J. (2017). Turing on Common Sense: Cambridge Resonances, In *Philosophical Explorations of the Legacy of Alan Turing – Turing 100*, edited by Juliet Floyd and Alisa Bokulich. *Boston Studies in the Philosophy of Science*. Dordrecht: Springer pp. 103–149.
- Floyd, J. (2019). La quête culturelle: revisiter le test de turing, *Cités* 80(4): 15–30.
- Floyd, J. (2020). Wittgenstein on ethics: Working through Lebensformen, *Philosophy and Social Criticism* 46(2): 115–130.
- Floyd, J. (2023). Revisiting the Turing test: Humans, machines, and phraseology, In *Katz, Schiepers and Floyd*, eds pp. 75–114.
- Floyd, J. and Katz, J. (2023). Interview with Stephen Wolfram, In *Katz, Schiepers and Floyd*, eds pp. 117–136.
- Hodges, A. (2012). *Alan Turing: the Enigma. Centenary edition.*, Princeton, NJ: Princeton University Press.
- Katz, J. E. and Aakhus, M. A. (2004). *Perpetual Contact: Mobile Communication, Private Talk, Public Performance*, Cambridge, U.K.: Cambridge University Press.
- Langer, S. K. K. (1957). *Philosophy in a New Key; a Study in the Symbolism of Reason, Rite, and Art*, 3d ed. MA: Harvard.
- Laugier, S. (2024). Culture as experience from Dewey to Cavell, *Journal of Aesthetic Education* 58(4): 99–116.
- Mills, C. W. (1997). *The Racial Contract.*, Ithaca: Cornell University Press.
- Nurock, V. (2007). Le sens de la justice selon John Rawls: à la confluence de la philosophie et de la psychologie morale, *Cahiers Philosophiques* 110: 65–92.
- Nurock, V. (2008). *Rawls pour une démocratie juste*, Paris: Éditions Michalon.
- Proudfoot, D. (2014). Turing's three senses of "emotional", *International Journal of Synthetic Emotions* 5(2): 7–20.
- Pugh, A. J. (2024). *The Last Human Job: The Work of Connecting in a Disconnected World*, Princeton University Press.
- Rawls, J. (1963). The sense of justice, *The Philosophical Review* (Reprinted in Rawls 1999b, 96–116) 72(3): 281–305.
- Rawls, J. (1999a). *A Theory of Justice*, Revised edition. Cambridge, MA: Harvard University Press.
- Rawls, J. (1999b). *Collected Papers*. Edited by Samuel Freeman, Cambridge, MA: Harvard University Press.
- Rawls, J. (2001). *The Law of Peoples with the Idea of Public Reason Revisited*, Cambridge, MA: Harvard University Press.
- Rawls, J. (2005). *Political Liberalism. Expanded edition*, New York: Columbia University Press.
- Rawls, J. (2007). *Lectures on the History of Political Philosophy*. edited by Samuel Freeman, Cambridge, MA: Harvard University Press.
- Rose, K. (2024). Can ai be blamed for a teen's suicide, *The New York Times* 2024/10/23.
- URL: <https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html?searchResultPosition=2>
- Russell, B. (1940). *The Functions of a Teacher.*, Harper's Magazine (reprinted in *The Basic Writings of Bertrand Russell*, eds. eds. Robert E. Egner and Lester).
- Searle, J. (1980). Minds, brains and programs, *Behavioural and Brain Sciences* 3: 417–424.

- Searle, J. R. (1984). *Minds, Brains and Science: the 1984 Reith lectures*, Cambridge, MA: Harvard University Press.
- Sterrett, S. G. (2000). Turing's two tests for intelligence, *Minds and Machines* 10: 541–559.
- Thoreau, H. D. (2015). *Civil Disobedience*, New York: Open Road Media Integrated Media. Originally published as “Resistance to Civil Government”, in Elizabeth Peabody, ed., *Aesthetic Papers*, Boston and New York: The Editor and G.P. Putnam. pp. 189–211 (1849).
- Trauth, E. and Quesenberry, J. (2023). *Handbook of Gender and Technology: Environment, Identity, Individual*, Cheltenham, UK: Edward Elgar Publishing.
- Turing, A. M. (1936/37). On computable numbers, with an application to the entscheidungsproblem, *Proceedings of the London Mathematical Society* 2(42): 230–265; Corrections, *Proceedings of the London Mathematical Society* 43(7): 544–546.
- Turing, A. M. (1944). The reform of mathematical notation and phraseology, *In Cooper and van Leeuwen, eds.* pp. 245–249.
- Turing, A. M. (1947). Lecture on the automatic computing enging, a lecture to the London Mathematical Society 20 february 1947, *In Cooper and van Leeuwen, eds.* pp. 486–498.
- Turing, A. M. (1948). Intelligent machinery: Lecture to the London Mathematical Society 20 february 1947, *In Cooper and van Leeuwen, eds.* pp. 501–516.
- Turing, A. M. (1950). Computing machinery and intelligence, *Mind* 59: 433–460.
- Wikipedia (2025). Women in Bletchley Park.
URL: https://en.wikipedia.org/wiki/Women_in_Bletchley_Park
- Wittgenstein, L. (2009). *Philosophische Untersuchungen = Philosophical Investigations*. Trans. By G. E. M. Anscombe, P. M. S. Hacker, and Joachim Schulte. Rev. 4th ed., Chichester, West Sussex, U.K. ; Malden, MA: Wiley-Blackwell.