A HOLISTIC TEST FOR (ARTIFICIAL) GENERAL INTELLIGENCE

A PREPRINT

Danny A. J. Gómez-Ramírez*
Professor and Reseacher at Parque Tech at
Institución Universitaria Pascual Bravo
Medellín, Colombia
daj.gomezramirez@gmail.com

Judith Kieninger
Research Institute for Economics of Aging
Vienna University of Business and Economy
Vienna, Austria
judith.kieninger@wu.ac.at

August 25, 2023

ABSTRACT

We approach the notion of general (global) human intelligence as a prominently multifaceted concept which can be tested in at least seventy specific scenarios. We say that an agent has Artificial Global Intelligence (AGLI), if it is able to perform in an intelligent manner for at least the collection of tasks defining the former scenarios. In particular, based on Gartner's multiple intelligences theory, we describe the design of a concrete test for AGLI made in such a way that an 'average' young human being should, per diffinitionem, pass it. We compare our holistic test with several highly representative human and artificial intelligence tests for deducing that our AGLI test is the one with the broader scope of the whole range of intelligence. However, we claim as a plausible hypothesis and based on the wide thematic spectrum of the test that the most outstanding inventions of AI currently available are far from getting a high score on it. Finally, we open a discussion about the fact that, although there are artificial agents being able to simulate intelligent behaviour at a perfect level for several local tasks, this seems not to be the case for AGLI. This would imply, in particular, a kind of practical implausibility of the notion of intelligence explosion coming from philosophy of AI.[2]

***K*eywords** Artificial General Intelligence · Intelligence Test · Local Intelligence · Strong AI · Weak AI

1 Introduction

What is the key feature of the human mind? What is its central ability, and how wide is the pragmatic spectrum of such an ability? On the other hand, could we artificially replicate this ability in a general manner? These questions has been presented throughout centuries in different forms as part of the most seminal philosophical and scientific concerns of mankind. In fact, the last one has a particular interest in the modern academic and artistic world. Moreover, the answer to the first question is codified by the word 'intelligence', which represents, on itself, a puzzle. The answer to the third question is encompassed in the concept 'artificial general intelligence', which seems to be even more mysterious.

Etymologically, the term 'intelligence' comes from the Latin words 'intellingentia', which, at the same time, comes from the work 'intelligere', meaning 'to understand' or 'to perceive'. Sometimes, 'intelligere' has been understood as 'intus legere' or 'to read inside'. However, the meaning and usage of 'intelligentia' has varied widely among the treatises of philosophy, requiring a robust context for its better understanding [Ferrater Mora1958]. Moreover, historically speaking, the notion of intelligence has been closely related to the ones of understanding, intellect and reason. And, without going into every minor detail of the definitions proposed, one can say that, in synthesis, intelligence can be described as the global faculty of comprehension, understanding, discernment and effective thriving of the (human) mind (see, for instance [Harper et al.2001]). Furthermore, the most fundamental dimensions of intelligence that have

*Multidisciplinary Researcher, Founder and Leader of the Consortium ARMAINTE and the Meta-Project Artificial Mathematical Intelligence https://dagomez1982.wixsite.com/artmathintelligence. Personal Website www.DAJ-GomezRamirez.com been generally studied are the internal (or more related with consciousness) and the external one (or more related with practical activities). In other words, one could perceive intelligence (theoretically) in a human being by analyzing the cognitive processes taking place into his/her mind, or (more pragmatically) by evaluating the particular actions done by the subject of study.

Historically, the study of the internal dimension of intelligence has been the most difficult and subtle one, since it involves a very detail understanding and structuring of consciousness and a general classification of the fundamental cognitive abilities used by the mind during the creative process. In the contexts of formal reasoning, one of the most outstanding results obtained regarding a global taxonomy of cognitive abilities of the mind is described within artificial mathematical intelligence [Gomez-Ramirez2020, Part II].

Furthermore, on the human quest for defining, evaluating and replicating human global intelligence in artificial agents, this inner dimension of intelligence has the additional challenge that almost all of the agents constructed so far (e.g. computer programs and robots) possess no global consciousness and human-like perceptual behaviour like in humans. More specifically, on the cognitive architectures that we use on the generation of 'artificial minds' there is always a local, and usually small, functional and phenomenological resemblance to the behaviour of the human mind, but not a global general similitude [Lieto2021].

So, going back to the original question giving birth to the modern idea of artificial general intelligence (AGI), we see that for evaluating objectively AGI in an artificial agent, the second (external) dimension of intelligence is the central dimension to analyze. In fact, it could theoretically happens that an artificial agent can be construct with a initial level AGI, but no trace of (human-like) consciousness on it. Thus, one could say that such an agent possess AGI only regarding the second dimension. On the other hand, it is plausible that we can create a sort of level of consciousness in an artificial agent (for example, in terms of integrated information theory [Tononi2012] and [Tononi et al.2016]), without implying at the same time that we have intelligent behaviour.

So, due to the current challenges that exists regarding how to capture and to assess (artificial) conscious states, and even more, (artificial) internal 'subjective' intelligent acts; the most objective manner of evaluating general intelligence is by measuring external pragmatic actions of the potential (artificial) agents.

In the present paper, we will deal essentially with some aspects of the third question from a pragmatic manner. In other words, without going deeper into ontological and semantic issues about what intelligence is, we will develop a new systematic manner of testing if an (artificial) agent possesses human intelligence at a general level or not.

On this regard, the theory of Howard Gartner about multiple intelligence turns out to be extreme relevant for our purposes. In a nutshell, the theory of multiple intelligences proposes a global classification of the most outstanding forms in which (human) intelligence can be materialized. More specifically, the theory of multiple intelligences (MI) was originally introduced in 1983 in [Gartner2011]. In this seminal work, Gartner based on a considerable big amount of empirical studies psychology (e.g. psychometrics), neuroscience, cultural studies and anthropology approached the notion of intelligence in broader and more universal manner in comparison with classic psychological studies focused mostly on the logic and linguistic aspect of intelligence assessed by (multiple choice) written tests (e.g. IQ tests) [Gardner and Moran2006]. On the development of his theory, Gartner took into consideration the full range of diversity that intelligent actions and intelligence assessments can possess along different types of cultures around the globe; and enriched the classic mono-disciplinary studies about intelligence and intelligence assessment with this new methodological dimension and with neuro-biological relevant research, among others (see for example [Chen et al.2009]).

Furthermore, this broader and multidimensional conceptualization of intelligence is more coherent and more closely related with personal real-world success and global human development than the classic one assessed mainly by IQ tests [Moran and Gardner2006].

On the other hand, from a purely scientific perspective Gartner's theory was obtained by means of a formal process of general synthesis of a wide plethora of empirical data going far beyond classic psychology. So, MI theory cannot be judged and understood correctly if one is immersed only in a psychosomatic context, one needs a wider multidisciplinary standpoint including related disciplines. Now, this larger formal modus operandi brings with it a bigger experimental challenge, i.e., testing multiple intelligences in humans (and extensively in artificial agents) is more complex, more expensive and more demanding than performing the corresponding paper-and-pencil (IQ) tests. Nonetheless, despite of the former logistic challenge the MI theory has been highly well received in the educational surroundings. And, not only that but it has inspired dozens of teachers in finding more effective methodologies in the classroom [Armstrong2009].

Due to the former reasons, and avoiding any further standard academical discussion about the semantics of the concept of intelligence, we adopt here in general terms the MI theory as a suitable formal framework for understanding general intelligence, firstly on humans and extensively on artificial agents.

Before proceeding with the pillars of our deductive treatment, and for obtaining more sensibility regarding the notion of intelligence, let us present an enlightening initial example in the form of a question.

Who was more intelligent Albert Einstein or Charles Chaplin? If we want to give an objective and specific answer to this question, we should say immediately: ‘it depends on the kind of intelligence we are talking about’. So, as we mentioned before, let us consider ‘intelligence’ at least in the very broad sense of H. Gartner, i.e., an initial taxonomy of this concept should contain at least the following types of intelligences: logical-mathematical, linguistic, spatial, musical, bodily-kinesthetic, interpersonal, intrapersonal, naturalistic and existential [Gartner2011], [Gartner1995]. Here, it is fundamental to clarify that for the purposes of this paper, we are essentially concerned with finding a global and pragmatic way for measuring a lower bound of the ‘general intelligence at human level’ that any (potentially intelligent) agent could have. Therefore, the meta-physical question regarding the meaning of the concept of ‘intelligence’ will not represent our main focus of interest. In fact, independently of the definition of intelligence that one uses, and of how many types of intelligence there are, our approach will be based more on the multifaceted way in which intelligence is materialized among humanity, with a huge amount of ‘acts of intelligence’ being performed through several channels by means of the body-mind (physical) structure.

So, going back again to our last question, the answer would require a deeper qualitative description taking into account the global diversity of intelligence. For example, we could require measuring how good was Einstein in acting and how good was Chaplin in understanding physics and mathematics.

In general, a plausible answer would be that if Einstein and Chaplin were alive, we should perform a long test with both of them, where we design particular tasks regarding at least the former types of intelligence.

This kind of general test would give us a global measure of the ‘intelligence’ of Chaplin and Einstein in a very comprehensive manner. In fact, due to the quantitative aspect of each of the tasks, we could obtain a specific numerical value of the (general) intelligence of each one in order to be able to compare them globally. So, we could finally get a concrete answer to our original question. In fact, let us assume that, hypothetically the performance of Chaplin on the linguistic, spatial, bodily-kinesthetic, interpersonal and music intelligences was notably better than Einstein’s, and that Einstein’s performance was better in the logic-mathematical and intrapersonal parts. So, we would conclude that, according to the performances of these two personalities during, let us say, more than 200 testing hours, Charles Chaplin is more intelligent than Albert Einstein. In fact, the broader we understand intelligence, the better we understand that Chaplin could even be more intelligent than Einstein under those circumstances.

Note the ontological enhancement that our pragmatic conceptualization of ‘intelligence’ possesses in comparison with the initial notions that some of the fathers of AI used (see for example [Newell and Simon1975]). Explicitly, the pragmatic aspects of the origin of AI were strongly influences by the generation of domain-specific intelligent agents in just a few areas like concrete logic/mathematical theories, language and strategy creation [McCorduck2009, Russell and Norvig2016].

Another fundamental point for judging intelligence is the one of specialization. A very common argument for (non-) comparing intelligences can be described as ‘each person is good in something very unique, which is quite specific and differs between people, therefore one cannot make comparisons concerning (global) intelligence’. To this argument there is a quite general remark: It used to happen that people that are very good at very specific fields, simultaneously start to develop an ability to be able to understand faster other fields, sometimes related with their specialized fields, sometimes not.

For example, the work of Descartes on obtaining better models of geometrical objects led him to use methods of classical algebra in order to create his ‘analytic geometry’ [Descartes1954]. Andrew Wiles’ effort on proving the insolvability of Fermat’s Last Problem (a pure arithmetical problem at the first sight) led him to find more precise connections between the areas of elliptic curves and modular forms [Wiles1995]. The creative enterprise of Carl Levi-Strauss of making anthropology into a more formal science guide him to use Saussure’s structural linguistics and some aspects of combinatorics for developing his structural anthropology [Levi-Strauss2008]. Furthermore, Bertrand Riemann in order to get a deeper understanding of the distribution of the prime numbers had to develop a new way of seeing functions with complex variables [Mazur and Stein].

Now, it is important to clarify that these phenomena occur, only if a person is flexible enough, intellectually speaking, to grasp deeply into other fields. Thus, a practical and vivid 'intellectual openness' would be necessary for getting broader scopes in our comprehension of very specific intellectual fields. [3]

We want to draw attention to the fact that some people are able to learn regularly within some months (in the university) several kinds of scientific and artistic disciplines, to play different kinds of sports and music (instruments), to interact with many people at a wide range of levels in order to obtain specific and differentiated goals with any of them, to move to several places in minutes just by following quite abstract symbolic information expressed in oral and written language, to select at a quite unique and creative way of styling, food and places to live in; to integrate themselves relatively fast into a very sophisticated system of services, products, artificial laws and cultural schemes; to obtain a world wide perspective of the current political, social and economic state of the world just by processing light and sound configurations of an electronic device (e.g. smart phone, computer); to create quite innovative solutions in order to improve the life's quality of other people and to modify whole pieces of natural environments in prefixed extraordinary ways; among many, many others.

So, any agent having general intelligence at human level should be able to perform at least the sort of activities described here, since these tasks are typically made by humans, and evidently require intelligence for being performed successfully. Now, what is, objectively speaking, the state of the art of the current AI existing agents regarding general (global) intelligence? Here, it is worth to mention that in the particular field of AGI there are at least two (non-scientific) external sources which strongly influence the objective estimations of people concerning this issue, namely, the cinematographic and the machine industries.

In particular, it is quite plausible to think that films like 2001: A Space Odyssey; Artificial Intelligence; I, robot; Ex-machina and The Star Wars Saga, among others; have a strong influence, at least cognitively speaking, on people's own perception of the current state of the art of AI (devices). Effectively, the goal of some of these films (e.g. Ex-machina) is precisely to produce the temporal illusion that a hypothetical researcher was (almost) able to produce AGI. Of course, the directors wish to present their stories in a very contextualized and realistic way and we could say that as pieces of art these movies are very creative contributions. However, the objective contribution of these films to the development of AGI is a quite different story. On the other hand, in the case of companies trying to sell some 'intelligent' devices, it is used to happen that there is sometimes an oversized description of the 'abilities' of the corresponding devices. One of the main causes of this phenomenon is clear: they want to sell a product, so they apply all kinds of marketing strategies to obtain this goal. Now, how could we distinguish between scientific, fantastic and marketing-driven results on AI research?

The former remarks are completely relevant for the main line of argumentation presented here. Effectively, AI is currently one of the few fields of research that is deeply influenced and, sometimes shaped, by politics, the arts, the military, the industry, and the entrepreneurial field; so, the objective perspective that a person could have about the real state of the art of artificial (general) intelligence is, in most of the cases, influenced by a lot of subjective, non scientific and biased information.

Thus, one of our main purposes is to bring objective light on the suitable assessment and the state of the art of (artificial) general intelligence, implicitly based on a cumulus of scientific studies from a wide collection of systematic disciplines.[4]

2 Designing a Global Intelligence Test

Nowadays, there are thousands of academic publications, research institutes, movies, videos, devices, machines and advertisement regarding the further progresses we are obtaining towards constructing 'intelligent machines', i.e., machines that perform intelligence at a human level. In fact, strong artificial intelligence (also known as artificial general intelligence (AGI)) is, in some contexts, defined as "machine intelligence with the full range of human intelligence" [Kurzweil2006]. Now, if AGI is pursuing to construct a machine which can perform and, at least simulate, human-level intelligence in its whole spectrum then, what kind of (general) test should such a machine pass in order to get in an authentic way the attribute of 'being intelligent at a full human-level'?

In order to be able to answer this question in a very concrete way let us start to review, at least very globally, the collection of activities where humans can perform an intelligent behavior.

[3] Here, the expression 'intellectual field' encompasses all the possible areas related to at least one of the former types of intelligence, from classical sciences to any kind of sport and artistic activity.

[4] We say "implicitly" due to the fact that the MI theory is the concrete carrier of this cumulus of fields of study.

First, the classical tasks where one can perform an intelligent behavior are related with sciences and engineering, e.g., solving a mathematical problem, a problem involving theoretical physics, involving a basic knowledge in chemistry, biology; writing an algorithm for solving a concrete computational question; estimating the approximate age of a particular object; answering a question regarding historic information; winning in a (virtual) game, estimating the age of an observable star; producing a material with some fixed predetermined physical properties; designing an experiment for verifying a physical theory; constructing a machine which performs a repetitive physical work; finding a (new) medication for a decease; building a house, designing a web page for advertising a product, among many others.

Second, the tasks related with text comprehension and text production: answer concrete questions referring to a short text, writing a summary of a book, checking grammar and style of a text, describing verbally some particular event (e.g., a sport game), among others.

Third, the activities regarding an artistic task: playing a particular instrument, composing a piece of music, dancing to a particular kind of music, painting a landscape, performing a role (acting) based on a determined screenplay, sculpting a human body, among others.

Forth, the kinematic intelligence of a person can be tested in any kind of sport and some of the former artistic activities, among others. For example, winning the final of a football/basketball championship, performing an almost perfect score in (Olympic) gymnastics, winning a tennis Open, setting/breaking a domain-specific kinematic World Guinness record, and creating a choreography for a dancing championship.

Fifth, the social human intelligence can be tested in activities like teaching at school, doing politics with concrete (definable) purposes, acting as a diplomatic (e.g. peace) intermediary, performing a (standard and successful) psychotherapy, performing a stand-up comedy show that effectively makes the audience laugh, selling a product or service and gathering a specialized collection of people which can form a robust team for creating a concrete kind of enterprise.

We can continue with the description of a lot of other daily activities requiring human intelligence to be performed, but we included here the former ones in order to sensitize the reader regarding the huge amounts of (intelligent) actions that are done by humans at a daily bases, and that should be potentially learnt by any artificial agent possessing the status of being 'intelligent' at a human level.

Even more, since our discussion goes in the direction of evaluating as objective and real as possible (human) general intelligence, the activities of our test should be performed in the most natural and realistic manner. In other words, we do not need to reduce ourselves to a paper-and-pencil activities only, since general intelligence goes far beyond that. So, we can add everyday life activities in order to evaluate more rigorously the attribute of 'intelligent' in man-made agents.

Now, let us design a test to prove that a specific (artificial or human) physical agent 'X' (e.g., a person, a robot, an humanoid, a computer, a machine) performs *artificial global intelligence* at a human level (AGLI).

First, let A denote a concrete activity where the intelligence of any generally healthy human being can be tested. We will provide a complete list of such activities included in the test later.

We choose seven (professional) experts on the activity A to form the specialized committee C_A. Furthermore, C_A will design the following sections of this part of the test:

1. A learning section, where X learns the very basics of A and will be evaluated in accordance with their performance.
2. A basic section, where X carries out a concrete test in order to show what they have learned.
3. A creative section, where X tests their ability to solve a challenge, whose solution requires the combination of the former basics of A in a (more) creative way (in comparison with the former sections).

The second section would require a standard application of the basic tools acquired in the former one in order to be passed. It is similar to the final exam performed by a first semester student concerning the topics learned on one/two terms. The third section would require a little more of creative genuine thinking, since it would involve blending the former basis cognitive capabilities in a new way for finding the solution. If we restrict ourselves for a moment to an activity involving mathematical reasoning, then this section is similar, to the most basic problem of a classification test in order to form a national team of the well-known International Mathematical Olympiad (IMO) [imo].

Further, the first and second sections are designed to evaluate the ability of X to learn and to integrate the basic patterns regarding the activity A, assuming that X has no previous formal knowledge or training on A.

Lastly, the last section is drawn to measure any kind of more specialized ability regarding A beyond the average.

The maximal possible grade is 10 and the minimal is 0,01. The intermediate results are described with two decimal digits.

The contribution of each section is 40, 40 and 20 percent, respectively. Moreover, the grade 0,01 will be given to an agent X who is basically still during the whole test, i.e., the C_A cannot see any kind of perceivable answer of X.

Each of the C_A will design the specific test as if X was an 'average' healthy 25 years old person, when the vast maturity in the network of the brain and the human mentation is most generally reached [Pujol et al.1993]. In fact, the test will be designed such that this person would obtain a score of at least 6 assuming that they are just able to learn the very basics related with A.

Here, we are just interested in intelligence at a human level regarding any of the activities. Therefore, we restrict ourselves to agents having a 'human form'. It means, the physical-mechanical structure of the agent X should be isomorphic to the one of a human body, i,e, it should have one mechanical (m) head with two m-eyes, a m-nose, a m-mouth, two m-ears, 32 m-teeth; two m-forearms, two m-arms, two m-hands, two m-legs and a m-trunk, 10 m-fingers, a m-voice, among others. The reason for this is that we want to test the agent X performing A at a human level. And, for example, when we test the performance of X regarding 'acting', X should be able to perform facial 'human like' emotions. Besides, when we test 'soccer', X should be able to perform the basic kinds of moves that a standard football player does with their two feet.

It is important to mention that for any activity A we choose a specialized committee C_A having a well-accepted and long enough experience on A. This is in order to guarantee a minimal degree of 'objectivity' in the design and evaluation of the corresponding test. For instance, if we are considering the task of 'composing a little piece of music', then it is clear to assume that a professional composer would evaluate the quality of such a composition more objectively than an AI-researcher, who in principle is not so trained in music and could be influenced partially by the fact that he or she is pursuing to design agents with intelligence at human level.

In most of the specific activities to be evaluated, one can imagine that, for avoiding a bias by the specialized committee, the agent X (either a human being or a robot) could cover their whole body with clothes, (cloth) masks, gloves and glasses. The main motivation for this kind of technical measure is to minimize as much as possible any kind of bias that could appear by the domain-specific testers in the case that they could know in advance that X is not a human. Nonetheless, even if they know in advance the (non-)anthropological origin of X, they would be able to apply without major problems the specific test, since *per definitionem*, they would deal with X as if it were a human.

Regarding an activity like acting, the former measure would be not so straightforward to apply since a considerable amount of the quality of an acting performance depends deeply on the specific gestures and facial movements of X, which could not be deeply appreciated (and evaluated) with a mask, for instance.

A lot of intellectual tasks (considered in the former wide spectrum of materializations) that we used to admire, appreciate and reward in modern society require of a very sophisticated constellation of 'cinematic steps' (or movements), mostly before, during and after the achievement of the desired goal. Thus, it is fundamental to point out the fact that our test pays a very close attention to the (cinematic) manner in which X perceives, processes and returns information (in the form of instructions, for example). We are not just testing a kind of abstract intelligence that provides conceptual and invisible solutions to problems, but also a physical intelligence that moves in the spatiotemporal environment and through these movements provides intelligent and pragmatic solutions to a multifaceted collection of tasks (like the human being used to do).

At this point, it seems straightforward to say that our holistic (artificial) general intelligence test, when applies to human beings, serves as well for testing intelligence in the whole sense of the MI theory. In other words, putting the financial and logistic conditions of the trial aside, the test could be perfectly taken by any healthy human being without any problem.

3 Specific Tasks

We describe in this section a minimal collection of tasks that an agent X should be able to perform in order to obtain the attribute of 'being (globally) intelligent at a human level'.

Let us fix some terminology: LB(Y) would be an abbreviation for 'learning the basics of Y (e.g. main concepts and results) and solving tasks in this subject at the level of an undergraduate book for freshmen[5], or at the level of a standard beginner of a private academy (or suitable institution)' [6].

[5]In the case of an academical subject or some types of arts.

[6]In the case of a sport, some arts and others.

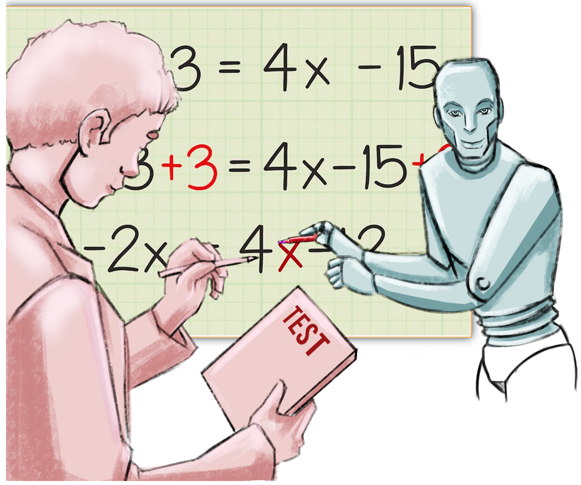

Figure 1: Logical-Mathematical Intelligence

SB(L) abbreviates 'learning the basics of the language L and speaking L at the basic first level'.

PB(I) stands for 'playing the instrument I at a basic level and being able to play a basic piece of music (alone and within a musical group), to compose at a basic level and to improvise'.

CB(F) signifies: 'being able to cause the mood (resp. feeling, reaction) F at different levels in an arbitrary person through a conversation'.

The following local sub-tests are the specific activities that X should globally approve in several domain-specific areas in order to pass our test.

Now, regarding the particular choice of the activities to be evaluated, let us keep in mind that, in principle, we do not have here the same spatio-temporal restrictions that the classic intelligences do. In other words, we can include hundreds of hours in our test in order to be more accurate in the whole assessment. On the other hand, the more realistic the testing sections are with the corresponding real-life counterpart the more precise the test is.

Therefore, the following choices of the representative specific activities in the test to be described in the next subsections were done with the following criteria: a) to possess pragmatical resemblance to actual tasks in (human) daily life. b) to belong to at least one of the forms of intelligence in the MI theory. c) To exhaust the whole taxonomy of intelligence described in the MI theory. And d) To solve the most outstanding intellectual challenges that an adult intelligent human being in modern society should face.

So, we the former criteria in mind, we see that there is flexibility in the specific choice of activities that can be tested. So, our choice represent less a canonical one, and more a proto-typical one.

3.1 Logical-Mathematical Intelligence

1) LB(Mathematical Logic) 2) LB(Elementary Mathematics) 3) LB(Elementary Physics) 4) LB(Elementary Biology) 5) LB(Elementary Chemistry) 6) LB(Elementary Computer Science) 8) LB(Chess) 9) LB(Go).

Even though mathematics and logic can be thought to be more elementary than physics, biology, and chemistry, the specific knowledge in these applied disciplines also contributes to a particular way of thinking and therefore cannot be separated from a general account of intelligence.

3.2 Verbal-Linguistic Intelligence

1) SB(Mandarin) 2) SB(Spanish) 3) SB(English) 4) SB(Hindi) 5) Being able to write a coherent summary about a novel. 6) Being able to do paraphrasing. 7) Being able to write and to declaim a poem in English related with a given topic. 8) Being able to construct at least one example of conceptual blending of two given concepts. 9) Being able to describe a specific situation or topic in terms of abstract metaphors.

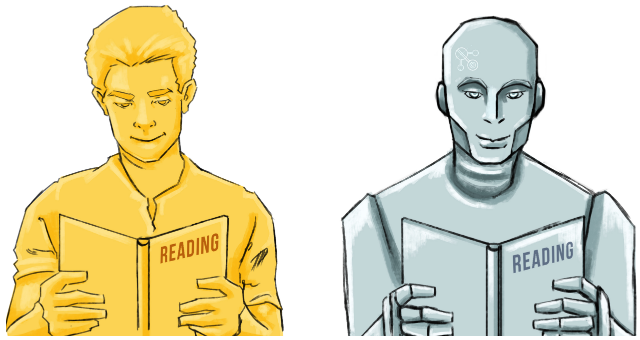

Figure 2: Verbal-Linguistic Intelligence

Figure 3: Musical-Rhythmic and Harmonic Intelligence

The last two tasks were chosen because of the central role that formal conceptual blending [Fauconnier and Turner2003], metaphoric reasoning [Lakoff and Johnson1980], and analogy making [Gust et al., Schmidt et al.2014] plays in human intelligence. Besides, the choice of the languages is based on the fact that an average human being is able to learn, with the appropriate training and with some effort, around three or four languages. In further, the four chosen languages are the most representative ones nowadays.

3.3 Musical-Rhythmic and Harmonic Intelligence

1) PB(Piano) 2) PB(Violin) 3) PB(Guitar) 4) PB(Drums) 5) PB(Percussion) 6) PB(Singing) 7) PB(Flute) 8) PB(Trombone) 9) Being able to perform at a basic level musical improvisation. 10) Being able to play 'harmonically' in a band in two scenarios: fixed compositions and group improvisation. 11) Being able to direct a musical assemble (e.g. a choir) at a basic level.

The former instruments were chosen in order to have at least one representative regarding to the form, material, and the way the instrument is played. It is clear that we could expand the list of instruments to other traditional but non-standard ones, but for the sake of effectiveness and concreteness in our test we restrict ourselves to this list trying to capture in a minimal way a general spectrum of musical performance and composition.

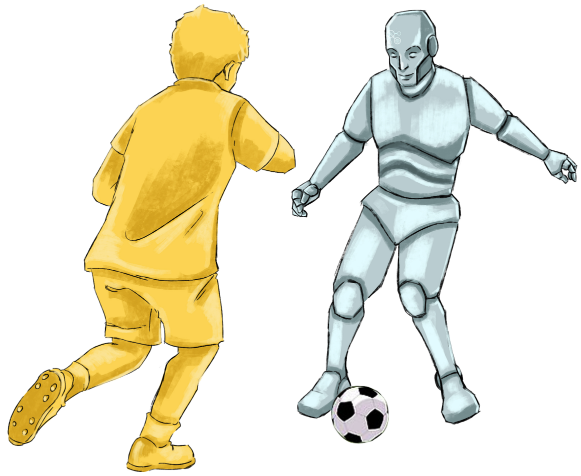

Figure 4: Bodily-Kinesthetical Intelligence

Figure 5: Visual-Spatial Intelligence

3.4 Bodily-Kinesthetic Intelligence

1) LB(Soccer) 2) LB(Basketball) 3) LB(American Football) 4) LB(Gymnastics) 5) LB(Swimming) 6) LB(Bouldering) 7) LB(Baseball) 8) LB(Climbing) 9) LB(Body Building) 10) LB(Ballet) 11) LB(Break dance) 12) LB(Salsa) 13) LB(Judo) 14) LB(Acting) 15) LB(Sculpting)

The same kind of criteria applied to the construction of the former list hold here regarding sports, dance styles and others, e.g. elementary capabilities (walking, balancing, hand-eye-coordination, and the like).

3.5 Visual-Spatial Intelligence

1) LB(Euclidean Geometry) 2) LB(Drawing) 3) LB(Painting) 4) LB(Photography) 5) LB(Architecture) 6) LB(Differential Geometry in 2 and 3 dimensions) 7) Being able to find a specific location just with the help of a map.

3.6 Interpersonal Intelligence

1) Being able to initiate and to keep an informal conversation (about virtually any topic) like a human being. 2) Being able to make general suitable inferences (in the form of coherent qualitative estimates) about other people's moods,

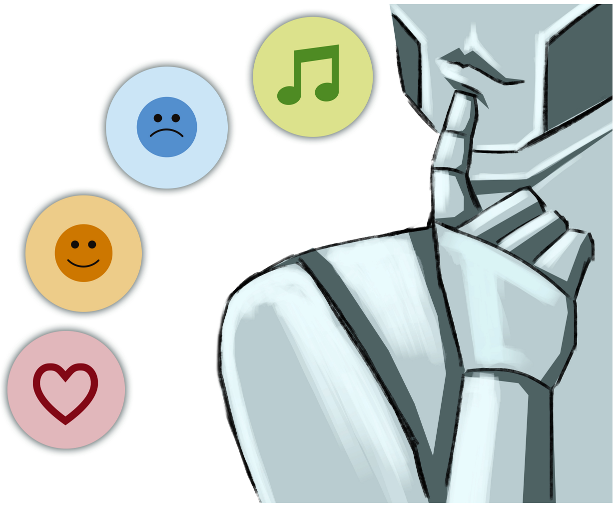

Figure 6: Visual-Spatial Intelligence

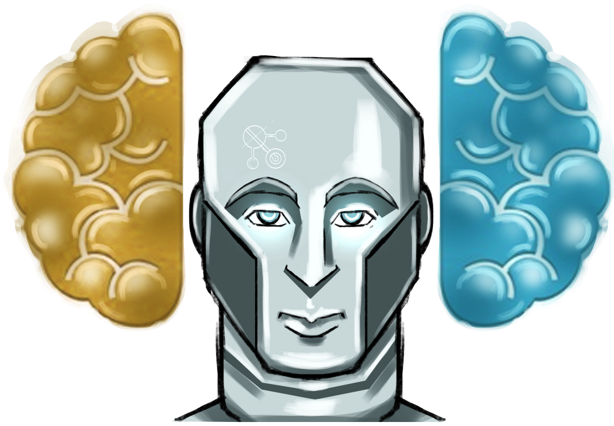

Figure 7: Intrapersonal Intelligence

feelings, personalities, and motivations after having a (casual) conversation with them. 3) Being able to perform successfully various basic and concrete charity activities. 4) CB(Laughing) 5) CB(Admiration) 6) CB(Surprise) 7) CB(Happiness) 8) CB(Courage)

The first activity of this specific kind of intelligence includes conceptually the classic Turing Test [Turing2009]. So, with the current terminology developed so far, the original way in which Turing conceived conceptually the notion of 'intelligence', being measured through (and only through) his test possesses more a local than a global character.

3.7 Intrapersonal Intelligence

1) Being able to answer open questions of any kind regarding its origins, physical and conceptual abilities, purposes, among others. 2) Being able to make general suitable and coherent qualitative estimates about other people's self-perspectives.

In order to be able to test objectively this kind of intelligence in an agent X, the corresponding C_A should have a relatively complete amount of information about how X was constructed, in the case that X were a non-human agent. This is necessary because C_A should be able to compare this information with the answers that X can give about itself.

3.8 Naturalistic Intelligence

1) LB(Hunting) 2) LB(Farming) 3) LB(Botanic) 4) LB(Practical Zoology) 5) LB(Practical Marine Biology) 6) LB(Astronomy) 7) Being able to do a kind of (m-)hygiene and related measurements in order to keep a functional (m-)body and functional (m-)mind. 8) Being able to make a verbal detailed description of the particular natural surrounding where the agent is located.

3.9 Existential Intelligence

1) LB(Philosophy: Metaphysics) 2) LB(Philosophy: Theology)

3.10 Successful Intelligence

In this section, we will include a quite foundational aspect of intelligence related with being able to set goals on the short-, middle- and long-term; and being able to generate (and modify) suitable plans in order to achieve such goals.[7] We could have included the tasks of this section within the inter-personal (or social) intelligence's section, but due to their own importance in the human development, we decided to include them in a single category (see for example [Sternberg2015] and [Albrecht2006]).

1) LB(Setting Coherent Goals/Visions) 2) LB(Generation of a Realistic Plan for the achievement of a Goal/Vision) 3) LB(Evaluating the Degree of Fulfilment of a specific Plan of Action) 4) LB(Thriving and Updating Goals/Visions and the corresponding Plans of Action according to a New Contextual Conditions)

As we mentioned before, it is clear that there is some relative flexibility in the choice of the specific activities. What seems to be necessary is that a minimal list of tasks for evaluating an agent X in the very wide spectrum of what human intelligence could encompass, should contain at least around the quantity of tasks described here, if not more.

Due to the fact that any concrete materialization of the former test would require that each group of experts design in a specific and measurable way each of the particular sub-tasks related with the corresponding activity, our test would not have one of the typical shortcomings that a considerable amount of artificial intelligence test possesses (for example, the Turing Test), i.e., a low level of accuracy in the description of the tasks [Li et al.2018].

As the reader can now appreciate in more detail, due to the universality of the ontology proposed by the MI theory, the full collection of former chosen activities represents a suitable plethora of endeavors for testing (artificial) general intelligence. Even more, as we mentioned before, the resemblance to a real-life situation of our test is bigger in comparison with all the standard intelligence tests proposed and implemented so far. So, although the materialization of our holistic test would require a considerable amount of people, time and space, it is completely plausible to come the test into being with the best similitude to a real-world scenario.

3.11 A fundamental Additional Security Measure for Testing the Artificial Genuineness of a Specific Testing Agent

In order to test if a specific artificial agent X possesses global intelligence at a human-level or not, one should not only conduct (all) the former (local sub-)test(s), but also one should be able to prove genuinely that X is truly an artificial autonomous device, which, for example, is not remote controlled and moved through any kind of internal sub-device.

Currently, this security measure is totally necessary due to the outstanding progresses in nanotechnology and remote control's devices [Zhang et al.2009, Mamalis2007, Valcárcel et al.2008, Gaw2016, Choudhury and El-Nasr2019, Ginzburg et al.1966]. In other words, one should be completely sure that X is neither a remote controlled device which is operated 'live' by intelligent humans controlling its outputs, nor an actual human being (assuming that X was declared as artificial in advance).

[7] In this paper, we will not enter into the discussion about whether this kind of intellectual capacity is, in fact, a type of intelligence *per se* or not. We just use the denomination 'success intelligence' for the sake of succinctness and similarity with the other parts of our holistic test.

3.12 Global and Local Intelligence

We say that an agent X has a specific level L of *local intelligence*[8] , if there exists a particular task A (like the ones described before or other ones preserving the level of specificity of the former ones), such that X performs A with a punctuation of L after passing the corresponding particular test as described in the last section.

Additionally, we say that X has level G of *general intelligence* (or *global intelligence*), if the punctuation of X is G after performing the former general intelligent test.

4 Comparison with Outstanding (Artificial) Intelligence Tests

Due to the fact that we judge a (human or artificial) agent regarding its performance with respect to human general intelligence (as an standard initial point of reference), we will assume in our discussion that a (human) intelligence test represents also an artificial (general) intelligence test. In other words, if we need to verify that a concrete agent X possesses human intelligence at a global scale, then X should be able to pass, at least, the typical (human) intelligence test that we have designed to examine human intellect.

With this clarification in mind, first we contrast our holistic test with other relevant human intelligence tests, and subsequently, we make the comparison with a collection consisting of the some of the most relevant artificial (general) intelligence tests.

Due to the fact that the vast majority of psychologically-based intelligence tests are framed in a passive question-answer (or stimulus-response) methodology, they will implicitly avoid most of the sorts of (materialization of) intelligence already mentioned, e.g. dynamic assessments, neurologically based instruments and psychometric-ability instruments [Daniel1997]. In fact, the narrow scope of the (standard) types of (psychological) intelligence tests is something that has been pointed out before explicitly in psychology [Daniel1997]. Of course, it is easy to understand that due to pragmatic reasons (like duration and financial costs) such intelligence tests have been narrowed from their very initial conception. At the same time, it is important to clarify at this point that those pragmatic constrains are not bounding considerably our multi-faceted test, since one of our main goals is being able to create an initial globally-conceived intelligence test that not only can be potentially done in a plausible and coherent manner, but also that can measure quantitatively and qualitatively global (resp. general) intelligence in an effective manner through a minimally robust concrete taxonomy of its several materializations/forms.

Let us start our travel throughout classic and modern intelligence test (strategies). For example, one of the most widely used IQ tests that has also been modified and adapted several times until today is the Wechsler intelligence scale for children and adults assessing important intelligence factors like verbal comprehension, perceptual organization, processing speed, freedom from distractibility, and so-called fluid intelligence (i.e. the ability of doing effective reasoning with novel information) [Daniel1997].

The Stanford-Binet intelligence scale measures crystallized intelligence (i.e. verbal reasoning and quantitative reasoning), abstract-visual reasoning and short term memory. It depends mostly on verbal skills and thus presupposes formal education [Daniel1997]. Similar to Wechsler's IQ test, another widespread intelligence test, originally developed in 1936 puts less emphasis on educational background, language, reading or writing skills. The Raven's Progressive Matrices (RPM) is based solely on geometric figures and was used by armed forces to measure nonverbal intelligence of entrants [Raven2000]. The Kaufman-Brief Intelligence (KBI) test, first published in 1990 comprises both, nonverbal knowledge such as the ability to solve novel problems (fluid thinking) and verbal knowledge expressed by vocabulary and definitions (crystallized ability). The KBI is used for screening high-risk individuals in institutional settings, for gifted and talented screening or when longer assessments are not possible [Hays et al.2002].

Now, by doing a general analysis of the former intelligence tests, we deduce that all the essential features of each of them are implicitly evaluated several times during the different sections of our holistic test, mostly in a local and multi-formed manner. For example, the fluid intelligence is widely evaluated during all the (74) creative sections, since in this part of the local tests the agent should be capable to deal creatively with new pieces of information that were not necessarily included in a complete manner in the former sections. All of the defining factors, reflected in the mentioned IQ tests are measured in some of the specific sorts of sub-tests, e.g. the ones belonging to logical-mathematical, verbal-linguistic and visual-spatial intelligence. Regarding short-term memory, our test measures throughout all its parts this necessary ability for human intelligence, due to the fact that each of the members of the specific specialized committees speaks with the agent in a human-style manner. Even more, this is tested additionally in the learning part

[8] Here, the meaning of 'local intelligence' goes in the same direction of the meaning of the classic (AI) notion of 'weak intelligence' like in [Newell and Simon1975, Kurzweil2006].

of the logical-mathematical sub-test, for example, in exercises involving the solution of several types of arithmetic problems.[9]

Another popular intelligence test used at (pre-)schools and other applied settings is the Woodcock-Johnson Tests of Cognitive Abilities based on the Cattell-Horn-Carroll theory. First developed in 1977, the test was revised three times with the latest adjustments in 2014 referring to as WJ IV [Mather and Jaffe2016]. The broad abilities measured on one or several of the three types (batteries) of Woodcock-Johnson tests are: fluid intelligence (e.g. quantitative reasoning), visual-spatial ability (e.g. mental rotation), processing speed (e.g. perceptual speed), Long Term Retrieval (e.g. memory for names), Auditory Processing (e.g. phonetic coding, word fluency), Short Term Memory and Crystallized Intelligence (e.g. listening ability, lexical knowledge) [Abu-Hamour et al.2012].

Regarding tests evaluating explicitly emotional intelligence (EI), note first that our test captures the essential aspects of this fundamental manifestation of intelligence within the sub-tests related to inter- and intrapersonal intelligence (see for instance [Mayer and Salovey1993]). So, the Mayer-Salorevy-Caruso emotional intelligence test, for example, will be implicitly covered by the two former types of intelligence described before [Brackett and Salovey2006].

A rather different theory of intelligence testing goes beyond memory-based skills and analytical abilities and focuses on a combination of analytic, creative and practical intelligence resulting in Sternberg's triarchic theory of human intelligence [Sternberg et al.1985], [Sternberg2015]. The model captures ability and intelligence in a more comprehensive way than standard tests do. Creative knowledge refers to the production of new ideas (such as written short stories, the design of a product or a scientific experiment) and to coping with unconventional problems. Practical intelligence is embodied in tacit knowledge and involves shaping, selecting and adapting to different environments and situational contexts. It involves abilities that people internalized based on scenarios, and challenges they encounter in their everyday life such as work-related problems or leisure-related problems (e.g. figuring out how to make friends). Analytical intelligence in his context encompasses essentially the analysis, judgement, evaluation, contrast and comparison of information. Sternberg later introduced the idea of successful intelligence as the combination of all three components needed in order to achieve important goals in life and to succeed within a given socio-cultural context. He attaches greater weight on successful intelligence than on high scores of conventional IQ tests [Sternberg2015]. Finally, the corresponding test emerging from this approach (i.e. the Sternberg Triarchic Abilities Test) is included in the successful Intelligence part and, to a very low extent, in the creative sections of the some of the sub-tests[10] [Sternberg et al.2001].

On the other hand, let us consider representative tests for artificial general intelligence. So, starting with the Turing test, we see that this foundational AI test is structured mainly in a linguistic environment and exclude explicitly a lot of kinesthetic, musical, artistic, environmental, mathematical, visual and existential dimensions of human intelligence. Thus, the globality of our test is barely considered in the structure of the Turing test, since it is conceived only within a conversational environment. It follows from this, that the Turing test remains mainly a local test for artificial general intelligence.

Other simple and famous test for AGI is the "coffee test" proposed by Wozniak, in other words "going into an average American house and figure out how to make coffee, including identifying the coffee machine, figuring out what the buttons do, finding the coffee cabinet, etc." [Mikhaylovskiy2020, pag. 4] and [Goertzel et al.2012]. In this case, it is clear that the range of abilities evaluated are reduced to visual-spatial, kinesthetic and (minimal features involving) logic-mathematical intelligence.

There is another test due to B. Goertzel relevant in our discussion, i.e., the robot college student test. It refers to an (artificial) agent that enrolls in a university as standard student and passes all the courses and tests, and finally obtains a degree [Goertzel et al.2012]. Although this test possesses a lot of qualitative similarities with our holistic test, it is not specific enough in order to be able to make a genuine estimation of its level of globality. This results from the fact that the exact academic/artistic degree the agent should obtain in college, is not specified in the test. In fact, for most of the potential bachelor's degrees that exist, the agent would be evaluated in a relatively narrow collection of signatures and intelligence's types regarding the activities described in our test. Moreover, in most of the signatures, our agent would receive only learning and basic sections involving a specific topic, and the creative section in average would be dropped out. Even more, from a pragmatic point of view this test could require years to be accomplished. This, without taking into account the cognitive impact and the huge bias from the perspective of professors and classmates that a tested artificial agent would have in its surroundings.

[9]Nonetheless, note that the degree of short-term memory of the agent on its own does not so say much about the human component on the (artificial) intelligence of it. For example, one could say that a personal computer possesses (at some level) a perfect short-memory. However, virtually all of them (in the current state-of-the-art) are far away from being globally intelligent at human-level (according to our test. See for example the next section).

[10]This is due to the fact that, for example, the Sternberg test possesses a highly reduced kinesthetic component in comparison with lots of our sub-tests.

Let us consider now the employment test of N. Nilsson, i.e., measuring progress in A(G)I by calculating the fraction of jobs that machines are progressively fulfilling at human level [Nilsson2005]. In fact, this way of evaluating the development of AGI should be considered more like a meta-test consisting of an undefined numbers of sub-tests that can be passed by different artificial agents. So, strictly speaking it does not count as a concrete test with the level of pragmatism that we are requiring here. However, if one specifies exactly the concrete job that a clearly defined machine is doing at a human level, then this concrete form of the Nilsson proposal could be compared with our test. Again, in this case, for most of the specific jobs that we have at our disposal, the range of types of intelligence used is, in average, highly restricted.

There are several additional (human and artificial) intelligence tests that could be considered in our discussion. However, we chose the ones that, we believe, possess the most representative and global features regarding artificial global intelligence.

Now, in the following table, we summarize the main (human and artificial) intelligence tests presented above with a roughly estimate of the corresponding degree of generality (resp. locality) taking our ontology of 10 types (or materializations) of (human) intelligence as global framework.

Furthermore, we assigned the scores in the table approximately based on the following guidelines: a) we compare each of the sub-tests of our holistic test with the corresponding test and we assigned a rate between zero and one, measuring the local percentage of resemblance between both. b) we multiply each of the former local rates by the corresponding percentage of globality of the sub-test in consideration. c) we add all the quantities obtained in b) for obtaining the global percentage and we round it. d) for obtaining the local (general) percentage in the second column of the test, we subtract the global percentage obtained in c) from the 100%.

Intelligence Test	% Local Test	% Global Test
The Raven's Progressive Matrices [Raven2000]	90%	10%
Coffee Test (Wozniak) [Goertzel et al.2012]	85%	15%
Stanford-Binet Intelligence Scale	84%	16%
Wechsler Intelligence Scale [Wechsler and Kodama1949]	80%	20%
Kaufman-brief Intelligence Test [Kaufman1990]	78%	22%
Mayer-Salorevy-Caruso Emoc. Int. Test [Brackett and Salovey2006]	77%	23%
Woodcock-Johnson Test [Mather and Jaffe2016]	71%	29%
Turing Test [Turing2009]	70%	30%
Specific Forms of the Employment Test [Nilsson2005]	68%	32%
Sternberg Triarchic Abilities Test [Sternberg et al.2001]	60%	40%
Robot College Student Test [Goertzel et al.2012]	59%	49%
Holistic (Artificial) Intelligence Test	0%	100%

Estimation of Intelligence Tests by their degree of generality (resp. specificity).

5 Outstanding Qualitative and Quantitative Results

We can make now a relatively straightforward estimate of the levels of local and general intelligence of the most outstanding agents to be considered, namely, human beings and artificial agents. First, since each of the local tests is designed by the corresponding committee C_A having in mind that an average 25 years old human being should obtain at least a grade of 6, assuming that he/she has just a minimal ability to learn the very basics of A, it is expected that such a person would obtain a grade of 6 in our general intelligence test. The reason is that the final score is computed as the average of the local scores, i.e., the sum of local grades divided by the total number of tasks (74).

On the other hand, let us imagine that X is any of the most sophisticated AI devices and systems that we have nowadays, for example, IBM‘s WATSON [High2012], Wolfram Alpha [Hoy2010], Google's DeepMind [Babuschkin et al.], Honda's all-new ASIMO [Shigemi et al.2018], Boston Dynamics' PETMAN [Nelson et al.2012], and Open AI's GPT-3 [Dale2021], [Floridi and Chiriatti2020]; among others. So, if we are very optimistic regarding the performance of X, we could suppose that X would obtain a perfect grade of 10 in 5 of the tasks. Now, because of the local nature of most of these agents, we would expect local grades of 0.1 in all the tasks except by some the logical-mathematical ones (resp. at most one of the bodily-kinesthetic ones), among a few others.

One particular case deserving additional consideration can be ChatGPT [Radford et al.2019]. In fact, after using systematically the chat for several intellectual tasks, one can obverse that ChatGPT is highly efficient in some particular assignments (like rephrasing and programming), and simultaneously highly inefficient in other basic ones (solving basic conceptual mathematical problems). Additionally, it seems to make the same elementary (domain-specific) mistakes over and over again, showing along the way a very limited ability to learn in a complete and global manner. So, one cannot establish that it learns in, at least, the coherent way that a human being learns a basic skill, since it seems to repeat regularly the same simple (domain-specific) errors independently of the number of lessons/conversations that it has with the user.

So, just by doing a straightforward arithmetical estimation, we conclude that the grade of X (excepting ChatGPT) regarding general intelligence would be at most of around 1.0/10.0 On the other hand, due to the former remarks one can give to ChatGPT approximately a 4.0/10.0, assuming, in general, a very poor performance in creative sections of most of the local sub-tests, as well as a poor/middle-poor performance in the learning and basic sections. For example, for almost all the activities belonging to the bodily-kinesthetic, musical-rhythmic and harmonic, visual-spatial, intra-personal, and existential intelligences the performance of ChatGPT is essentially zero due to the static question-answer structure of it.

Moreover, let us assume that we want to expand our former global test by adding extra tasks coming from the currently most outstanding intelligence tests. Now, since most of this tests have a predominant local nature in their design, we could put for instance 7 new tasks consisting of entire particular tests or specific parts of them, and we re-scale the results in order to obtain outputs between 0.1 and 10. So, assuming again that our artificial agent X is able to obtain the same expected results on the initial 74 tasks and an average score of 6 in each of these new tasks, its score in the expanded version of the test with 81 tasks would be approximately 1.3-1.4. In any case, the score obtained would be very low from a global point of view.

6 Open Discussion and General Remarks

We have lots of perfect local agents, namely, agents having perfect local grades in several tasks. However, our best current agents seems to have very low grades as general agents, i.e., as agents performing general (global) intelligence. So, a plausible hypothesis to be subsequently tested in more concrete experimental scenarios and emerging as a very natural consequence of the whole collection of former facts presented here is that at the beginning of the second decade of the twentieth-one century we seem to be located in the very beginning of plausible materializations of artificial general (global) intelligence at a human level. In particular, the former hypothesis would implies that the question concerning an (hypothetical) intelligence explosion, leading to a kind of 'singularity' [Chalmers2010], seems to have a more metaphysical speculative nature in our current state of technological development, than a well-grounded and scientific status. On the other hand, regarding weak or local artificial intelligence, we seem to be located in some epoch of the 'golden age' with respect to some specific, but not all, tasks. For example, regarding mathematical formal research, we are currently advancing quickly towards the fulfillment of artificial mathematical intelligence [Gomez-Ramirez2020].

At this point, we would like to open an objective and enhanced discussion about the realistic and constructive perspectives of the present and the future of AGI, going beyond any political, military or financial interest and focussing on the scientific, technical and ethical issues of this matter. So, in this context, we would like to offer our holistic test as a methodological impartial starting point.

This distinction between the genuine state of the art of global and local intelligence is something that should be clarified more deeply in the literature, and, thus, our holistic (AGI) test represent a useful conceptual initial tool to clarify and crystallize this issue structurally.

Furthermore, one of the seminal and transparent hypothesis presented here is to give an objective and more robust answer to the fuzzy question regarding the real and contextualized state of the art of general (strong) and local (weak) intelligence in terms of a holistic and mature conceptualization of intelligence (MI theory), grounded by lots of empirical studies in psychology, neuroscience, anthropology and cultural studies.

Additionally, a very compelling and promising manner of enriching the present test for taking into account the inner (phenomenological) dimension of intelligence is to develop an additional collection of tasks which can measure if an (artificial) agent can have (resp. simulate) a functional, creative. active and intelligent consciousness, being able of processing information, imagining abstract and fantastic scenarios and answering wisely and astutely to a wide range of questions, among many other things.

Even more, with enough resources at disposition, it seems to be highly plausible to materialize a particular form of this test in a concrete physical place, with a concrete group of juries, and using a concrete and robust sample of artificial and natural agents for performing the test. This would give a lot of enlightening information about all the crucial issues discussed here so far.

So, we strongly believe that an open, deep and multidisciplinary academic discussion on this matter would be very useful regarding the future of the research on cognitive (human) and artificial intelligence and its constructive impact on mankind. The important point here is that a wide spectrum of scientists, entrepreneurs and leaders can be involved on the discussion and that the main goal can be a global, peaceful and impartial benefit for the whole society. An inspiring example of this kind of initiatives is the Future of Life Institute[11], which have made a tremendous effort for the improvement of the ethical principles that should guide any potential materialization of local and global artificial intelligence.

References

[imo] International mathematical competition. *https://www.imo-official.org/?language=en*.

[Abu-Hamour et al.2012] Abu-Hamour, B., Al Hmouz, H., Mattar, J., and Muhaidat, M. (2012). The use of woodcock-johnson tests for identifying students with special needs-a comprehensive literature review. *Procedia-Social and Behavioral Sciences*, 47:665–673.

[Albrecht2006] Albrecht, K. (2006). *Social intelligence: The new science of success*. John Wiley & Sons.

[Armstrong2009] Armstrong, T. (2009). *Multiple intelligences in the classroom*. Ascd.

[Babuschkin et al.] Babuschkin, I., Baumli, K., Bell, A., Bhupatiraju, S., Bruce, J., Buchlovsky, P., Budden, D., Cai, T., Clark, A., Danihelka, I., et al. The deepmind jax ecosystem (2020). *URL http://github. com/deepmind*.

[Brackett and Salovey2006] Brackett, M. A. and Salovey, P. (2006). Measuring emotional intelligence with the mayer-salovery-caruso emotional intelligence test (msceit). *Psicothema*, 18:34–41.

[Chalmers2010] Chalmers, D. J. (2010). The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies*, 17:7–65.

[Chen et al.2009] Chen, J.-Q., Moran, S., and Gardner, H. (2009). *Multiple intelligences around the world*. John Wiley & Sons.

[Choudhury and El-Nasr2019] Choudhury, P. K. and El-Nasr, M. A. (2019). Insight into remote monitoring systems. *Journal of Electromagnetic Waves and Applications*, 33(7):795–798.

[Dale2021] Dale, R. (2021). Gpt-3: What's it good for? *Natural Language Engineering*, 27(1):113–118.

[Daniel1997] Daniel, M. H. (1997). Intelligence testing: Status and trends. *American psychologist*, 52(10):1038.

[Descartes1954] Descartes, R. (1954). *The Geometry of Rene Descartes: with a facsimile of the first Edition*. Dover Publications, Inc.

[Fauconnier and Turner2003] Fauconnier, G. and Turner, M. (2003). *The Way We Think*. Basic Books.

[Ferrater Mora1958] Ferrater Mora, J. (1958). *Diccionario de Filosofía*. Sudamericana, Buenos Aires.

[Floridi and Chiriatti2020] Floridi, L. and Chiriatti, M. (2020). Gpt-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4):681–694.

[Gardner and Moran2006] Gardner, H. and Moran, S. (2006). The science of multiple intelligences theory: A response to lynn waterhouse. *Educational psychologist*, 41(4):227–232.

[Gartner2011] Gartner, H. (2011). *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books, New York.

[Gartner1995] Gartner, M. (1995). Reflections on multiple intelligences: Myths and messages. *Phi Delta Kappan*, 77:200–209. `https://learnweb.harvard.edu/WIDE/courses/files/Reflections.pdf`.

[Gaw2016] Gaw, D. C. (2016). Remote sensing device, system, and method utilizing smartphone hardware components. US Patent 9,401,977.

[Ginzburg et al.1966] Ginzburg, S. A., Lekhtman, I. Y., and Malov, V. S. (1966). *Fundamentals of automation and remote control*. Oxford.

[Goertzel et al.2012] Goertzel, B., Iklé, M., and Wigmore, J. (2012). The architecture of human-like general intelligence. In *Theoretical foundations of artificial general intelligence*, pages 123–144. Springer.

[Gomez-Ramirez2020] Gomez-Ramirez, D. A. J. (2020). *Artificial Mathematical Intelligence: Cognitive, (Meta)mathemticial, Physical and Philosophical Foundations*. Springer International Publishing. ISBN 978-3-030-50272-0.

[11]https://futureoflife.org

[Gust et al.] Gust, H., Krumnack, U., Kühnberger, K.-U., and Schwering, A. Analogical reasoning: A core of cognition. *KI-2008*.

[Harper et al.2001] Harper, D. et al. (2001). Online etymology dictionary.

[Hays et al.2002] Hays, J. R., Reas, D. L., and Shaw, J. B. (2002). Concurrent validity of the wechsler abbreviated scale of intelligence and the kaufman brief intelligence test among psychiatric inpatients. *Psychological reports*, 90(2):355–359.

[High2012] High, R. (2012). The era of cognitive systems: An inside look at ibm watson and how it works. *IBM Corporation, Redbooks*, 1:16.

[Hoy2010] Hoy, M. B. (2010). Wolfphram— alpha: a brief introduction. *Medical reference services quarterly*, 29(1):67–74.

[Kaufman1990] Kaufman, A. S. (1990). *Kaufman brief intelligence test: KBIT*. AGS, American Guidance Service Circle Pines, MN.

[Kurzweil2006] Kurzweil, R. (2006). *The Singularity Is Near: When Humans Transcend Biology*. Penguin (Non-Classics).

[Lakoff and Johnson1980] Lakoff, G. and Johnson, M. (1980). *Metaphors We live By*. Chicago University Press.

[Li et al.2018] Li, L., Lin, Y.-L., Zheng, N.-N., Wang, F.-Y., Liu, Y., Cao, D., Wang, K., and Huang, W.-L. (2018). Artificial intelligence test: a case study of intelligent vehicles. *Artificial Intelligence Review*, 50(3):441–465.

[Lieto2021] Lieto, A. (2021). *Cognitive design for artificial minds*. Routledge.

[Levi-Strauss2008] Levi-Strauss, C. (2008). *Structural Anthropology (Revised Edition)*. Basic Books.

[Mamalis2007] Mamalis, A. (2007). Recent advances in nanotechnology. *Journal of Materials Processing Technology*, 181(1-3):52–58.

[Mather and Jaffe2016] Mather, N. and Jaffe, L. E. (2016). *Woodcock-Johnson IV: Reports, recommendations, and strategies*. John Wiley & Sons.

[Mayer and Salovey1993] Mayer, J. D. and Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence*, 17(4):433–442.

[Mazur and Stein] Mazur, B. and Stein, W. *Prime Numbers and the Riemann Hypothesis*. `http://wstein.org/rh/rh.pdf`.

[McCorduck2009] McCorduck, P. (2009). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. AK Peters/CRC Press.

[Mikhaylovskiy2020] Mikhaylovskiy, N. (2020). How do you test the strength of ai? In *International Conference on Artificial General Intelligence*, pages 257–266. Springer.

[Moran and Gardner2006] Moran, S. and Gardner, H. (2006). Extraordinary achievements: A developmental and systems analysis.

[Nelson et al.2012] Nelson, G., Saunders, A., Neville, N., Swilling, B., Bondaryk, J., Billings, D., Lee, C., Playter, R., and Raibert, M. (2012). Petman: A humanoid robot for testing chemical protective clothing. *Journal of the Robotics Society of Japan*, 30(4):372–377.

[Newell and Simon1975] Newell, A. and Simon, H. A. (1975). Computer science as empirical inquiry: Symbols and search. *Philosophy of Psychology*, page 407.

[Nilsson2005] Nilsson, N. J. (2005). Human-level artificial intelligence? be serious! *AI magazine*, 26(4):68–68.

[Pujol et al.1993] Pujol, J., Vendrell, P., Junqué, C., Martí-Vilalta, J. L., and Capdevila, A. (1993). When does human brain development end? evidence of corpus callosum growth up to adulthood. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 34(1):71–75.

[Radford et al.2019] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., et al. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.

[Raven2000] Raven, J. (2000). The raven's progressive matrices: change and stability over culture and time. *Cognitive psychology*, 41(1):1–48.

[Russell and Norvig2016] Russell, S. J. and Norvig, P. (2016). *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited,.

[Schmidt et al.2014] Schmidt, M., Krumnack, U., Gust, H., and Kühnberger, K.-U. (2014). Heuristic-driven theory projection: An overview. In Prade, H. and Richard, G., editors, *Computational Approaches to Analogical Reasoning: Current Trends*, pages 163–194. Springer, Berlin, Heidelberg.

[Shigemi et al.2018] Shigemi, S., Goswami, A., and Vadakkepat, P. (2018). Asimo and humanoid robot research at honda. *Humanoid robotics: A reference*, pages 55–90.

[Sternberg2015] Sternberg, R. J. (2015). Successful intelligence: A model for testing intelligence beyond iq tests. *European Journal of Education and Psychology*, 8(2):76–84.

[Sternberg et al.2001] Sternberg, R. J., Castejón, J., Prieto, M., Hautamäki, J., and Grigorenko, E. L. (2001). Confirmatory factor analysis of the sternberg triarchic abilities test in three international samples: An empirical test of the triarchic theory of intelligence. *European Journal of Psychological Assessment*, 17(1):1.

[Sternberg et al.1985] Sternberg, R. J. et al. (1985). *Beyond IQ: A triarchic theory of human intelligence*. CUP Archive.

[Tononi2012] Tononi, G. (2012). The integrated information theory of consciousness: an updated account. *Archives italiennes de biologie*, 150(2/3):56–90.

[Tononi et al.2016] Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience*, 17(7):450–461.

[Turing2009] Turing, A. M. (2009). Computing machinery and intelligence. In *Parsing the Turing Test*, pages 23–65. Springer.

[Valcárcel et al.2008] Valcárcel, M., Simonet, B., and Cárdenas, S. (2008). Analytical nanoscience and nanotechnology today and tomorrow. *Analytical and bioanalytical chemistry*, 391(5):1881–1887.

[Wechsler and Kodama1949] Wechsler, D. and Kodama, H. (1949). *Wechsler intelligence scale for children*, volume 1. Psychological corporation New York.

[Wiles1995] Wiles, A. (1995). Modular elliptic curves and fermat last theorem. *Annals of Mathematics*, 142:443–551.

[Zhang et al.2009] Zhang, X., Guo, Q., and Cui, D. (2009). Recent advances in nanotechnology applied to biosensors. *Sensors*, 9(2):1033–1053.